Toward Shared System Knowledge

An empirical study of knowledge sharing policy and practice in Systems Engineering Research in the UK

By

Paola Di Maio

A thesis submitted to the University of Strathclyde for the degree of

Doctor of Philosophy

DMEM
Department of Design Manufacture and Engineering Management
University of Strathclyde
Glasgow, Scotland, UK.

November 2011

DECLARATION OF AUTHENTICITY AND AUTHOR'S RIGHTS

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:____

Date:_________________________

ABSTRACT

Research in Open Access (OA) to Scholarly Publications has flourished in recent years, however studies published to date tend to be quantitative, statistical analyses over undifferentiated corpuses, that monitor the overall uptake (Björk et al. 2010; Laakso et al. 2011). This doctoral thesis explores a different path of inquiry: it examines the effectiveness of OA policies in relation to the perspective of a 'knowledge seeker' and considers them in the context of the wider regulatory landscape that motivates their existence, specifically monitoring the availability of shared resources – journal publications, as well as other knowledge sharing artefacts adopted in technical domains – in relation to systems engineering research in the UK. Research Funding Councils adopt Open Access policies and display them prominently on their website, yet not all funded research projects seem to share knowledge by publishing Open Access resources. The main hypothesis driving this thesis is that a gap exists between Open Access in theory and Open Access in practice. A unique research methodology is devised that combines evidence based research (EBR) with a wide range of mixed method techniques, including FOI (freedom of information) requests. A novel collection instrument, a set of heuristic indicators, are developed to support the empirical observation of the gap between 'Open Access policies in theory', corresponding approximately to what the funding body state on their website, and 'Open Access policies in practice', corresponding to the level of adoption of these policies by grant holders. A systematic review and a meta-analysis of a 100 publicly-funded projects are carried out. The research demonstrates empirically that in the majority of the audited publicly-funded projects, no Open Access resources can be located.

ACKNOWLEDGEMENTS

This doctoral research has been made possible thanks to the knowledge and expertise generously shared by practitioners and communities of practice, especially in the fields of Open Access, Knowledge Sharing, Web Technologies, Systems Science and Knowledge Representation to which the author is profoundly indebted to. Thanks to Alex Duffy, for securing funding and academic freedom, and to Iain Boyle, for ongoing critique on drafts, and thanks to Jonathan Corney and Umit Bitcici of DMEM for additional suggestions and tips. Thanks to Google for making it possible to locate all the necessary documents and supporting the research with a reliable search algorithm and a valuable suite of online applications.

Thanks to the Freedom of Information Act for making it possible to get at least some straight answers, and to all the civil servants and research officers who patiently and gracefully have sustained my relentless inquisitions and repeated requests. Innumerable individuals have contributed usefully to various aspects of the research and thesis development on an ongoing basis, by answering questions, providing pointers, or peers reviewing drafts, especially in those aspects of the research where limited expertise and knowledge was otherwise available, listed below in no particular order, and apologies for any omissions.

I take this opportunity to also thank the lecturers who first introduced me to Knowledge Systems and Systems Engineering when I first started studying them academically over a decade ago, whose teachings are vivid in my mind and still central to my work today (in particular Nancy Johnson, in memory, and Lester Gilbert, now at Southampton University), as well a big thank you to all elders and teachers who selflessly and unconditionally share knowledge openly and generously.

Thanks also to anonymous reviewers, the independent auditors who helped to collect data, to the copy-editor for assisting with pagination and formatting of the thesis, and to friends and family for encouragement and support.

Donna Rhodes, MIT

Fran Ackermann, University of Strathclyde

Milena Dobreva, University of Portsmouth

Simon Buckingham Shum, Open University

Jim Wretham, Head of Information Policy UK

Peter Burnhill, EDINA

Philip Davies, Oxev Consulting

Eugenia Vassillakaki, University of Manchester

Ken Turner, University of Stirling

Robert Raeside, Napier University

Angus Whyte, University of Edinburgh

"There is only one good, knowledge, and one evil, ignorance"

Socrates

FOREWORD

This doctoral dissertation results from three years of socio-technical investigation, and builds on previous and ongoing related work. It is partly motivated by personal observations and by first-hand experience of some rather paradoxical situations that led to the formal inquiry reported herein. The breadth and scope of work undertaken required an adaptive, hybrid methodology that could make use of accumulating and emerging evidence to evolve the research design (Berry, 2010). Personal experiences and previous research can contribute, among other factors, to influence decisions regarding research topics and structure (Alter & Dennis, 2002). This is particularly true in the qualitative and interpretative portions of the study (Creswell, 2007) in qualitative research where *"writers can bring themselves into the study, the personal pronoun 'I' can be used, and even engage in storytelling form of narration"* (Joubish, 2011). The resulting research design is interdisciplinary, integrating information, data, techniques, tools, and perspectives from more than one discipline (Committee for Interdisciplinary Research, 2005), which evolved into a 'critical inquiry' (Packer, 2010), carried out in the context of an evidence based research (Fitzallen & Brown, 2007). Key motivations and personal experiences are reported as first-person accounts in the exploratory and reflective portions of the thesis, alongside a more formal technical report of the study. The core inferences and main conclusions however, mainly to facilitate evaluation, are constrained within the boundaries of positivist hypothesis testing (PHT) approach.

> *"While an empirical-analytic paradigm presumes that researchers should have an objective stance (neutral and detached), and that research advances through hypothesis formulation and testing, while interpretive paradigm assumes that the researcher should have a participatory stance,*

and that research requires the description of specific cases (persons and communities), through narrative articulation and interpretation, the critical paradigm asserts that the researcher should have a critical stance, alternately participating and objectifying, that research involves both measurement and narrative, and that research leads to suggestions for action." (Packer, 2010)

Snippets of correspondence and private exchanges mentioned in this thesis are pasted in anonymous form to protect the privacy of the informants, however copies are held in a personal research log for documentation purposes, and for future work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

"We are not students of some subject matter, but students of problems. And problems may cut right across the borders of any subject matter or discipline."

Karl Popper

CHAPTER 1

INTRODUCTION

Anecdotal evidence and a personal, interpretive perspective provide the initial background and motivation for the research questions, and justify the overall direction and key research design choices. This chapter includes an overview of the thesis structure, together with general orientation such as a basic working glossary and a concept map.

1.1 BACKGROUND

My ‘Journey into Knowledge’ started early in life. One of the few certainties I had as a teenager was that I wanted to acquire as much knowledge and understanding as possible about the world and everything in it. The pursuit of ‘all knowledge’ could seem futile and troublesome.

I learned something in school, but conventional ‘average’ schools tend to be indoctrinating. Intellectual refinement included acquiring as much knowledge as possible, from whatever means (including comic books and the dreaded television) and by comparing knowledge sources for different points of view, including empirical observations, to verify that what is passed down as ‘knowledge’ actually corresponds to fact. That's a lesson I learned early.

Inevitably some knowledge leads to more questions, and that's how the never-ending quest of this knowledge seeker began. Nonaka's spiral, as per Fig. 1 (Nonaka & Takeuchi, 1995) is perhaps the most commonly used metaphor that captures the cyclic and recursive nature of a knowledge acquisition and creation process, although it is generally used to represent an organisational knowledge transfer as a flow, it can also serve as a generic metaphor to illustrate the endlessness of knowledge cycles.

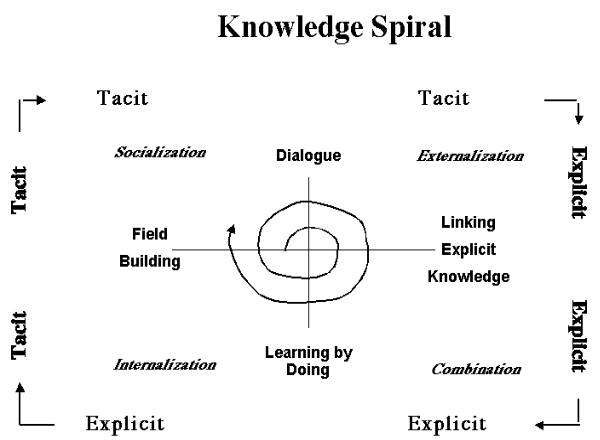


Fig. 1: Knowledge Spiral (Nonaka, 1995)

At least three specific events contributed to shape the trajectory of this researcher's journey into knowledge, and constitute much of the 'wind behind the sails' of this doctoral thesis: a lecture in systems theory, the arrival of the internet, awareness of the importance of access to scholarly knowledge, shared briefly below.

1.1.1 A Lecture on Systems Theory

While attending a course in the Faculty of Law and Jurisprudence, UNAM (Mexico City, 1992) in 'International relations', I learned how systems theory (Bertalanffy, 1968) can help to frame issues in the broadest possible context, grasp complex problems, get the big picture and think systemically. The cognitive approach that leverages the principles of systems theory is known as 'systems thinking' or 'systemics' (Umpleby, 2001). More recently the notion of 'systems being' has emerged from system research (Lazlo, 2011). I then went on to study 'information systems' (Master of Science in IT), and eventually found an interesting niche in socio-technical systems, which constitute both the epistemic perspective of my research, as well as my primary field of practice. Now everything is either 'a system' or 'part of a system'. Caveat lector.

1.1.2 The Arrival of the Internet

A few months after the systems theory lecture, I learned about a new public system - at the time it was relatively new - called 'the internet'. I later was informed that appropriate terminology applies, and that the internet and the web are two different layers of the public data exchange network, sometimes used (incorrectly) interchangeably or as synonyms. This 'new' system was about to change the life on the planet for good, and became my subject of study, my place of work, my research topic. The web is the knowledge environment *par excellence*.

Every object, information, whether physical or intangible, has a knowledge dimension attached to it, and the web serves as the most dynamic and pervasive environment for knowledge (about everything) to become accessible and co-created (Collis, 2008). In this doctoral thesis the internet and the web are considered a de-facto optimal knowledge exchange environment, therefore all the axioms, assumptions conventions and constructs that pertain to 'web science'[1], apply.

1.1.3 Access to Scholarly Knowledge

When I became employed full-time by an international university[2], which subscribed to *all* the scientific electronic libraries and journals services in the world, and learned what a difference access to scholarly knowledge can make; accessing scholarly outputs (in plain words, research papers published by academic journals) made it possible for me to learn the 'state of the art' on any given subject, and to make novel contributions accordingly.

Thanks to a limited teaching load, and a full-time lectureship post, I spent most of my non-teaching hours between 2005 and 2008 finding, and avidly reading, as much original research materials as possible: ACM, IEEE, SpringerLink and Emerald.

Thanks, among other factors, also to access by subscription of the research institution I was employed by, to vast amount of sources of scholarly knowledge, I could finally became a scholar myself, and published my first academic paper (Di Maio, 2006).

1 http://webscience.org
2 http://www.mfu.ac.uk

Recursively, perhaps, in one of those funny loops of life, the object of my research is knowledge itself, in particular, knowledge modelling and representation on the web, which can be considered the largest knowledge base and the largest socio-technical system that ever existed. Article after article, all the secure *PDFs* that until then had been beyond my reach, would open, a bit like entering Aladdin's Cave by pronouncing 'Open Sesame'.

In turn, access to knowledge opened new partitions in my cognitive and intellectual configuration, and led me to pursue a doctorate, the structure and main outcomes of which are summarized in this thesis. This research would not be possible - or would be radically different - without an institutional account, currently Shibboleth (not long ago it was Athens). In the future, myself and others hope that every individual online will be able to access scholarly knowledge without necessarily having to be employed by an academic institution, especially when the knowledge and papers in question are generated by research paid for with taxpayers or other public money. This research contributes directly to the realisation of that vision. Watch this space.

1.2 THESIS STRUCTURE

Chapter 1 of this thesis consists of an introduction, which starts with a personal narrative account of some background experiences pertinent to the research, and provides an outline and basic orientation to the rest of the thesis. Chapter 2 starts with key 'observations' that together with the introductory notes of Chapter 1, constitute an interpretive basis that justifies and in part motivates the line of work undertaken, and which contributes to formulate the research questions and help to shape methodology and the research challenges that follow.

In Chapter 3 essential literature is reviewed, from the variety of disciplinary perspectives that constitute the 'corollary' for the research and some gaps in literature are identified. Although the core of the literature review is presented in Chapter 2, the entire thesis is underscored by scholarly references throughout, and some of the empirical findings presented in RC1 are obtained by triangulating information obtained from a review of pertinent documentation and sources, appropriately referenced, with other sources such as FOI requests.

Chapter 4 presents and discusses the collection instruments, and in Chapter 5 the two main research components, RC1 and RC2, are presented together with an analysis of the main findings. Chapter 6 contains quality and evaluation for the study, and Chapter 7 concludes with recommendations, discussion and final remarks. A complete set of references is provided, together with supporting documentation in the corresponding appendices for each section.

RESEARCH COMPONENT	PARADIGM /APPROACH	TECHNIQUE	OUTCOME
Observations	Interpretive	First person narrative	Motivation and Justification for further research
Survey Policies	Exploratory	Survey/ FOI	Independent Variable
Do people know about these policies?	Evaluative	Ethnographic Observation	Extent of policy impact
Field work	Quantitative (PHT)	Audits	Statistically Significant Research Data
Guidelines	Reflective	Model building Good practices integration	Conclusion

Table 1: Summary Overview of Main Research Components

1.2.1 Concept Map

The field of enquiry against which this thesis develops is vast, and highly interdisciplinary, somewhat conflicting with a stereotype view that doctoral research must have a narrow focus. To help visualise the broad scope of this enquiry, a concept map is provided, Fig. 2, where the central themes tackled in the chapters that follow are graphically represented. The arrows indicate a level of recursiveness and circularity between the two main themes: knowledge and systems. The relation between knowledge and system science deserves to be explored in full, falling under the broader heading 'systemic knowledge', defined as "the holistic understanding of interpersonal expectations or norms, the technical system, and the relationships between the two" (Sheffield, 1995). The scope of this research is constrained by limited resources and a looming deadline. For practical purposes, the rest of this thesis is narrowed to answer the research questions, as explained in Chapter 2.

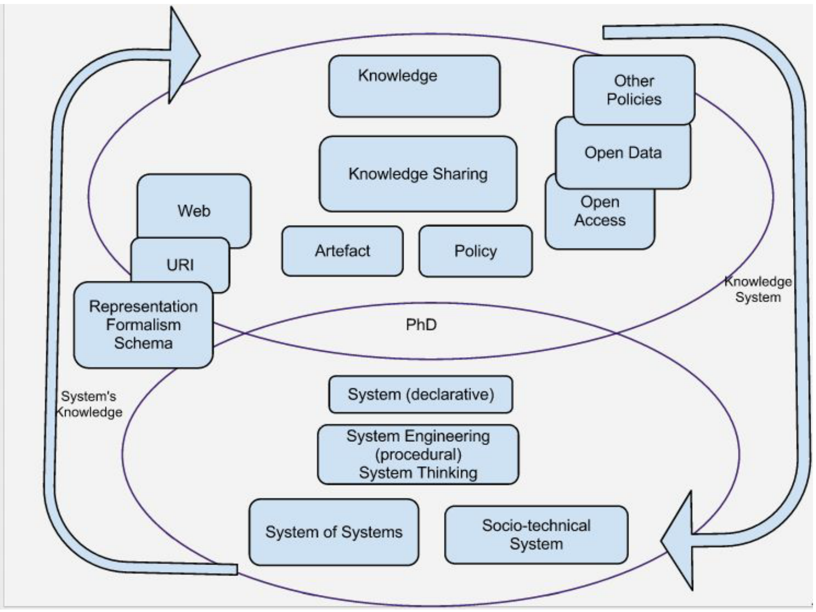


Fig. 2: PhD Thesis Concept Map

1.3 PREVIOUS AND RELATED WORK

Not much happens in a vacuum. This thesis isolates and tackles a specific set of challenges and issues that relate to knowledge sharing and re-use of systems knowledge, but it is a 'snapshot' of a lifelong inquiry, equivalent perhaps to what a single photogram is to a movie. The foreword introduces glimpses into personal motivations. This section provides a brief background into the research published in previous years, to which parts of this thesis are a logical sequel.

Just over ten years ago – but it feels like a much longer in internet time – the author of this doctoral thesis chose Knowledge Based System (KBS) – a type of expert systems and a disciplinary branch of Artificial Intelligence (AI) – as the elective for an MSc in IT, learning the ropes of 'Knowledge Engineering':

> *"KE is an engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise."* (Feigenbaum & McCorduck, 1983)

An expert system was developed using Object Oriented modelling and technology, and Common KADs[3] as the knowledge modelling methodology. Core Principles of knowledge representation and modelling acquired during the development of expert systems are used in this research to devise a sample of shared knowledge model of systems engineering proposed in the final recommendations.

---

[3] http://www.commonkads.uva.nl/

An FOAF based model of expertise, FoafX[4] (Di Maio, 2007) leveraged a similar modelling approach. In 2009 an agile methodology for knowledge modelling called 'JEOE', recently published in ACM Proceedings (Di Maio, 2011).

Combining an interest in knowledge representation, collaboration and emergency management, due to a personal involvement in the 2004 Tsunami in South East Asia, resulted in the first special track ever held in Ontology for Crisis Management (Washington DC, ISCRAM, 2008), where the case for open shared ontologies in support of networked capabilities research was made (Di Maio, 2008). The research above has provided background knowledge relevant to this thesis. A list of publications and presentations related to the research presented in this thesis is provided in Appendix A.

1.4 GOALS OF THE RESEARCH

This research identifies and addresses a 'pragmatic gap' that exists between knowledge sharing 'in theory', corresponding roughly to an exploration of the policies devised to support access to scholarly research outputs, and the practice, corresponding to an audit of a comprehensive sample of publicly-funded research projects. To relate these two otherwise disconnected areas of inquiry, it is necessary to devise a novel methodological approach and collection instruments.

The intended goal of the research is therefore the discovery of a new pragmatic dimension, that emerges from the combination of (at least) two important fields of enquiry: policy/governance, in this study specifically knowledge sharing policies, and web science, specifically the adoption of the technical artefacts (such as URIs) designed to make knowledge sharing and openness feasible and inexpensive to implement.

[4] http://itschool.mfu.ac.th/expertfinder/

Figure 3 illustrates the 'future research' space that emerges when visualising the pragmatic gap which is the object of this study, which finds some correspondence in the Japanese metaphors 'BA' (Nishida 1990; Shimitzu 1995; Nonaka & Konno 1998) and 'MA' (Medeni et al., 2008) defined respectively as:

Ba (場): *"creative and collaborative space that is Real/Physical, Digital/Virtual, and Mental/Intellectual at the same time"* (Von Krogh et al., 2001).

Ma (間): *"a spacio-temporal interval, in-between-ness that both separates and connects as a permeable membrane for creative and collaborative interaction"* (Song & Kondou, 2006).

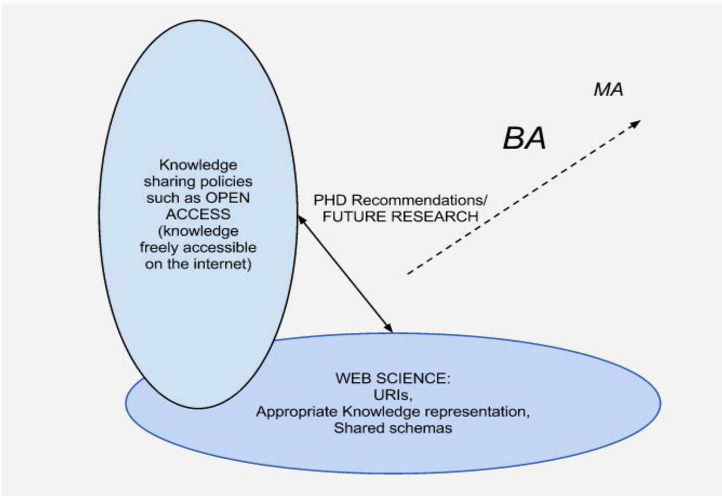


Fig. 3: PhD Recommendations/Future research

1.5 WORKING DEFINITIONS

Some key terms listed below are defined loosely, primarily for reference, however their usage in the body of the thesis should be considered in the context of the narrative.

Entanglement: co-dependence between two or more factors

Evidence Based Research: evidence gathered through scientific method using a multiplicity of research techniques, carried out for the specific intent to inform and influence policy making

Interdisciplinary: approach and means to enable the realisation of successful systems (INCOSE)

Knowledge: a mental process, a high level cognitive function, and an objective resource resulting from learning

Knowledge Artefact: cognitive and/or technical device for codified knowledge

Knowledge Asset: repository of intellectual capital access to which is limited or restricted

Knowledge Resource: repository of intellectual capital which is widely accessible, although some restrictions can apply

Knowledge Sharing: the process of communication in which members exchange and discuss knowledge through discussion and internet to increase the knowledge value (Hendriks, 1999)

Knowledge Sharing Policies: regulatory instruments devised to optimise knowledge flows and access to knowledge

Mandate: a policy or other organisational instrument

Mixed Method Research: a research methodology undertaken under more than one paradigm of inquiry, and that adopts a combination of research techniques

Open Access: a movement, and a series of initiatives, including open access policies (mandates) devised to support knowledge sharing

Paradigm of Enquiry: a logical intellectual construct that depends on a given epistemological stance

Shared Knowledge: explicit, codified knowledge access to which is not conditional or restricted

Socio-Technical System: system whose boundaries include social, technical and environmental dimensions

System: a functional whole, either existing in nature, or engineered by humans

Systemic Knowledge: the holistic understanding of interpersonal expectations or norms, the technical system, and the relationships between the two

"And those who were seen dancing

were thought to be insane

by those who could not hear the music."

Friedrich Nietzsche

CHAPTER 2

MOTIVATION, RESEARCH PLAN, METHODOLOGY

OVERVIEW OF THE SECTION

This section provides an insight into personal experiences that motivate and justify the research questions and methodology adopted to answer them. 'Observations' are a first-person narrative of events that took place at the beginning of the doctoral research programme, when seeking to review project knowledge in relation to 'knowledge re-use and learning in network capabilities research', as part of an EPSRC (Engineering and Physical Sciences Research Council) funded programme, and of exchanges with peers in the context of various systems engineering research projects, and the reflections that emerged. These observations do not contribute directly to the findings, results and final recommendations, but did help to shape the research hypothesis and questions, and to justify the methodological approach, detailed in the subsequent paragraphs. This chapter concludes with ethical considerations.

2.1 OBSERVATIONS

The work carried out in the years and months preceding the start of this doctoral research contributed to the shaping of goals and expectations in relation to systems knowledge and its re-use. The opportunity to make contextualised observations in relation to knowledge sharing and re-use of engineering knowledge materialised at the very beginning of the doctoral programme, encapsulated in the three examples illustrated below.

2.1.1. Aserp

In 2008 an EPSRC Doctoral Training Account (DTA) was made available through the research program Aserp (A Systems Engineering Research Project) to pursue research in 'Knowledge Re-use and Learning ' in the context of Networked Capabilities research. The first question that needed to be answered was: what knowledge had been created by the Aserp project that should be used as the starting point for the research ahead? The academics in charge of administering the DTA, suggested contacting the Aserp project leader, another academic in a university part of the same research consortium, who in turn suggested contacting the relevant project officer at ABIP (A British Research Partner), who suggested a formal request to access such information had to be made in writing.

Not much knowledge or information was available on the project website other than a static list of publications. It took several email exchanges with Aserp project officers to learn that despite the project being funded with public money, there was no project documentation available in the public domain. Despite the substantial public budget for Aserp and the participation of high profile research institutions including top universities, it appeared that:

- Terminology and concepts were not explicitly defined, nor agreed by project partners (this was gathered from email or face to face exchanges with some of the project participants).
- The project documentation was not publicly available.
- It was not clear (or at least it was not public knowledge) who was making project decisions, especially in relation to knowledge sharing, according to what project policy/brief.

The underlying, endemic problem in relation to sharing systems engineering knowledge tackled by this research at large is well illustrated by the Aserp example. As a DTA holder tasked with advancing the state of the art in 'knowledge re-use and learning in networked capabilities research for systems engineering' and receiving doctoral research funding from EPSRC public funding, it was essential for this researcher to acquire and examine existing project knowledge before the state of the art could be advanced. Despite the fact that the project was publicly-funded by EPSRC, which has an 'Open Access Policy', access to project knowledge seemed to be constrained by contractual obligations originating from some agreements with the industry partner ABIP, a private company which operates a policy of strict knowledge control.

ABIP did not share or publish any system diagrams, or vocabularies or data dictionaries in relation to Aserp, and the academic research partners had to ask for permission to ABIP before any decision in relation to knowledge sharing could be taken.[5]

Although some of the papers listed on the static project pages could be retrieved from scholarly repositories via web searches, they did not contain structured, technical knowledge that could be re-used to carry out engineering tasks. In other words, no structured technical project knowledge was openly accessible.[6]

An endless sequence of emails to obtain access to the knowledge artefacts related to the project between the doctoral researcher and entire hierarchies of academics and individuals working for the defence industry generated no outcomes - the 'target knowledge' was never obtained, and emails ended up being unanswered. Some of these are kept as documentation in a research log.

Outcome of the Observation: Despite an 'Open Access Policy' published on the funding body website, the list of publications on the project was 'static', that is, no hyperlink was accessible from the project home page to redirect (dereference) the knowledge seeker to papers or other knowledge repository. No obvious Open Access project resource was accessible on the project website, and no technical 'project knowledge' resources such as diagrams and drawings were accessible either via the project website or via web searches. This prompted one of the questions that has been driving much of the inquiry for the rest of this research:

5 various items of private correspondence

6 After the issue was raised in a scholarly article this has changed, as new links and resources have been added to the website.

if this research is publicly-funded via EPSRC, which like all other UK Research Councils embraces 'Open Access' policies, why can papers, and technical knowledge resources not be accessed freely over the internet?

2.1.2 Exploratory Conversations

A number of informal exploratory exchanges took place in 2009 with researchers in engineering faculties and research centres. These are recorded as 'conversations', and their purpose was to provide a sense of direction and input into understanding challenges related to knowledge sharing, therefore to help structuring the research plan. Some of the questions were derived from a pilot survey (Appendix B). Additional observations were carried out using elements of ethnography, the outcomes of which are described in a later section of the thesis.

Outcome of the Observation: The key behaviours observed relevant to the research revealed a lack of use of consistent terminology, lack of understanding of 'what is knowledge', no awareness of knowledge modelling principles, no awareness or understanding of Open Access policies and principles.

2.1.3 A University Research Centre (AURC)

A University Research Centre (AURC) is a pseudonym used to preserve the anonymity of one of the AURCs under observation. The purpose of the observation was to understand what provision and mechanisms were being developed by the centre to support knowledge exchanges. Two people were informally approached, one was the researcher directly responsible for devising and implementing a knowledge management system to facilitate different levels of access to knowledge to different 'research customers', and the other was a director.

The director explained how the business plan leverages different levels of 'subscriptions', and a multi-tier knowledge management facility was being developed to provide different levels of access to knowledge based on the subscription level of the customers. Both the director and the researcher denied knowledge of public funding coming into the Research Centre.

However thanks to an FOI request, it emerged otherwise: that there is substantial EPSRC funding involved in the activities of the Research Centre. After the FOI request was issued, the Research Centre started producing an email newsletter, which contains mainly publicity and advertising for the Centre itself, but it does not contain any research outcomes and project knowledge resources. The funding body (in this case EPSRC) does not specify what form 'dissemination' should take, nor provide any specification attached to 'Open Access' resources generated via public funding.

Outcome of the Observation: The information gathered from interviewing personnel conflicted with the information gathered via FOI request. Personnel were either not informed correctly about the level of public funding going into the project (they initially denied the project was funded with public money and said it was solely funded by paying customers, and there was no hint on the project home page about sources of funding) or they were deliberately providing vague information when answering questions, to avoid going into more detail. The case is currently open, and further FOI requests have been issued to gather additional documentation and understanding of the balance between 'Open Access' vs. 'commercial exploitation' of knowledge resources, although due to constrained resources further analysis of this particular case may be remanded to future work.

Due to the complex contractual arrangements however, this case has not contributed to the outcomes and conclusions presented in this thesis, but serves mainly as justification for further research.

2.1.4 Reflections

These observations are part of the exploratory part of field work. The first observation was serendipitous, as this researcher tried to access, and was denied, specific project knowledge, despite the project being publicly-funded through a research council that declares to abide to Open Access policies.

This outcome raises issues about the scope and the effectiveness of existing Open Access policies, and motivated further research into how these are implemented and monitored. Various conversations with peers and academics followed, that helped to gain an insight into knowledge sharing attitudes of engineering researchers, which contributed to shape some aspects of the knowledge audit templates adopted in the research instruments, presented in the relevant chapter of the thesis.

Finally, attempting to gather insights into the knowledge sharing practices of AURC showed that researchers have limited or no awareness of Open Access policies, and limited awareness of public funding their institutions receive, and the Open Access obligations that go alongside public funding. The observations above are exploratory. The outcome of these observations directly motivates and justifies to the shaping of the hypothesis, research question and methodology, presented in the next section of this thesis, but no conclusion is drawn based on the observations alone. The proposition and initial research question that emerges from the initial observations is summarised in Fig. 4:

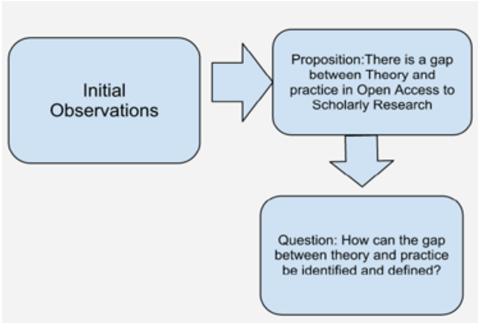


Fig. 4: Initial Observations Leading to Main Research Question and Hypothesis

2.2 QUESTIONS AND HYPOTHESES

From the preliminary observations outlined in the previous section, the following considerations emerge:

- UK systems engineering researchers, project leaders an public grant holders in engineering research seem to be insulated from the wider 'Open Access' movement to scholarly research which is pervading scholarly research debate in other disciplines.
- Engineering research knowledge outputs in the UK seem to be constrained by contractual commercial agreements in place between universities and industrial partners, the extent and modalities of which are unclear and not transparent.
- 'Knowledge' is still a poorly understood construct among the SE community, despite it being painstakingly, and at times even pedantically, well specified in literature and knowledge-related disciplines such as knowledge management (KM) and artificial intelligence. A working research question is therefore formulated as follows:

Given widespread knowledge sharing policies such as 'Open Access', to what extent:

a) are these effective in granting access to knowledge for the purpose of knowledge re-use?

and

b) are these adhered to?

Some aspects of the knowledge sharing/Open Access policy problem space are inter-related, and when tackled in combination, *systemic traits* - emergent characteristics that become apparent only when considered 'as a whole' - become visible. This systemic inter-relatedness can be viewed as 'entanglement', illustrated schematically in the diagram below and a broad proposition can therefore be articulated as in Fig. 5:

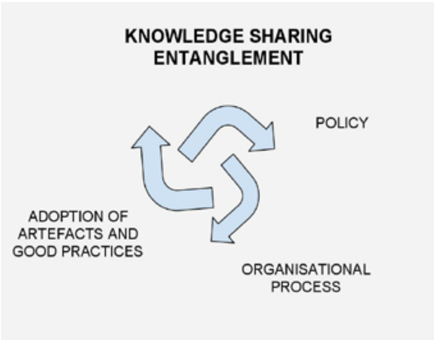


Fig. 5: Knowledge Sharing Entanglement

There is a gap between theory (T) (knowledge sharing policies) and practice (P) (knowledge resources accessible freely on the internet) in publicly-funded projects. An equivalent/alternative formulation can be worded as follows:

Policies such as Open Access are:

a) not necessarily supporting knowledge sharing requirements

and

b) not always adhered to in practice.

To test the proposition above, this research asks the following questions Q1, Q2 and Q3 and formulates the corresponding hypotheses H1 and H2, and H3:

===

Q1. How can the gap between T and P be identified?

H1. By gathering and comparing evidence of T and P.

===

Evidence of T can be gathered by surveying and analysing the scope and modalities of knowledge sharing policies (knowledge sharing 'in theory')

and

Evidence of P is gathered by auditing publicly-funded research projects (knowledge sharing 'in practice').

If after answering Q1 the hypothesis H1 is proven, then a research question that follows is:

==

Q2. Is there a gap between T and P?

H2. Despite knowledge sharing policies, such as Open Access mandates being adopted by research funding councils, knowledge seekers cannot locate Open Access knowledge resources via online web searches for every publicly-funded project.

===

This research addresses the question and tests this hypothesis in relation to a specific research sector, in a specific country (publicly (EPSRC) funded systems engineering research in the UK).

A further research question and hypothesis can be derived as follows:

===

Q3. How can the gap between T and P be reduced?

H3. By devising integrated socio-technical measures and interventions.

===

The research plan that follows is devised primarily to verify whether the main proposition is true (that there is a gap between the theory and the practice in knowledge sharing policies), Q1 and Q2, and to test H1 and H2 respectively. Q3 and H3 are formulated for completeness, and addressed mainly in the final chapter, as part of the recommendations.

Evidence is gathered from a mixture of empirical and non-empirical methods such as literature, surveys and correspondence and face to face exchanges, and audits (cases) to evaluate whether the main proposition is true. Further secondary hypotheses are formulated and tested at the data analysis and validation stage, see Chapters 5 and 6.

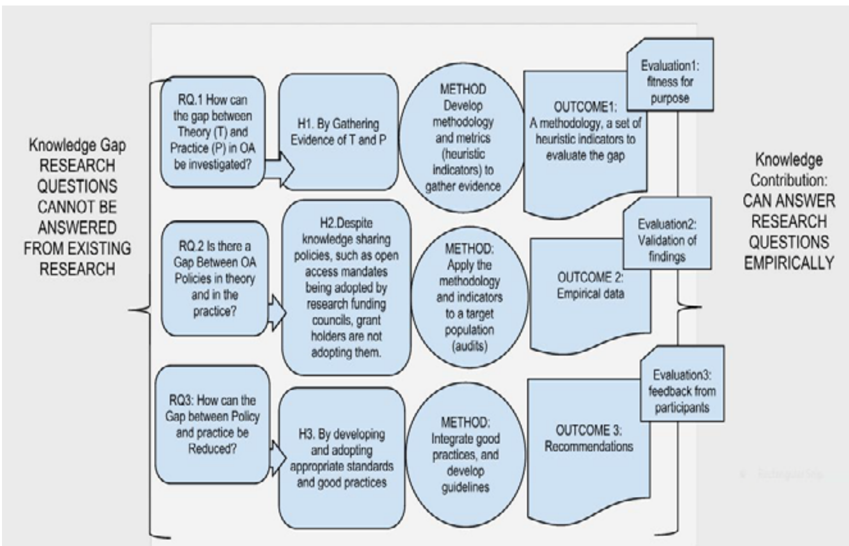


Fig. 6: Research Design

The research and methodological challenges associated with the systemic socio-technical aspects of knowledge sharing and re-use of publicly-funded research outputs are many-fold and cross several disciplinary boundaries. They could be formulated and phrased in a number of ways. Some of these challenges are discussed in Section 2.3.

The questions and hypotheses as formulated here are the result of some necessary compromise: they are generic enough to allow for a creative mix of approaches and methods to be applied in answering, yet sufficiently defined and pragmatic to be feasibly and credibly addressed within limited resources, despite the lack of established methodologies. After all:

> *"research question(s) should be 'doable', so that they can realistically be achieved within real-world constraints of time, money, expertise, access and ethics."* (O'Leary, 2004)

2.3 RESEARCH CHALLENGES

Thanks to the increasingly rapid and intense data and information exchanges, knowledge is developing exponentially, however the ability of humans in general and especially for researchers, to acquire, process and apply vast amounts of knowledge may not be developing as fast. Equally, organisations, research institutions and funding bodies, as bureaucratic entities, do not evolve as rapidly as required to adapt their structures and infrastructures quickly enough to the constant state of flux to achieve and maintain optimal functionality.

An overarching challenge for this research is that it spans several disciplinary boundaries, not supported by any individual methodology or language;

even assessing the state of the art of related topics via literature review can be disorienting and may result in what can appear as a superficial and fragmentary overview of the issues at stake:

> *"In an interdisciplinary project, one cannot start from a single, coherent, agreed-upon set of assumptions and proceed in a logical, linear fashion from premises to conclusion as in most disciplinary projects."*
> (Newell, 2007)

Further challenges are discussed below.

2.3.1 Methodological Challenges

A 'systemic' approach to study a phenomenon (in the case of this research the effectiveness of knowledge sharing policies in place) tackles simultaneously (at least) two dimensions of a question:

> *"Regardless of any specific definition, we observe that the concept always distinguishes between (at least) two different levels of abstraction, or systems levels: the system as a functioning unit and the system as a set of interacting parts. We regard a system as a primary unit when we treat it as a black box and ask about its overall behaviour -- i.e. what it does or accomplishes. As a set of parts or components (which somehow work together to produce the system's overall behaviour) we can examine the system's construction -- i.e. its internal structure and processes. It is this distinction between system levels -- between the behaviour of the system as a whole and the specific relationships between its parts -- which is fundamental to the concept. The idea of a system would be meaningless without this distinction."* (Ritchey, 1996)

In this inquiry, 'knowledge sharing' challenges in relation to scholarly outputs are considered at one level *as a whole* (do existing knowledge sharing policies serve their purpose 'effectively'?) and at another level how they are composed and combined (what knowledge sharing policies are there and does their combination serve their purpose?).

The methodological challenge is faced here by drawing from a rich and eclectic tradition of 'system research' which adopts heterogeneous, hybrid approaches to problem formulation and solving, and is becoming increasingly popular (and necessary) to study complex, large scale socio-technical challenges.

As a result, 'mixed methods' approaches, which have gained acceptance even in the most rigorous scholarly circles in recent years, are the only possible methodological pathways to tackle 'messy' problems (Horn, 2001).

2.3.2 Knowledge Domain Challenges

Systems engineers are not new to mechanisms for formalising knowledge, yet when it comes to defining the systems engineering domain itself – in terms of theory and practice, vocabularies, axioms, shared conceptualisations – efforts are still fragmented and largely not in the public domain. Despite being increasingly widely practised in all fields and across many disciplines, systems engineering is still associated with the defence sector, especially in the UK, where knowledge sharing is not encouraged, and is constrained by a secretive, proprietary and retentive culture (Kim, 2005). Thanks also to the adoption of the web as an 'always on', unconstrained knowledge exchange hub, 'system of systems' architectures are taking shape at various levels (social, technical, topical, data and information levels) as the result of various combinations and meshups that constitute the fabric of the open web, and a 'global' web-centric information sharing culture is trying to become mainstream.

The systems engineering body of knowledge is trying to find its shape and place in this 'open web' landscape, facing many challenges:

> *"The discipline of systems engineering (SE) has been recognized for 50 years as essential to the development of complex systems. [...] SE is still treated primarily as heuristics learned by each practitioner during the personal experimentation of a career. The heuristics known by each differ, as shown by the fractured development of SE "standards" and SE certification. [...] As a result of this heuristic understanding of the discipline, it has been nearly impossible to quantify the value of SE to programs. [...] The differences in understanding, however, just as typically result in disagreement over the level and formality of the practices to include. In response to the uncertainty in the heuristics, some efforts have been made to identify "leading indicators" that provide early tracking information as to the worth of a project's SE efforts."*
>
> (Honour & Valerdi, 2006)

In contrast to traditional engineering and science fields, doctoral research in systems engineering is characterised by unique factors, for example, a relatively young tradition of systems engineering academic programs, the necessity for hybrid research methodologies, the existence of strong links with industry and government (Rhodes & Valerdi, 2007). One of the challenges that should be added to the characterisation as it emerged from this research, is the difficulty in accessing systems engineering research scholarly outputs, as noted in the preliminary observations. In addition, most of the grand challenges for systems engineering research, summarised from literature and highlighted in Table 2, contain an underlying shared 'knowledge dimension', culminating in Grand Challenge 5, which is directly related to accessibility and reliability of information and knowledge, tackled directly in this research.

It can be said that knowledge representation and sharing are core issues in systems engineering challenges. The issues identified above are only the tip of a much bigger iceberg; massive shifts are happening at all levels of the social sphere, research and scientific paradigms and assumptions are being challenged at an unprecedented rate. And things are changing very fast; many of the phenomena under empirical investigation – including policies and data about their adoption – have shifted radically throughout the duration of the research, making it like (just a minor interpretive parenthesis) trying to stand on quicksand.

SE GRAND CHALLENGE	KEY ISSUES	RELATION TO THIS RESEARCH
"Ultrascalable Heterogeneous Systems"	"Open and scalable systems architecture"	"• Ubiquitous communications and user access • Management of ultra-scalable data – ensuring only relevant data of interest is accessed • Data provenance – knowing the exact status of information "
"Ultrascalable Autonomous Systems"	"To function, an autonomous system incorporates disciplines such as AI, networked computing, object-orientation, engineering, economics, sociology and organisational science."	"• Information provenance – an ability to take full accountability for action Integration of the autonomous system in an environment where human controlled systems coexist"
"System Verification, Validation and Assurance (VV&A) of Extremely Complex System"	" Functional correctness (such as conformance to functional specification, type correctness, consistency of data, numerical accuracy etc.) as well as many non-functional issues (such as dependability, safety, security, timeliness of response, availability, maintainability etc.)"	"Unless, the system has been designed for reliability from the outset a single bit error can have catastrophic consequences for the whole system. System reliability, resilience engineering"
"Modelling & Simulation (M&S) - Total System Representation "	"Challenges relating to re-use of system models open and unresolved. high levels of knowledge and skill required , support of domain and subject matter experts in order to interact with model components"	"• Accurate representation of the human (physical, cognitive and performance) "
"Through Life Information and Knowledge Management"	"Information and knowledge relating to the product can expand, including design decisions taken during iterations, could potentially include all the email traffic associated with the product, telephone calls, meeting memos, diagrams, photographs and other forms of data."	"• Ensuring the data is secure and reliable • Information provenance – ensuring what is real remains real data • How to retrieve information in an intelligent and usable manner • Development of a diagnostic consultant that is able to search the data for relevant information • Development of an information agent tool • Reliable and efficient storage of data "

Table 2: Excerpted and Adapted from: Research Grand Challenges (Kalawski, 2009)

2.3.3 Systemic Challenges

Formulating and addressing the problem of a complex, entangled problem space is riddled with 'systemic challenges' which in literature have been addressed as 'wicked problems' (Rittel & Webber, 1973), typically identified by the following characteristics:

- There is no definitive formulation (defining wicked problems is itself a wicked problem).
- Have no stopping rule (neither the problem nor a solution can be framed within a fixed parameters)
- Solutions to wicked problems are not true-or-false, but better or worse
- There is no immediate and no ultimate test of a solution to a wicked problem
- Every solution to a wicked problem is a "one-shot operation"; because there is no opportunity to learn by trial and error, every attempt counts significantly
- Wicked problems do not have an enumerable (or an exhaustively describable) set of potential solutions, nor is there a well-described set of permissible operations that may be incorporated into the plan
- Every wicked problem is essentially unique
- Every wicked problem can be considered to be a symptom of another problem

Along similar lines, traits similar to 'wicked problems' have been identified as 'social messes' (Horn, 2001) defined as:

- Different views of the question and contradictory solutions
- Most issues are connected to other issues
- Data is missing
- Conflicting values

- Cultural and socio-economic constraints
- Different logics (that can seem illogical under a single perspective)
- Multiplicity of opportunities for intervention
- Unpredictability, uncertainty
- Change aversion and resistance

Social messes are:

> *"More than complicated and complex. They are ambiguous, contain considerable uncertainty - even as to what the conditions are, let al.one what the appropriate actions might be, are bounded by great constraints and are tightly interconnected, economically, socially, politically, technologically."* (Horn, 2001)

The complexities of 'wicked problems' and social messes are described using a number of philosophical metaphors, and re not dissimilar from what the social sciences refer to as 'intractability' and 'incommensurability' (Kuhn 1970; Feyerabend 1975). Intractability, in the context of socio-technical systems and risk assessment is defined as: a system or a process is intractable if the principles of functioning are only partly known or even unknown, if descriptions are elaborate with many details, and if the system may change before the description is completed (Hollnagel, 2008).

The 'knowledge sharing' challenge described here in a wider socio-technical context, shares the traits of intractable systems. As such, no single 'off the shelf' method or approach exists for their solution.

The tension between different related forces that characterise a problem space, can result is what is called 'fragmentation' (Conklin, 2005) illustrated in Fig. 7:

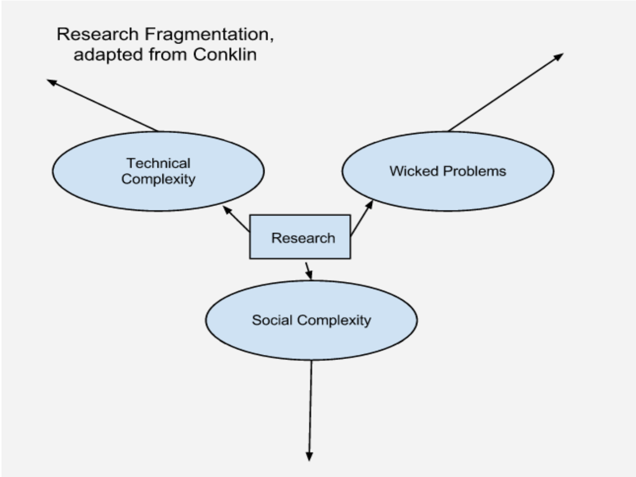


Fig. 7: Research Fragmentation, (adapted from Conklin, 2005)

> *"The antidote for fragmentation is coherence. How, then, do we create coherence? In organizations and project teams - in situations where collaboration is the life blood of success - coherence amounts to shared understanding and shared commitment. [...] Coherence means that a wicked problem is recognized as such, and appropriate tools and processes are constantly used to 'defragment' the project."* (Conklin, 2005)

This doctoral thesis, compatibly with the challenges identified above, intentionally frames the research problem and questions as being *wicked* and *a mess* (with all the methodological uncertainties this may imply), taking into account fragmentation of the various aspects of the problem space, and points to a set of pragmatic and systemic solutions, intended to contribute to increase 'shared understanding' not just of meaning and context, but also of the dimensions and issues in the problem, to promote a shared commitment of diverse stakeholders to finding measures that realistically address the challenges.

Among the research instruments devised to support the integration of disparate angles of the field work, a morphological box (aka Zwicky box) is developed and presented in Chapter 3 of this thesis.

2.3.4 Hidden Agendas

Important political and economical imbalances are becoming visible at global and local level and more difficult to justify, thanks to the increased transparency and access to knowledge that impact the governance structures of organisations and countries alike. Research – agendas and funding – is an important constituent of the wider social machinery. The shifts of power that are taking place thanks, among other forces, to growing awareness and better knowledge, are inevitable.

However, red tape and measures and interventions can be deliberately designed with unnecessary built-in costs and dysfunctionalities to at least slow down and muddle up a process of societal transformation and change that is otherwise probably irreversible and with long term transformative impact in all areas of human and social activities. This can result in perfectly logical, pragmatic and cost-effective solutions to research problems and questions being avoided and dismissed or even refuted based on one or other theoretical viewpoint, or scholarly dogma.

2.4 SIGNIFICANCE AND RELEVANCE

The study presented in this thesis has been devised to answer directly the research questions and test the hypotheses outlined above, however it also addresses generic knowledge gaps exposed in the research agendas of different but related, disciplines. Below a short list of pointers as to 'why this research is needed' and 'what knowledge gap does the research address' from different fields of study:

1. 'Knowledge sharing on the web' is a relatively novel topic for the scientific community, which has developed in the past twenty years or so, and which is studied, researched and widely debated with different emphases across different disciplines. By exploring the correlation between different novel disciplinary dimensions (the policy and the practice), and by devising novel instruments and techniques and deploying them in the field, this research aims to contribute to the framing and optimisation of otherwise disparate aspects of initiatives that are intended to be central to human and social development and, overall, to the progress of science. More importantly, this research relates the field of study of knowledge sharing to the context of socio-technical systems, and system of systems engineering research, as discussed in more detail in the sections that follow.

2. The outcomes of this research (the proposed instruments and its recommendations) are developed taking into account open questions in relation to knowledge sharing being raised in the socio-technical systems research agenda, such as:

> **How can different types of knowledge be captured at low cost and maintained in an accessible way?**
> *"The problem of low-cost knowledge capture was, we believe, one reason why many attempts to implement organisational memory systems in the 1990s were ineffective. Capturing knowledge for the future distracts people from their everyday work so we need to discover techniques that capture information from normal work activities with minimal intervention from the people involved in these processes."* (Baxter & Sommerville, 2011)

and

How can the use of organisational memories and other support for organisational learning be embedded in the socio-technical systems engineering (STSE) process?

"Organisational memories and learning from experience can only be effective if they are actually used. We need to invent ways of easily accessing such information as part of routine processes and ensuring that the information can be updated with accounts of practical usage experience." (Baxter & Sommerville, 2011)

In consideration of the open research questions above, this inquiry contributes:

a) a socio-technical perspective to addressing Open Access issues

and

b) a mechanism toward standardising knowledge sharing for system knowledge, which includes systems science, systems engineering and System of Systems Engineering Research SoSE).

The Systems Engineering paradigm in the last decade has moved toward System of Systems (SoS) research (Ackoff, 1971), however few knowledge processes and models take into account the SoS requirements in the socio-technical dimension highlighted earlier. It is said that:

"It is only through respecting the needs of the various stakeholders, looking at the SoS from a holistic viewpoint, and balancing these perspectives that an SoSE team can achieve an environment of cooperation, transparency, and trust needed to successfully evolve an SoS to meet new capability needs." (Lane & Valerdi, 2010)

This research widens the understanding of the knowledge sharing requirements, knowledge model to facilitate the transparent knowledge flows in SoSE, as well as in systems engineering research in general.

3. Existing studies in Open Access research, referenced in Chapter 3 of this dissertation, as well as the background research that they reference, examine Open Access from a 'macroeconomic' perspective, and are based on quantitative, statistical methodologies over a large, aggregate dataset relating to generic corpora of scholarly publications. The methodology presented in this dissertation, by contrast, is designed to search the existence of Open Access resources through the lens of the knowledge seeker who may not know the title of the paper they are looking for, and to provide evidence from the 'bottom up' and using broader and more qualitative parameters than purely statistical ones.

Rather than searching for journal articles using their title, as done in the majority of studies to date (a variable that must be known to perform the search) the methodology used in this doctoral thesis and presented in Chapter 4, Open Access Monitor (OAM), searches for funded project name and grant number to see what journal papers, or other resources (web pages, videos, etc) match the search term. In our study the methodology is applied to target a specific subject (systems engineering research) in a specific country (UK); however, it can be applied to any field, sector and funding council, contributing a novel methodological perspective to Open Access research.

4. In Chapter 3, some omissions (a knowledge gap) in a recent journal article 'knowledge sharing, future research directions' are identified and addressed, by proposing the addition to future research agenda the policy aspect.

5. The study is also intended to contribute to addressing, at least in part, Research Utilization (RU) challenges: how to make research outputs accessible, useful and relevant to wider audiences (Davies et al., 2005). RU issues are well summarised by the following statement[7]:

> *"The purpose of research is to be of use, to change current practice, or to confirm it. Yet the process of moving new understandings and new products from research to practice usually takes years, decades, or even generations. Although there are good reasons for moving carefully—new research needs to be evaluated, replicated, and refined—too often the pace of change is set, not by a rigorous process of review and refinement, but by the gap between the research community and the world of practice. Research on dissemination, or knowledge utilization as it is sometimes called, has yielded a wealth of information about what does and does not work.*
>
> *But, due to this gap, those understandings for the most part have not moved from the research community—those who study the process of knowledge use—to the practice community—those responsible for adopting and applying research outcomes. As a result, most dissemination practices are still based on a mechanistic, linear conception of dissemination as a process of "getting the word out."* (RUSH)

[7] RUSH, Research Utilization Support and Help, retrieved September 2011 http://www.researchutilization.org/learnru/welcome2ru/

2.5 METHODOLOGY

2.5.1 Introduction to the Methodology

Research and evaluation methodologies can be determined by the paradigm of inquiry adopted (Guba & Lincoln, 1988), however the paradigm of inquiry adopted does not necessitate a single, inflexible methodological position (Patton 1988; Shulman 1988; Salomon 1991).

The initial months of this research were devoted to explore, tentatively and experimentally, the problem space via exploratory observations, as reported in Chapter 1.

The difficulties encountered then constituted one of the constant challenges throughout this research, at the same time they also provided an immediate snapshot of the problem under investigation: knowledge retention, and knowledge sharing aversion; almost a cultural norm, a code of silence. Instead of being discouraged by the poor outcomes of the exploratory observations, these became instrumental to the decision to opt for a research plan that would not rely solely on the cooperation of engineers and researchers. The observations pointed to a dichotomy between the theory and the practice in knowledge sharing of publicly-funded research, and to possible paradoxes: on the one hand research institutions make claims of excellence, yet few of the research data and research outcomes are publicly accessible, or even verifiable (the public has to take their word that they have conducted science, but nobody can actually check how sound their data or their methodology is, other than what is presented in publications, which is often just an outline).

Public Research Councils declare that they support 'Open Access', and boldly display prominent links to online policies on their websites, but do not often, and surely not as a matter of course, share statistics or data sets to evaluate and monitor as to how these policies are implemented.

Academics employed by award-winning research institutions are reluctant to discuss their knowledge sharing behaviours in relation to the research projects they work for, and when eventually they accept to answer some questions, do so telegraphically and under conditions of limited disclosure, often without justification and provide answers that when cross-referenced with documentation disclosed under the FOI Act, appear to be incorrect (or incomplete, or too vague to be conclusive). Some of these exchanges are reproduced, in anonymised form, purely as a documentary example, in Appendix C.

An ethnographic approach was therefore opted for; rather than trying to elicit information directly, systems engineers were observed and approached indirectly, in the course of normal academic or professional activities, and relevant questions asked in the context of casual conversations. A list of specific systems engineering events where observations and conversations have taken place is provided in Appendix D.

The main methodology is structured by integrating an 'Evidence Based Research' (EBR) approach with 'Mixed Methods Research' (MMR), illustrated schematically in Table 3, and described in the paragraphs that follow.

EBR	MMR			
FOMULATE Q				
SEARCH FOR ANSWERS/ GATHER EVIDENCE	OBSERVATIONS and ETHNOGRAPHY	LITERATURE REVIEW	correspondence + FOI REQUESTS	AUDITS
APPRAISE EVIDENCE	SYSTEMATIC REVIEW OF THE FINDINGS M E T A - A N A L Y S I S			

Table 3: Illustration of Integration of EBR and MMR

2.5.1.1 Elements of ethnography

Three of the main principles of ethnography (Hammersley, 1990 cited in Genzuck, 2000) are:

- Naturalness – whereby participants are observed in their natural settings to minimize interference with natural behaviour
- Understanding – must gain an understanding of the cultural perspective

- Discovery – the research process is inductive or discovery-based, rather than limited to the testing of hypotheses: if one approaches a phenomenon with a set of hypotheses one may fail to discover the true nature of that phenomenon, being blinded by the assumptions built into the hypotheses.

Other established tenets of ethnographic research methods are adopted as follows: (a) Knowledge sharing behaviours are studied where possible (for example via audits) in their natural context.

(b) In addition to gathering data via literature review and audits, direct observations and/or relatively informal conversations have been used to gather evidence.

(c) Some of the evidence was gathered using unstructured approaches, for example by collating and comparing information gathered in different email exchanges, rather than in a single structured email questionnaire.

> *"…unstructured in the sense that it does not involve following through a detailed plan set up at the beginning; nor are the categories used for interpreting what people say and do pre-given or fixed.*
>
> *This does not mean that the research is unsystematic; simply that initially the data are collected in as raw a form, and on as wide a front, as feasible."*
>
> (Genzuck, 2000)

The early fieldwork, which shifted from approaching individuals and asking direct questions to more subtle observations, did not lead to particularly conclusive outcomes per se, but provided evidence and justification for a more pragmatic autonomous approach to data gathering, that later resulted in an 'auditing framework' used to carry audit Open Access adoption.

2.5.1.2 Evidence Based Research

Evidence Based Research (EBR) is the systematic study of evidence-based methods and strategies, known as 'Evidence Based Practice' (EBP) and increasingly used in the clinical and social sciences (Paynter, 2009).

Evidence Based Practice, widespread in the medical community, encourages researchers and professionals in their respective disciplines to make:

> *"conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients."*
> (Sackett et al., 1996)

The rationale for EBR is rooted in clinical practice in the health and medical domains, however a methodology has grown out of it, that has been adopted by other social science disciplines.

> *"The term evidence or evidence-based, as it relates to research-based knowledge, pertains to the summative collection of research on a specific topic that answers specific and important questions (e.g., questions regarding relationships, why problems exist or persist, or what is the best decision for policy making."* (Raudenbush 2002; Shavelson & Towne 2002)

and

> *"While research quality pertains to the scientific process, evidence quality pertains more to a judgment regarding the strength and confidence one has in the research findings emanating from the scientific process."*
> (Mosteller & Boruch 2002; Shavelson & Towne 2002)

A typical EBP research process includes, for example, the following steps:

(1) Formulate the question.

(2) Search for answers.

(3) Appraise the evidence.

(4) Assess the outcome

(Gray, 2004)

A typical research design for EBR is illustrated by the 'pyramid metaphor' (Paynter, 2009), Fig. 8, whereby qualitative studies, in the case of this research ethnographic observations, serve as the basis for quantitative studies, the outcomes of which are then analysed systematically to issue guidelines and recommendations.

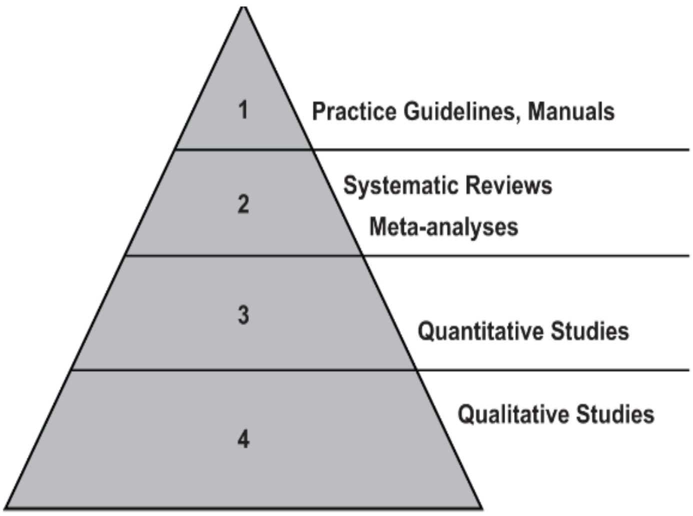


Fig. 8: Evidence Based Research Pyramid (Paynter, 2009)

Although EBR in recent years has gathered the status of a pragmatic, 'standalone' research approach, gathering evidence has always been the core of any scientific inquiry as well as technical investigations. Gathering different types of evidence has been done in case study research, Table 4 (Yin, 1994).

Source of Evidence	Strengths	Weaknesses
Documentation	• stable - repeated review • unobtrusive - exist prior to case study • exact - names etc. • broad coverage - extended time span	• retrievability - difficult • biased selectivity • reporting bias - reflects author bias • access - may be blocked
Archival Records	• Same as above • precise and quantitative	• Same as above • privacy might inhibit access
Interviews	• targeted - focuses on case study topic • insightful - provides perceived causal inferences	• bias due to poor questions • response bias • incomplete recollection • reflexivity - interviewee expresses what interviewer wants to hear
Direct Observation	• reality - covers events in real time • contextual - covers event context	• time-consuming • selectivity - might miss facts • reflexivity - observer's presence might cause change • cost - observers need time
Participant Observation	• Same as above • insightful into interpersonal behavior	• Same as above • bias due to investigator's actions
Physical Artifacts	• insightful into cultural features • insightful into technical operations	• selectivity • availability

Table 4: Types of Evidence, (Yin, 1994)

The emphasis in EBR is on a pragmatic approach that uses findings obtained from the evaluation of evidence to directly formulate interventions; in the case of our study these take the form of guidelines and recommendations. An evidence-based approach helps to cut through the challenges of heterogeneous research designs, especially when the eclectic combination and variety of methods used makes it hard to find a single theoretical framework that can be validated and agreed upon.

It is said that:

> *"those who work within a single theoretical framework find others unintelligible."* (St. Pierre, 2002)

There is consensus that evidence-based research is valuable and possibly the most adequate method to inform government policy and guide organisation (Fitzallen & Brown, 2007). In this research, evidence has been gathered and validated using a variety of methods in combination, as explained in the research design and evaluation sections of the thesis below, respectively.

The research design takes into account a typical set of tasks in EBR (Levitt et al., 2010):

a. scoping
b. gathering evidence
c. analysis
d. discussing evidence
e. reaching findings and conclusions
f. making judgements
g. reporting

2.5.2 Mixed Method Research (MMR)

Mixed methods research is:

> *"A class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study. Philosophically, it is the "third wave" or third research movement, a movement that moves past the paradigm wars by offering a logical and practical alternative."*
>
> (Johnson & Ongwubuzie, 2004)

In mixed method inquiry the switching back and forth between different questions and propositions is perfectly admissible:

"In MMR design takes the form of an iterative inquiry into the nature of a situation to build and test hypotheses.
Inherent is the recognition that our design will get some things wrong, and that a big part of design is redesign to clarify and strengthen our thoughts by continuously subjecting our hypotheses to critical review. Research design should unfold most naturally as a discourse among stakeholders."
(Schmitt, n.d.)

and

"... the result obtained with one method can be used to select and justify the use of another method or, qualitative and quantitative methods can be used in combination to provide more comprehensive datasets."
(Creswell et al., 2007)

The recursiveness and iterations between thought paradigms as well as techniques and methods used in this thesis find correspondence in hermeneutics (Heidegger, 1927). A cyclic research design inspired by the 'hermeneutics' (Heidegger, 1962) has been recently explored in relation to mixed methods (von Zweck et al., 2008).

2.5.2.1 Why MMR?

Pure research problems that can be solved with individual conventional methodologies are increasingly rare; instead, socio-technical problem solving requires the integration of more than only one approach, and methodology and research can have significant social, economic, psychological and political implications beyond the targeted scope (Creswell, 2008):

"As the complexity, interaction and rate of evolution of human systems increases exponentially, the adoption of an ad hoc, mixed method approach to research has become necessary, and it is increasingly used to support the depth and breadth of research problems and questions, although there are still limited skills and institutional support for MMR techniques"
(Creswell & Plano Clark, 2010)

MMR fits a problem whereby the research dimensions are essentially *unknowable*, and no amount of information collection or analysis will reveal objective truth or provide the ability to make with certainty (Creswell, 2008):

"The complexity of our research problems calls for answers beyond simple numbers in a quantitative sense or words in a qualitative sense.
A combination of both forms of data can provide the most complete analysis of problems. Researchers can situate numbers in the contexts and words of participants, and they can frame the words of participants with numbers, trends, and statistical results. Both forms of data are necessary today. In addition, qualitative research has evolved to a point where writers consider it a legitimate form of inquiry in the social and human sciences."
(Denzin & Lincoln, n.d. cited in Creswell, n.d.)

This research, in 'looking for evidence' for both the theory and the practice of Open Access and knowledge sharing, adopts a wide combination of methods.

2.5.2.2 Worldview

A mixed method research plan should include a philosophical worldview (Creswell, 2008), which can refer to a 'basic set of beliefs that guides action (Guba, 1990) as well as to the 'preferred conceptual model of the reality under observation' more or less coinciding with the definition of 'paradigm' (Guba & Lincoln 1988; Mertens,). The term 'paradigm' originates from the Greek *paradeigma* meaning pattern, model or example (Stanage, 1987).

Research (Creswell, 2008) identifies four classical worldviews/paradigms: positivism, constructivism, participatory and pragmatic. The worldview adopted in this research is 'pragmatic', whereby its main research dimension is identified through an evaluation of the 'practice'.

A pragmatic worldview implies that the value of an idea, action, or object is determined by its practical impact (Fishman, 1999). The pragmatic worldview is concerned with application – what works – and solutions to problems (Patton, 1988) making it an appropriate fit for the evidence-based approach discussed earlier. This research can also be considered 'systemic', and to some extent 'transformative', intended respectively as specialisations and refinement of Creswell's pragmatic worldview, whereby 'systemic' refers to a worldview that identifies the widest possible boundary (considers various aspects of the problem space as a whole 'system'), and transformative, proposed initially as a framework of belief systems that directly engages members of culturally diverse groups with a focus on increased social justice (Mertens, 2010) is here generalised as a worldview capable of delivering systemic 'social' change, as tentatively illustrated in Fig. 9, adapted from Creswell (2008).

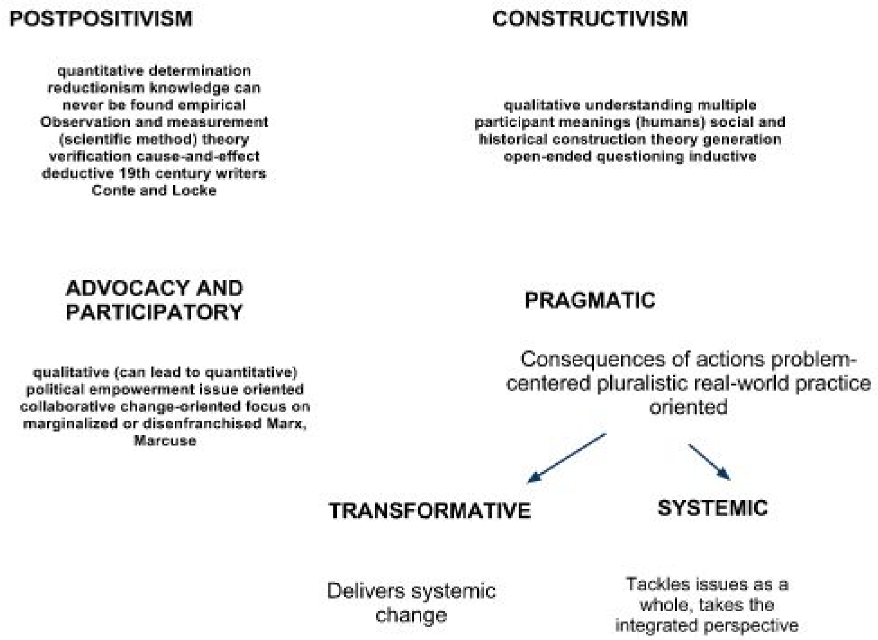


Fig. 9: Systemic and Transformative Specialisation of the Pragmatic Worldview

(Di Maio 2011, adapted from Creswell, 2008 and Mertens, 2010)

2.5.3 Research Design

Research designs for mixed methods studies can be divided into two main groups: **component designs**, where each component remains a discrete aspect of the research design and the integration of these aspects happen at the interpretation and conclusion stage, and **integrated designs**, where there is tighter interplay of different methods from the outset (Caracelli & Greene, 1997).

The research plan for this study adopts a component design, whereby each research component has its own internal logic and evaluation mechanism. MMR designs can be classified according to two major dimensions (Creswell, 2008):

1. Time order (i.e. concurrent versus sequential) and
2. Paradigm emphasis (i.e. equal status versus dominant status)

The time order in our MMR plan is hybrid, with most research activities taking place simultaneously (concurrent), while some activities, especially in relation to validation taking place sequentially.

The dominant paradigm emphasis is pragmatism, consistent with the 'worldview' discussed earlier. The analytical part of study consists of two empirical research components:

a) a survey and a critical appraisal of existing policies and legal instruments, providing evidence for the existence of knowledge sharing policies that should be in place (T).

and

b) a systematic review of funded projects (the audits), selected to specific inclusion criteria, described in more detail in Chapter 5, providing evidence of the degree of adoption of such policies in the practice.

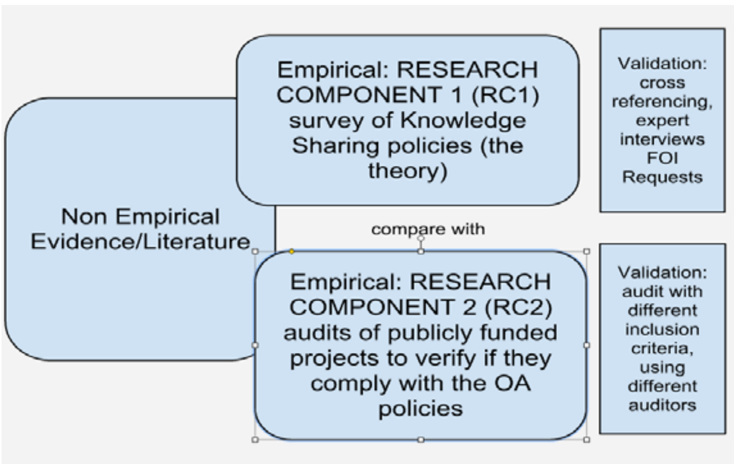


Fig. 10: Research Design

2.5.3.1 Research components

A survey and a critical appraisal (evaluation) of Knowledge Sharing policies that guide and regulate academic practice in the UK (Davies et al., 2005), is aimed at answering the following questions:

Exploratory evaluation: what policies are there to support knowledge sharing of (and access to) scholarly outputs?

The method employed in this research component is a survey of public documentation provided by funding bodies, validated by correspondence exchanged with civil servants and experts, and where necessary, FOI requests. The steps followed are:

- Identify public research funding bodies.
- Review existing literature and related research of knowledge sharing policies.
- Audit policies and, where possible, policy management strategies.

Impact evaluation: to what extent are these policies implemented?

The methods employed in this research component are audits and exchanges with selected case study participants. The steps are:

- Identify a representative sample of publicly-funded research projects in systems engineering in the UK
- Audit the sample using the auditing instruments devised to carry out data collection, presented and discussed in Chapter 4.

2.5.3.2 Chain of evidence

The 'chain of evidence' (Yin, 1994) for this research is illustrated as follows:

Observations (Aserp, Conversations) >>> Initial motivation, formulation of question/hypothesis

Field Work/Ethnographic >>> Gauge awareness of OA policies among the engineering community

Literature review >>> Previous studies confirm the findings from observations and fieldwork

Policy Analysis >>> Gather evidence of policies 'in theory' from funding councils

Audits >>> Gather evidence Systemic survey 'in practice' from the field (ongoing)

2.5.3.3 Validation of the evidence

In MMR, the multiple methods can be adopted as a way of mutually cross-referencing and triangulating the respective outcomes. For example, gathering evidence of the existence of Open Access policies has been carried out via a survey of research councils' websites, and analysis of literature where these are discussed, followed by correspondence and conversations or email exchanges to clarify the content of the policy where necessary.

This data is considered inherently 'valid' as it is 'given' by the authority. Indirectly however, the audits are designed as a validation mechanism to test empirically the degree of adoption of the policy. Furthermore, thanks to the FOI Act, it is possible to request access to supporting documentation as additional validation of the outcomes of correspondence and email exchanges. The overall quality of this research, and the more detailed aspects of the validation of its outcomes, such as internal and external validity and reliability, are addressed in Chapter 6.

2.5.3.4 Combining the evidence

In MMR the outcomes of each research component can be mixed according to different recipes, for example:

1. Merging or converging the two datasets by actually bringing them together,
2. Connecting the two datasets by having one build on the other,
3. Embedding one dataset within the other so that one type of data provides a supportive role for the other dataset. (Creswell, 2008)

In our research the main evidence is collected and validated as follows:

1. Research Policies >>>> data collection by survey of research funding bodies, and correspondence.

Validation: Obtain via FOI requests documentation to support/corroborate what the policies state.

2. How the research project share knowledge >>>> get list of funded projects by the funding body, then select a subset according to inclusion criteria, then physically audit (survey) the subset to find which knowledge resources including scholarly publications are publicly accessible.

Validation: The auditing process, specifically Phase 3 which prescribes emailing the grant holder to validate findings, is designed as 'built-in' validation, however validation is also carried out by having the audit repeated independently by two (or more) different independent auditors. The first evidence set serves as a baseline value (independent variable) for a comparative evaluation of the compliance of the second evidence set (the dependent variable).

The first evidence set shows the existence of a knowledge sharing/Open Access policy, and it is considered as an invariant, while the second evidence set shows the level of adoption of the policy, by auditing the individual artefacts and their location on the web which is considered the main variable for this research component.

2.5.3.5 Levels and units of analysis

The purpose of this research is to understand the relation between the theory and the practice in as far as knowledge sharing policies adoption is concerned. In terms of research design, this translates into the mapping of different levels of analysis: at one end of the research plan the level of analysis is the funding body, and the unit of analysis is its 'policy', and at the other end the level of analysis is the individual publicly-funded research project, and the unit of analysis is individual artefacts (seeking URIs of websites, journal papers and other knowledge resources), Fig. 11a.

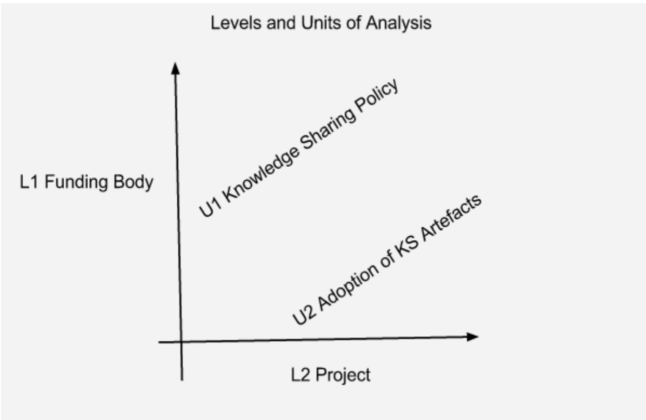


Fig. 11a: Levels and Units of Analysis

One of the novel and unique contributions of this research consists of exploring the relation between these two dimensions with a systematic review, to the best of the author's knowledge, at the time of writing, it is the first of this kind.

This research complies with the central tenets of what constitutes a 'systemic review' method as laid out by the Evidence Based Policy and Practice Institute (EPPI):

· Explicit and transparent methods are used following a standard set of stages.

· It is accountable, replicable and updateable.

· User involvement is built into the research design.

2.5.4 Other Research Techniques Used in This Research

EBR and MMR can be considered the pillars of the overarching methodological framework for the overall research design, however various techniques have been used to gather evidence. These are outlined below.

2.5.4.1 Systematic review and meta-analysis

A systematic review, also sometimes called an overview, consists of a summary of all the evidence gathered to answer a specific question, which frequently involves the effectiveness of an intervention; in the case of this research, the effectiveness of knowledge sharing policies such as 'Open Access'. The main methodological benefits offered by a systematic review (versus, for example, a narrative review) is that it aims, by a process of rigorous critical synthesis, to overcome different kinds of biases, for example by using explicit, preset criteria to select studies for inclusion on the basis of relevance and validity. A second way is by having more than two people independently make study selection decisions, compare results, and discuss discrepancies before moving on to independently extract data from the studies (Ciliska et al., 2001). Both techniques have been applied to the study presented in this thesis.

Meta-analysis consists of applying quantitative and statistical techniques to produce an overall summary across the study. In this research, meta-analysis is used to provide an overall analytical summary of the data gathered via the audits.

2.5.4.2 FOI requests

Some of the evidence in this research has been gathered via FOI requests, as a validation technique for the evidence collected via email exchanges. Doing research using FOI requests is relatively new, since the Act has been in place less than two decades, but it is increasingly adopted to complement other research methods[8]. Few formal guidelines exist on how to present and publish information obtained via FOI requests. Where necessary, essential documentary evidence is provided by reproducing key snippets of exchanges in anonymised form.

2.5.4.3 Knowledge model building

Models are 'abstractions', and the practice of 'model building' used to project, plan and devise artefacts has been used formally and less formally in all ages. Model building is loosely derived from 'Model theory':

> *"model theory is now a sophisticated branch of mathematics but in a broader sense, it is the study of the interpretation of any language, formal or natural, by means of set-theoretic structures, with Alfred Tarski's truth definition as a paradigm."* (Stanford Encyclopedia) [9]

Knowledge modelling techniques revolve typically around three steps, knowledge identification, specification and refinement (Schreiber & Wielinga, 1993):

Knowledge identification: consists of locating information and knowledge sources as the main references for the knowledge model, and developing simple artefacts such as a core glossary for domain terms and top level generic categories for the domain being modelled.

---

[8] Freedom of information: what's in it for researchers? Research Information Network http://www.rin.ac.uk/news/events/freedom-information-whats-it-researchers

[9] Stanford Encyclopedia http://plato.stanford.edu/entries/model-theory/

Knowledge specification: a specification language is a semi-formal, explicit and codified formalisation of the knowledge being shared. This step consists of a first attempt at 'encoding' – that is, representing formally – the artefacts devised in step one.

Knowledge refinement: in the final stage of knowledge modelling, attempts are made to validate the model and resulting artefacts as much as possible, via simulation based on some externally provided scenario (paper-based or via prototype).

Notions from knowledge modelling techniques have been adopted as methodology to devise the auditing templates of the collection instruments, as well as to devise the shared knowledge model and object outlined as an example in the final recommendation section. A more targeted effort related to this doctoral thesis consists of vocabulary extraction, and it has been contributed to the SEBOK (Systems Engineering Body of Knowledge) and reported in a separate paper published as ACM Proceedings (Di Maio 2011) and listed in Appendix A.

2.5.4.4 Morphological analysis

To tackle the complex entanglement of factors that make up the socio-technical problem space addressed in this research, a morphological analysis approach was undertaken, albeit experimentally, to support the theoretical exploration of the research field. Morphological analysis is a group of methods that breaks down a system, product or process into its essential sub-concepts, each concept representing a dimension in a multi-dimensional matrix. Thus, every product is considered as a bundle of attributes.

New ideas are found by searching the matrix for new combination of attributes that do not yet exist (Ritchey, 2005).

A morphological field including Theory, Methodology and Artefacts with Cognitive, Organisational and Technical aspects is constructed.

2.6 ETHICAL CONSIDERATIONS

This doctoral research is based on data and information collected using a variety of methods. The University of Strathclyde, the institution to which the thesis is being submitted for examination, has an ethical Code of Practice on Investigations Involving Human Beings[10]. The findings reported in this dissertation are exempt from this Code of Practice as they fall into the following categories to which the Code does not apply:

- Research which involves only working from historical and literary databases and documents and does not involve working with 'live' participants;
- Consultation with colleagues, experts or other stakeholders about the preparation or progress of an investigation, where those people do not contribute to the actual findings;
- Work which is part of routine practices in professional contexts or service evaluation.

Obtaining written informed consent is not always required. Given that no detailed interviews or research where personal, sensitive or confidential data have been gathered or stored, no requirement arises in this research to comply with the Data Protection Act. For general surveys and informal conversations, where no personal data are gathered and personal identifiers are removed from the data, information about the research project and the identity of the researcher have been provided to participants throughout the project, detailing the nature and

[10] http://www.strath.ac.uk/media/committees/ethics/Code_of_Practice_Oct_2009.pdf

scope of the study, and how the knowledge collected would be analysed (in aggregate anonymous form).

The research has evolved through a series of incremental steps, adopting a diverse mix of techniques. Individuals and organisations who came in contact with the study were always informed (sometimes in writing, sometimes verbally) that their contribution was being solicited in relation to the specified research context and were directed for more information to the relevant websites, initially KAF and then OAM, as well as to previous published papers relating to the research, and that their contribution, where relevant, would be analysed directly or indirectly in aggregate and anonymous form to progress the study in question.

The core aspects of the research as presented in the thesis do not directly involve human beings as primary subjects, since the research dimensions are organisational and institutions ("Levels and units of analysis", Section 2.5.3.5) through the remote (web-based) auditing of web-based knowledge resources. Where these were not available, or could not be located, email enquiries and telephone exchanges became inevitable and necessary, however the subjects of the same were not data or information pertaining to the individuals engaged in the conversations, rather objective facts pertaining to the knowledge sharing practices of the organisations they work for.

In summary, although no formal written informed consent was obtained, due consideration has been given throughout the research to ethical issues, and the necessary steps including the ethical check list, and humans who tangentially in contact with the study provided informed consent through various exchanges, some of which are in writing (emails), and the University Code of Practice in Investigation on Human Beings was adhered to.

Furthermore, ethnographic research can be exempt from written informed consent when these can contribute to distort the responses of participants (Fluher-Lobban, 1994 cited in Murphy & Dingwall, 2007). However every individual approached in formal emails and correspondence, or informal conversations during the course of the study was informed *a priori* that the nature of the interaction was motivated by the need to gather information and insights for the purpose of this scholarly investigation.

A tentative 'Research Design Map' that visually plots the techniques adopted against a theoretical scientific method framework[11] is provided in Fig. 11b.

2.7 CONCLUSION

> *"Design research is in a state of flux. The design research landscape has been the focus of a tremendous amount of exploration and growth over the past five to 10 years. [...] It is currently a jumble of approaches that, while competing as well as complementary, nonetheless share a common goal: to drive, inspire, and inform the design development process. Conflict and confusion within the design research space are evident in the turf battles between researchers and designers."* (Sanders, 2008)

Chapter 2 of this thesis provides an overview and justification for the research paradigm, methods and techniques adopted in this research, as well as a rationale for their combination and validation, and concludes with ethical considerations.

[11] Modelled on Beech, in Aytes

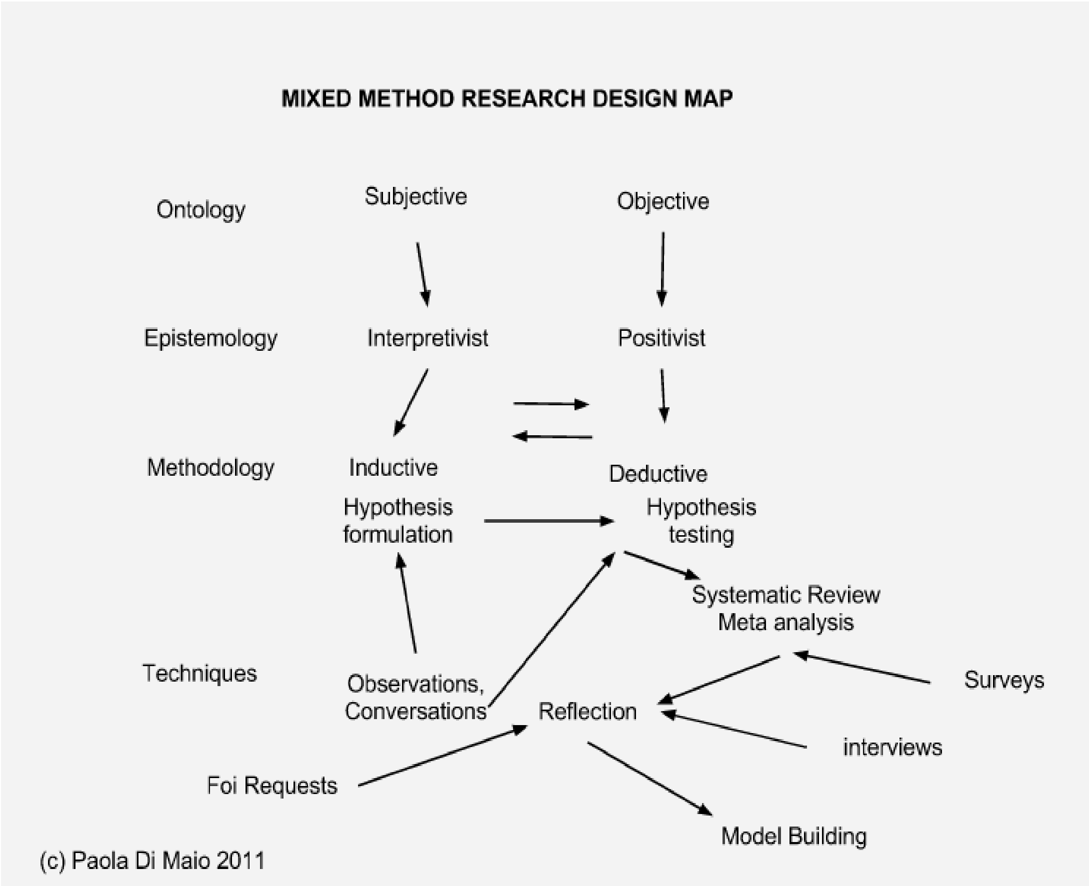


Fig.11b: Mixed Methods Research Design Map (Di Maio, 2011)

"Paradoxes can be viewed as opposing forces that are influenced by fluxes in realizing or recognizing a system"

Sauser and Boardman

CHAPTER 3

LITERATURE REVIEW

OVERVIEW OF THE SECTION

The main goal of this section is to introduce key conceptual constructs, definitions and knowledge acquired non-empirically, mostly via critical literature review, as well as to present related exploratory work, such as morphological analysis, that aims to somehow provide some rationale for such an exploration.

3.1 WHAT IS KNOWLEDGE (AGAIN)?

It is not uncommon to hear the occasional information scientist remark, even today, that 'there is no difference between information and knowledge' in disregard of decades – if not centuries, even thousands of years if one takes into account the long-term scholarly perspective – of key scholarly contributions in knowledge sciences (Zhang & Siekmann, 2007). The history of knowledge, and its place in human evolution, is a fascinating field of investigation going back as far as human history itself. Knowledge in philosophy has been considered 'justified true belief' (JTB[12]), and much has been said in relation to knowledge and truth (Russell, 1954; Russell, 1913), however in scientific discourse knowledge is understanding or 'true' belief derived from **evidence of some fact**:

> *"Truth is logical and parsimonious consistency with evidence and with other truth. Evidence is any and all perceived circumstances."*
> (Holz, 2003)

Knowledge is inherently 'polymorphic' (Aamodt & Nygård, 1995), and as such it can be characterised by many different concurrent dimensions and worldviews. Several analytical frameworks for knowledge sharing exist in literature, each with its own set of limitations. Knowledge has also:

> *"... intrigued the world's greatest thinkers from Plato to Popper without the emergence of a clear consensus."* (Grant, 1996)

---

[12] http://plato.stanford.edu/entries/knowledge-analysis/

Table 5 summarises the main perspectives, and the implications of each for Knowledge Management and Knowledge Management Systems (Alavi & Leidner, 2001).

Table 1. Knowledge Perspectives and Their Implications

Perspectives		Implications for Knowledge Management (KM)	Implications for Knowledge Management Systems (KMS)
Knowledge vis-à-vis data and information	Data is facts, raw numbers. Information is processed/ interpreted data. Knowledge is personalized information.	KM focuses on exposing individuals to potentially useful information and facilitating assimilation of information	KMS will not appear radically different from existing IS, but will be extended toward helping in user assimilation of information
State of mind	Knowledge is the state of knowing and understanding.	KM involves enhancing individual's learning and understanding through provision of information	Role of IT is to provide access to sources of knowledge rather than knowledge itself
Object	Knowledge is an object to be stored and manipulated.	Key KM issue is building and managing knowledge stocks	Role of IT involves gathering, storing, and transferring knowledge
Process	Knowledge is a process of applying expertise.	KM focus is on knowledge flows and the process of creation, sharing, and distributing knowledge	Role of IT is to provide link among sources of knowledge to create wider breadth and depth of knowledge flows
Access to information	Knowledge is a condition of access to information.	KM focus is organized access to and retrieval of content	Role of IT is to provide effective search and retrieval mechanisms for locating relevant information
Capability	Knowledge is the potential to influence action.	KM is about building core competencies and understanding strategic know-how	Role of IT is to enhance intellectual capital by supporting development of individual and organizational competencies

Table 5: Knowledge Perspectives (Alavi & Leidner, 2001)

Within the scope of this research, knowledge is considered in a pragmatic context of its role in open, networked information technologies - namely the internet and the web - as they provide the most efficient and economical environment for the sharing of explicit knowledge, which combines data, information and know-how, provided appropriate formalisms, conventional notation and common web standards and good practices are followed (Fikes et al., 1991). Sharing knowledge on the web presents a new, different set of challenges that did not exist before open and distributed knowledge systems and corresponding architectures became widely adopted.

Thanks to web-based knowledge sharing artefacts and techniques, many of the physical limitations that constrained knowledge flows in the pre-digital age are no longer relevant, and knowledge sharing acquires innumerable new dimensions:

> *"In a pervasively networked world, individuals are part of intersecting networks of interest and communities of practice. Knowledge becomes tangible as digitized content, as context that can be digitally shared, and through direct and indirect interactions. Knowledge can be created by asking a question and watching the responses provoke cascading conversations, responses, and interactions among network participants. The networked world continuously refines, reinvents, and reinterprets knowledge, often in an autonomic manner."* (Norris et al., 2003)

The term 'knowledge' and derived concepts, such as 'knowledge management' (KM) can sometimes be misunderstood and even misinterpreted (Tuomi, 1999). No single definition applies, and different epistemological stances yield different definitions (Alawi & Leitner, 2001). Two complementary definitions used in our research are:

> *"Knowledge is the sum or range of what has been perceived, discovered, or learned."* (Schubert et al., 1998)

and

> *"Knowledge is the result of a cognitive process."*
> (Fahey & Prusak 1998; Tuomi 1999)

Knowledge in this research is intended as mental state (to be in the know), a cognitive stream resulting from all information processing and a high level cognitive function.

Specifically, this research considers 'knowledge artefacts' as the resources deriving from the codification of explicit knowledge (Holsapple, 2004).

Since the eighties, when a shift from symbolic programming to knowledge-based architectures first took place, knowledge is to be understood as an architectural component (Newell, 1980). Knowledge is now a vital *layer* in intelligent, networked systems architectures (Batet et al., 2007), Fig. 12.

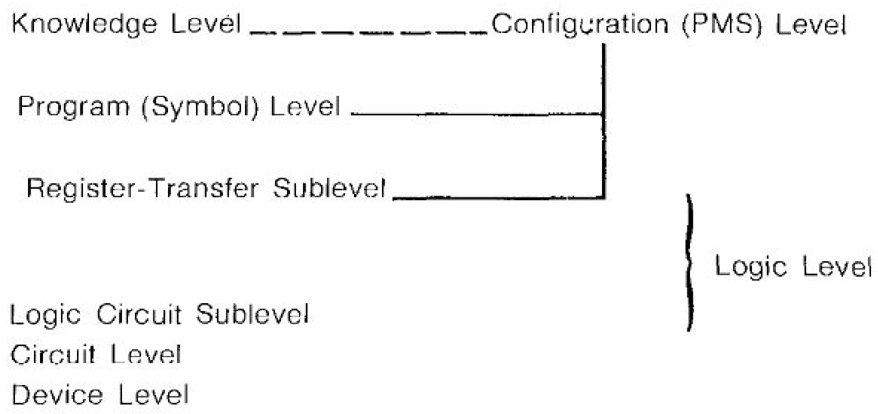


Fig. 12: Knowledge Level in Systems (Newell, 1980)

Researchers may not always unanimously agree on how to firmly distinguish between knowledge and information (Wang & Noe, 2010), however it is commonly accepted that organisations leverage knowledge rather than being simple information-processing entities (Dosi & Marengo 1994; Fransman 1994; Nonaka & Takeuchi 1995). Such a distinction is provided extensively in related literature (Ackoff 1989; Bellinger 2004; Sveiby 1997; Davenport & Prusak 1998), some excerpts reported below:

"Some of the present confusion concerning how to do business in the knowledge era would probably be eliminated if we had a better understanding of the ways in which information and knowledge are both similar and different.

The widespread but largely unconscious assumption that information is equal to knowledge and that the relationship between a computer and information is equivalent to the relationship between a human brain and human knowledge can lead to dangerous and costly mistakes." (Sveiby, 1997)

and

"Knowledge is neither data nor information, though it is related to both, and the differences between these terms are often a matter of degree. Confusion about what data, information, and knowledge are—how they differ, what those words mean—has resulted in enormous expenditures on technology initiatives that rarely deliver what the firms spending the money needed or thought they were getting." (Davenport & Prusak, 1998)

Further differences can be stated as follows:

"Knowledge should be viewed as an integrated totality, and it is the tightly connected network of interrelated sub-components that gives knowledge its power of data interpretation, information elaboration, and learning. Second, knowledge as the outcome of a learning process links knowledge to its potential use. Learning, as a process, is always related to a purpose, a way to make future use of what is learned. (Aamodt & Nygård, 1995)

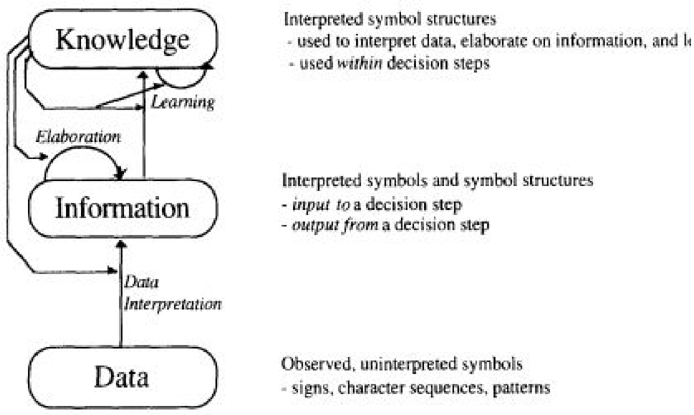


Fig. 13: Data Information and Knowledge (Aamodta & Nygård, 1995)

Figure 13 (Aamodt & Nygård, 1995) illustrates diagrammatically a typical view of the data-information-knowledge distinction. The classical data information knowledge continuum often cited above, can be extended to include intelligence, as shown in Fig. 14 (Tham et al., 2002).

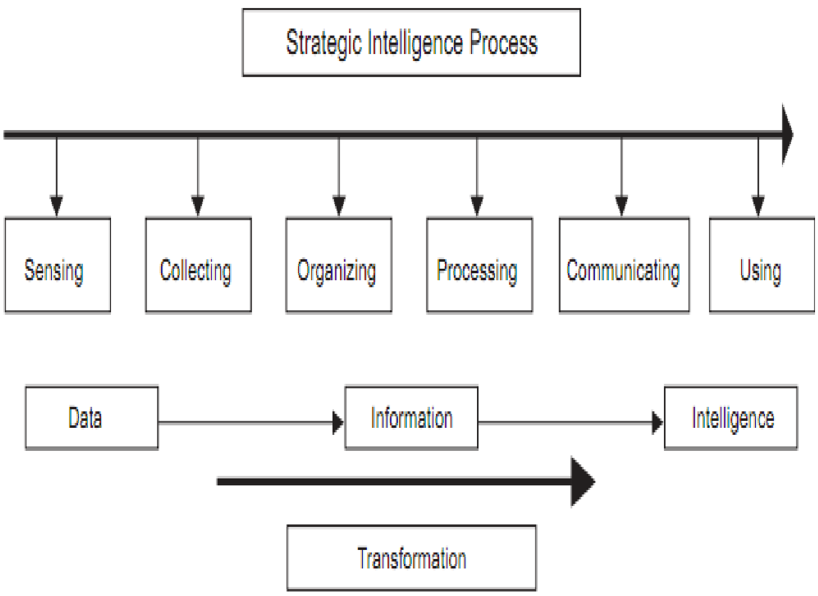


Fig. 14: From Data to Intelligence (Tham et al., 2002)

For completeness, it is noted that the data-information-knowledge hierarchy illustrated above can be turned around and viewed from the reverse perspective:

> *"Information can be created only after there is knowledge, and data emerges as a by-product of cognitive artefacts that assume the existence of socially shared practice of using these artefacts."* (Tuomi, 1999)

3.1.1 Explicit, Shared Knowledge

This research is concerned primarily with 'factual organizational knowledge' consisting of technological and market know-how, which is an accumulation of structured information and is transferable in formalized processes (Richter & Vettel, 1995).

This corresponds largely to the notion of 'Explicit Knowledge':

> *"Explicit knowledge is systematic and easily communicated in the form of hard data or codified procedures. It can be articulated in formal language including grammatical statements. This kind of knowledge can thus be transmitted across individuals formally and easily: for example, technological knowledge."* (Pan & Scarborough, 1999)

Also:

> *"It can be said that organisations share explicit knowledge, as opposed to tacit knowledge which is held by single individuals."* (Itami, 1987)

Technical knowledge is best shared using explicitly codified artefacts (Smith, 2001). Knowledge becomes an important organisational resource when it is shared (Davenport & Prusak, 1998).

Different factors influence knowledge sharing behaviour from different perspectives:

- trust and mutual influence (Nelson & Cooprider, 1996)
- organizational culture (Jarvenpaa & Staples, 2000)
- extrinsic motivators, social-psychological forces, and organisational climate (Bock et al., 2005)

Knowledge sharing can be defined as the process of communication in which members exchange and discuss knowledge through discussion and the internet to increase the knowledge value (Hendriks, 1999) however there is a polarisation and possibly a whole spectrum of mechanisms and modalities for knowledge sharing.

At the one end knowledge sharing is encouraged by the indiscriminate opening up and lowering the barrier to access knowledge resources, as advocated by the 'all source knowledge' approach, and allegedly supported by policies such as 'Open Access', and at the other end another knowledge sharing is highly directed and enforced via channelling, as supported in practice by knowledge transfer agreements which are designed to constrain and segregate knowledge flows according to strict economic and political organisational information agendas.

3.1.1.1 Knowledge sharing as enabler for re-use

Despite increasing quantities of valuable 'knowledge' being produced everyday in all fields of research and practice, the ability of systems and personnel to manage and apply existing knowledge to decision making and problem solving, is not easily qualified nor quantified, and limited on the one hand by the physical, psychological and cognitive capabilities (of human and machine), and on the other hand by the configuration of the information systems and social and organisational infrastructures in place which directly or indirectly regulate the knowledge flows.

A typical 'knowledge management lifecycle' breaks down the process into activity sets, for example discovery, development, sharing and organisation (Sang & Soongoo, 2002):

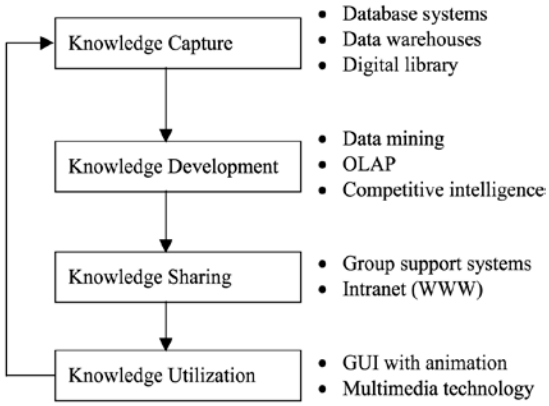


Fig. 15: Typical KM Phases (Sang & Soongoo, 2002)

Modelling such activities as discrete functions however is arbitrary, as in reality they are interdependent, blend together and often take place simultaneously. Knowledge 'Re-use' (KR) which within the classical knowledge management lifecycle is considered a 'phase' (Tao et al., 2005), actually is an underlying function of most other knowledge management activities, and should be modelled as such.

Loosely adapted from the software domain (Dusink & Katwijk, 1995), the notion of 're-use' can be defined as the systematic application of existing artefacts, or the physical incorporation of existing artefacts, during the process of building a new system, whereby artefacts are identified as pieces of formalised knowledge that can contribute to an engineering process, such as designs, transformation rules, as well as e.g. descriptions of how to detect abstract data types.

'Tacit' knowledge can be also a contributing factor to re-use (Harsh, 2008); for the purpose of this research we focus on knowledge re-use that takes place mainly through 'explicit' knowledge, and that as such is codified via explicit 'knowledge representation' formalisms and artefacts.

A Theory of Knowledge Re-use (Lynne, 2001) characterises four distinct typologies of situations: shared work producers, who produce knowledge they later re-use; shared work practitioners, who re-use each other's knowledge contributions; expertise-seeking novices; and secondary knowledge miners bringing to light the importance of distinguishing 'context' to all aspects of re-use. More recently, re-use is considered part of a 'knowledge transformation' process (Carlile, 2002), and viewed as a cycle, adopting an 'integrated framework' (Fig. 16) incorporating syntactic, semantic and pragmatic dimensions of knowledge transformation as a strategic and tactical approach to knowledge integration across disciplinary boundaries.

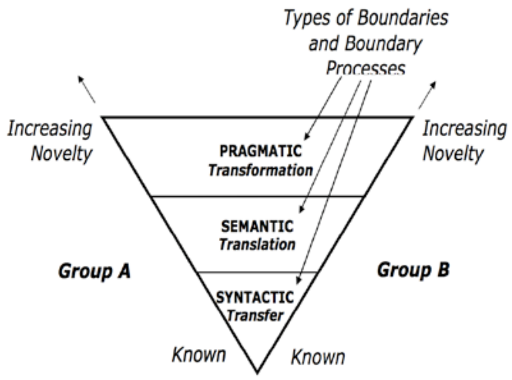


Fig. 16: Knowledge Integration Framework (Carlile, 2002)

The separation of 'semantic' and 'conceptual' layers for knowledge and its representation is already an established concept in the domain of operational simulations (Tolk, 2003); in a paragraph that follows the more 'pragmatic' dimensions of the research space are explored further.

3.1.2 Knowledge Codification

Structured, pragmatic knowledge is best represented explicitly (such as explicit knowledge models), and to be shared it should be represented using appropriate formalisms and notation (Fikes et al., 1991). Codification is defined as:

> *"The process of conversion of knowledge into messages that can be processed as information [...] The process of codifying knowledge entails three distinct but related steps: creating models; creating languages; and creating messages."* (Cowan & Foray, 1997)

Codified knowledge can be characterised as 'information-like' and objectified (Kogut & Zander 1992; Spender 1996); a conversion necessary to processing and management in the form of sets of identifiable rules and relationships, that can be transmitted to decision agents (Kogut & Zander, 1992).

> *"Codification contributes to reduce complexity when confining the description and analysis of a domain into what can be encoded."*
> (Boisot, 1998)

The benefits of codification can fall into two categories:

- Economical, whereby codified knowledge yields convenience as per traditional information models.
- Cognitive and pedagogical, emphasising the benefits of codification to facilitate learning (Prencipe & Tell, 2001).

3.1.3 Knowledge Sharing on the Web

At least two research directions have motivated (and funded) research in knowledge sharing on the web in recent years; one is the need to provide knowledge representation mechanisms that are 'shared', that is, which are commonly accessible and understood, so that knowledge can be re-used economically and web-based knowledge flow is optimised. Another is to provide mechanisms for artificial intelligent agents to perform reasoning functions. The latter can sometimes conflict with the former when generating intellectual property rights by protecting knowledge with patents, becomes a primary purpose for doing research. It is worth remembering that ontologies in general are devised to facilitate the sharing of knowledge, whether among restricted or open networks of agents, whether these are human or artificial. Generally, for knowledge to be re-used, it needs to be shared.

Ontologies provide sets of parameters for knowledge sharing, and rely on the assumption that all the constructs and artefacts are shared. Across large domains, taking into account the diversity of disciplines, paradigms, axioms, vocabularies, uses, standards, practices, and despite many years of knowledge management research and practice, knowledge sharing good practices (such as the adoption of shared vocabularies) are still not diffused, or only marginally so, outside the relatively small knowledge engineering community. A knowledge audit can provide an overview of the (explicit) knowledge, and its qualitative and quantitative characteristics, helping to identify the location where it resides, as well as other information such as people and roles involved in their creation and maintenance, and other organisational processes associated with it. The core principle central to knowledge sharing on the web however, is that every knowledge resource should have a URI (Uniform Resource Identifier):

> *"The essential process in webizing is to take a system which is designed as a closed world, and then ask what happens when it is considered as part of an open world. Practically, this effect on a computer language is to replace*

the names/tokens/identifiers for URIs. Thus, where before reference could only be made to something in the same document/program/module one can with equal ease make reference to something in a different one somewhere in that abstract space which is the Web." (Berners-Lee, 1998)

Also:

"There are a number of important aspects of URIs with respect to making the Web work: the definition of URI scheme, the convention that the first element is the server at which the resource named can be found, etc. From the Knowledge Representation point of view, however, there is another feature that is very crucial: any resource on the Web can be given an identifier, and any Web server pointed at that name will retrieve the same underlying representation of the resource." (Hendler & Harmelein, 2008)

The auditing component of this research, presented in Chapters 5 and 6, uses URIs to locate Open Access resources and assign a score to a degree of 'sharedness' of project knowledge, and in the final recommendations formulated in Chapter 7 it is suggested that Open Access policies provide guidelines to technical artefacts such as URIs to facilitate and lower the costs of knowledge sharing as well as of Open Access monitoring.

3.1.4 Semantic and Pragmatic Dimensions of Knowledge Sharing

New technologies, in particular web-based technologies, open up new prospects for knowledge management, especially re-use and learning; virtual environments, characterised by the lack of time/space boundaries, dissolve the physical barriers to knowledge transfer, therefore supposedly facilitating pervasive knowledge access, re-use and transformation. For example, what is defined as 'social learning', intended not only as individual cognitive processes but also as collective learning processes that include not only knowledge flows, but also interactions between actors and processes of negotiation and struggle (Stewart & Williams, 2005). In practice, however, other challenges come into existence:

- while physical barriers to knowledge access may be reduced by the pervasiveness of the web, cognitive barriers still exist, and potentially become greater as the amount and degree of knowledge increases.
- the inherent nature of web-based technologies, and the chaotic and virtually unbounded environments such as the web, constrain further the 'usability', hence the 'utility' of knowledge published on the web by increasing the complex challenge of sorting and retrieving meaningful, accurate and up to date data which is relevant to our queries, and to keep up and filter 'intelligence' from the vast amount of data that is being produced.

Additional obstacles to effective knowledge re-use may be partly due to inherent limitations of existing technologies and technology development patterns, and partly by socio-technical factors, such as organisational and social structures.

The Semantic Web

In order to address these, and other challenges, the 'semantic web' (SW) was envisioned (Berners-Lee et al., 2001). Semantic web technologies have been in development for over a decade, with the aim to support and leverage the logical relations that underlie knowledge represented in non-structured form. The semantic web technology corpus, which includes knowledge, techniques and tools, is devised directly or indirectly with the intent to maximise knowledge re-use.

It consists essentially of artefacts, such as vocabularies, explicitly coded to represent relations, properties, hierarchies and classes, represented and exposed in such a way that they aim to be accessible and re-usable. Formalised vocabularies are in fact considered the equivalent of 'simple ontologies'. Ontologies, in various degrees of formalisation, can be used to support knowledge management functions (Kitamura et al., 2004), and semantic web frameworks in support of Knowledge Re-use have been explored to some extent in scholarly literature (Stojanovic et al., 2002) but have not yet been adopted by mainstream and commercial implementations.

Current 'semantic web' standards issued by the governing consortium W3C are prescribed in the form of RDF and OWL, which are respectively the recommended formats for vocabularies and ontologies, and SPARQL, the corresponding query language. These recommendations represent 'attempts' to standardise novel approaches to semantic web, although they may not be the only way of supporting semantic capabilities on the web.

Semantic technologies have been in development for some years, yet limited measurements of advances in terms of their impact on knowledge re-use are available; as much of the current efforts are devoted to overcome core infrastructural challenges ('plumbing').

This further supports and justifies the investigation into how a pragmatic dimension contributes to the development of web-based technologies in relation to KR. The more the semantic web takes shape, the more the lack of adequate instruments to capture and represent social and human dimensions, and practical context, become evident. The role and impact of human factors in the development of the semantic web are still largely out of the scope of current SW research, prompting central questions that are often overlooked in technical debates.

A 'pragmatic web' (PW) has evolved in recent years to prioritise a different set of challenges and an entirely new perspective: although at 'physical' level there is only one web (that we know of) – the open 'network of networks' that links computers via http and uses IP protocols and hypertext (HTML) – the 'pragmatic web' approach, as described in literature, (Schoop et al., 2008) focuses on the intent to explicitly research, support, interpret and model 'context' to knowledge on the web. Technically, and with some limitations, 'contextual' information can be represented on the web with current SW artefacts and techniques (Reza-Tazari 2003; Guha et al. 2004). Pragmatically, however, challenges to knowledge re-use, and relevant contextual dependencies, are not merely technical, but belong to the realm of social and organisational systems design and management, and extend well into the boundaries of what is designated as 'policy' management.

Although it is impossible to pinpoint exactly where the semantic and pragmatic web overlap and where they actually differ, in the sense that semantic web research is increasingly adopting a stronger 'user-centric' approach, the focus of the pragmatic web as envisaged so far seems to emphasise:

1. Placing the human-to-human interaction, together with its social and behavioural complexity at the heart of web-based systems design and development.
2. The exploration of alternative frameworks and assumptions that can more adequately capture such complexities, and that are not restricted to RDF/OWL as they are known today.

The distinction between syntactic, semantic and pragmatic 'dimensions' was first made as part of the linguistic and communication theory (Morris, 1946).

Morris defined pragmatics as *"dealing with the origins, uses, and effects of signs within the total behaviour of the interpreters of signs"*; semantics as *"the relations between signs and the objects they signify"*, narrowing semiotic study to the strict literal meaning of signs and propositions. Syntactic *"concerns the formal relations between signs themselves, further narrowing semiotic study to the logical and grammatical rules that govern sign use."* Morris based much of his work on earlier studies of semiotics by Charles Pierce, and the 'social behaviourism' of Dewey and Mead. When considering the web as a virtually unbounded and chaotic information environment, structure and meaning acquire relevance mainly thanks to 'context', which is defined as the dynamic combination of factors and conditions that surround and affect an object or an event. On this premise, the pragmatic dimension is critical to information communication in that it allows the understanding and modelling of 'behaviours' and human 'experience', which in turn affect systems functionality and design.

The complementarity of the 'pragmatic approach' to the 'semantic web' is consolidating around a small and diverse community of interest, with the intent to develop environments and infrastructures designed primarily to support the automation of human-to-human web-based knowledge exchange, with particular emphasis on rule-based reasoning and intelligent applications based on natural language. While much of the 'semantic web' is still under construction, the pragmatic web is pointing to a further shift toward an adaptive and context sensitive real-time communication environment capable of supporting intelligent and dynamic applications. Web mediated human-to-human interaction is one of the purposes of pragmatic web research. Human-to-human interaction has so far been envisaged as essential to knowledge re-use.

In a study carried out at NASA Jet Propulsion Laboratory, (Majchrzak et al., 2000) all six case analyses of re-use for innovation, primarily rely on human-to-human contact for learning to occur - to find the right individuals with the right solutions, to query the individuals to assess the applicability and limitation of the solution, and to physically manipulate the solutions to personally assess their appropriateness. The importance of 'social knowledge networks' in knowledge re-use in the context of the architecture/engineering/construction industries has already been studied to some extent (Williams et al., 2005). One of the key issues emerging from research is that 'internal' knowledge - knowledge existing within a role or within an organisation - is usually complementary and necessary to make use of 'external' knowledge (Menon & Pfeffer, 2003). Since the web is the largest knowledge repository that has ever existed, to create support for knowledge re-use in practice, taking into account the social knowledge networks dynamics already under study elsewhere, points to new requirements to guide the development of web technologies and infrastructures.

A preliminary conclusion from analysis of existing work in this area points to the need for artefacts capable of extending the capability of the web from an instrument to disseminate and expose data and information, to an instrument capable of disseminating knowledge, know-how, and intelligence that is applicable to the evolution and advancement of society at large.

3.2 KNOWLEDGE ENGINEERING

Knowledge Engineering (KE) was introduced as a discipline in the eighties, as a first scholarly attempt to formalise the techniques to represent the 'knowledge level' discussed in a previous section. KE was defined in 1983 by Edward Feigenbaum and Pamela McCorduck as follows:

"KE is an engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise." (Fox, 1984)

Recently, in an editorial for the 25th anniversary of KER Journal special issue (Fox, 2011) it was proposed that:

"The engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise."

More succinctly, Knowledge Engineering can be considered the methodical construction of knowledge bases,[13] which is the preferred definition in this work. Different KE techniques apply, mostly consisting of developing and using explicit, symbolic representations of knowledge to carry out a variety of tasks that require some level of cognition. KE is a branch of 'cognitive science', which in turn has 'various definitions' including a human-centred meaning such as:

"*the study of human intelligence and of the symbol-processing nature of cognition'"* (Norman, 1980)

and a more neutral one such as:

"*the scientific project dedicated to understanding the processes and representations underlying intelligent action which accommodates insights from both artificial and natural intelligence."* (Fox, 2011)

13 www.cs.utsa.edu/~bylander/cs5233/notes/logickbhandout.pdf

As the web is the largest publicly accessible knowledge base that ever existed – albeit relatively unstructured and self-generated – knowledge engineering techniques are required to support knowledge representation for the purpose of sharing on the web (Studer et al., 1998). Since Knowledge Engineering first provided system developers with techniques to develop functional knowledge bases, 'cognitive engineering' (CE) has also become a field of practice in its own right (Vicente 1999; Hollnagel 2003). Today, the expression 'Knowledge Based Engineering' (KBE) is used by some to indicate a knowledge-based technique that uses a rule-based design approach similar to the expert systems generated in the artificial intelligence domain.

> *"Rather than simply a knowledge based system; however, knowledge based engineering focuses on the tasks of system design largely dependent on the geometric model.*
> *Knowledge based (expert) systems from the artificial domain are systems that can perform "a task normally done by an expert or consultant and which, in so doing, uses captured, heuristic knowledge."* (Cooper et al., 2009)

This invokes, and to some extent leverages, principles from both KE and CE to devise a knowledge auditing template and a reference model for knowledge sharing in Systems Research, described in the final recommendations of this thesis.

3.2.1 Knowledge Modelling

As mentioned in the earlier paragraph, one of the commonly adopted knowledge codification techniques is 'knowledge modelling'. Three models of knowledge sharing and re-use were proposed in the early days open networking protocols, from which the public and open internet and web based technologies in use today:

"1. The library model, in which bodies of formally represented knowledge are available as off the-shelf products, like books in a library. In this model, knowledge bases are designed artefacts, and the challenge is to make them available and reusable.

2. The software/systems engineering model. A standard approach to making software reusable is to decompose complex programs into modular pieces, and to provide a formal specification of the inputs, outputs, and functions computed by each piece, which also applies to knowledge sharing systems.

3. The reference model, typically used as an integration framework for the concepts in a domain and/or problem area that are common to the set of application tasks." (Fikes et al., 1991)

In the last twenty years, since the categories above were identified, research has evolved towards higher level knowledge representation frameworks, such as 'metamodels' (Becker et al. 2003; Holten 2000) and reference models (Geisberger et al. 2006; OASIS). Some schools of thought distinguish between a 'metamodel' and a 'reference model', for example as per the following definitions:

"A metamodel is a model of a modelling language. The metamodel defines the structure, semantics and constraints for a family of models."

(Mellor et al., 2004)

or

"A metamodel is a model of a language that captures its essential properties and features. These include the language concepts it supports, its textual and/or graphical syntax and its semantics (what the models and programs written in the language mean and how they behave)."

(Clark et al., 2008)

Elsewhere, a Reference Model is defined as:

> *"A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist."*
> (OAIS 2002; others)

It should be noted that in the modelling community there is no consensus as to what the actual differences between meta modelling and reference models are (Genova, 2009); for example in the UK, the MODAF metamodel is defined as 'a type of reference model':

> *"The MODAF Meta Model (M3) is the reference model that underpins MODAF. It, defines the structure of the underlying architectural information that is presented in the MODAF views."* (MODAF)

In summary, knowledge on the web is shared via knowledge models and corresponding artefacts. The explicit, shared and, to some extent, standardised codification and representation of these knowledge artefacts, for example URIs and the corresponding objects designed according to these models, such as knowledge schemas, are what makes (explicit) knowledge shared, accessible and 'reusable'. However, the adoption of these good practices 'in the real world' - for example outside of niche academic communities - is limited. Part of this research is devoted to audit publicly-funded systems engineering projects, to establish to what extent shared knowledge representation and codification and corresponding formalisms are adopted, and point to guidelines and good practices accordingly.

3.2.1.1 Principles of knowledge modelling

When we choose how to represent something, we are making design decisions. To guide and evaluate designs, objective criteria depending on the purpose of the resulting artefact, rather than based on a priori notions of naturalness or truth (Gruber, 1995). Some of the principles devised around 30 years ago, in the early days of knowledge and ontology engineering, still largely stand. They can be summarised as:

1. **Clarity**: should effectively communicate the intended meaning of defined terms. All definitions should be documented with natural language. [...]

2. **Coherence**: should be coherent, that is, it should sanction inferences that are consistent with the definitions. [...]

3. **Extendibility:** should be designed to anticipate the uses of the shared vocabulary. [...]

4. **Minimal encoding bias**: the conceptualisation should be specified at the knowledge level without depending on a particular symbol-level encoding.

An encoding bias results when representation choices are made purely for the convenience of notation or implementation. Encoding bias should be minimized, because knowledge-sharing agents may be implemented in different representation systems and styles of representation.

5. **Minimal ontological commitment**: should require the minimal ontological commitment sufficient to support the intended knowledge-sharing activities [...]

Principle 4, *minimal encoding bias*, states that ontology should be independent of its implementation, whereby the coding of the model using one formalism or other should not drive the modelling process.

Knowledge representation formalisms are known as 'standards' and sometimes formalisation can get in the way of the real purpose of a model, which is to communicate and make accessible its meaning and logic, for the purpose of being used and re-used.

3.2.1.2 Just enough knowledge modelling

'Just enough' approaches started to emerge when personal computing was merely a prospect, and 'structured' approaches to systems analysis and design promised to capture diagrammatic and schematic representation of essential aspects of systems components. This was a response to the growing demand for non-formal (mathematical) ways of expressing systems requirements and functionalities, and that would be more articulate than pure narrative description of the systems functionalities. That's when structured charts, data flow and data model diagrams, and data dictionaries started to come into use, with the aim to capture and represent 'what most counts' of design and modelling activities. The majority of information-centric systems today are designed to leverage knowledge expressed via natural language, where symbols and meanings (semiotics and semantics) need to be captured and represented adequately for these systems to function.

Domain and Knowledge engineering techniques discussed in the earlier sections in the last two decades have been adopted by 'Ontology Engineering', whereby ontologies are conceptual and semantic representations with different degrees of formalisation that capture and express knowledge. In this research 'ontology' is considered as a 'formalised model of knowledge', sometimes, also referred to as 'knowledge schema', where the knowledge in question has been codified and structured, although the development of an ontology is not within the immediate scope of this work.

The author is co-leading related efforts such as the development of conceptual Top Level Categories for the INCOSE System Science Workgroup[14] some ontology concepts and techniques are adopted in the recommendation section of this thesis. Since the 'knowledge layer' became a component of system architectures, ontology development methodologies have proliferated (Di Maio 2009; Di Maio 2011). This section provides an outline of the core notions that help define 'knowledge artefacts', as background knowledge to some of the recommendations provided in Chapter 7.

3.2.1.3 Knowledge artefacts

An ontology, as a formal model of knowledge, is defined by the boundaries that constitute it, such as conceptualisations, models, schemas, representations and frameworks. A formal knowledge model may take a variety of forms, but it necessarily includes a vocabulary of terms and some specification of their meaning, such as definitions, and an indication of how concepts are interrelated, which collectively impose a structure on the domain and constrain the possible interpretations of terms (Uschold & Gruninger, 1996).

When sharing project 'knowledge' in a technical context, or 'system knowledge', in addition to 'peer reviewed papers' which contain narratives, it can be helpful to use structured knowledge artefacts, adopting different degrees of formality as prescribed by ontology engineering. For completeness, below is a brief overview of some examples of common knowledge artefacts, such as vocabularies, concepts, relations and axioms, as they can all be used in producing formal models of knowledge.

---

[14] INCOSE System Science WG https://sites.google.com/site/syssciwg/projects/system-top-level-categories

Vocabularies: Encyclopaedias, dictionaries, thesauri and vocabularies are fundamentally lists of words and their definitions, which can include grammatical, phonetic and etymological annotations.

Thesauri are vocabularies where the semantic association between terms are mapped, while glossaries are alphabetised lists of terms with definitions usually appended at the end of documents or reports. Information systems adopt vocabularies to support design and documentation, 'data dictionaries' for example, are used to list the entries used in a database. Vocabularies are at the core of ontologies, to the point that sometimes they are referred to as being the ontology itself. They list terms that declare and represent every concept, relation, function and axiom; the more an ontology is formal, the stricter the definition of its vocabulary terms.

In an ontology, the vocabulary has more than one function: it serves as an index and a directory of content. Generic vocabularies can contain more than one definition for each term, but controlled vocabularies do not, as they allow only one definition per term and explicitly enumerated (numbered) terms, which must be unambiguous, and non-redundant.

Vocabulary creation is both an art and a science, which leverages principles of information and library science; the core notion however is to keep track of the terminology used in a research project, as well as to keep track of the discussions (the modelling choices) that lead to the adoption of such terminology.

The development of a standard vocabulary for the 'systems engineering' domain is currently underway in relation to INCOSE SEBOK[15] (systems engineering body of knowledge), in particular the collaborative authoring and publication effort via a 'wiki'.

Concepts: Concepts are fundamental to our ability to think, express, represent and communicate knowledge, however defining unambiguously and with certainty what constitutes a concept can be difficult. Concepts can correspond to things, but also to 'fuzzy clusters' of ideas and notions identified by words and related to a certain thing or subject. And even when referring to tangible things, concepts can be abstract, and difficult to capture.

The nearest techniques that can be compared to conceptual modelling are entity modelling, in database system design, or class modelling, in object-oriented technologies. Concepts can be broadly defined as cognitive artefacts that support categorisation and communication, and are necessary to support human and artificial thinking and reasoning. The purpose of ontologies and knowledge models is to make these concepts explicit and represent them so that they serve the intended goals.

Conceptual categories and thoughts are closely related to language. A concept model can be used to complement and extend a functional data model. Other key model artefacts include relations, particularly important in knowledge categorisation, and axioms, as they underpin 'rules', a form of knowledge representation.

The knowledge model proposed in the recommendation section of this thesis is an example of a knowledge sharing artefact that consists at this stage of concepts and vocabularies.

---

[15] SEBOK, http://www.bkcase.org/

A more complete knowledge model for systems knowledge is tackled as a collective effort in the context of related professional activities referenced earlier.

3.2.2 Domain Engineering

The research question, and the central hypotheses in this dissertation can be generalised and applied to any 'domain', however this study targets specifically 'systems research', since it has been devised to answer the research question in relation to a particular 'knowledge domain'. This coincides with another important parallel strand of work: the SEBOK Project, (Systems Engineering Body of Knowledge) to which this research has contributed an exercise in automated text extraction, reported separately (Di Maio, 2011). The auditing instruments presented in Chapter 5 of this thesis, as well as the recommendations are specifically modelled to audit knowledge in systems research, in particular, systems engineering (SE) research, due to the nature of the project the research was initiated.

This section introduces the notion of 'domain', 'domain engineering', and 'domain analysis' as precursors to Knowledge and Ontology Engineering.

Before the 'knowledge layer' shaped information systems architectures as we know them today, systems engineering used to model 'domains' intended as 'a universe of discourse':

> *"An area of human activity (business, industry, service), characterized by its own set of professional terms and otherwise pragmatically separable from but possibly interfaced to (incl. overlapping with) other domains."*
> (Bjorner, 2007)

A domain can also be defined as an area of knowledge scoped to maximize the satisfaction of the requirements of its stakeholders, including a set of concepts and terminology understood by practitioners in that area, and including knowledge of how to build software systems (or parts of software systems) in that area.

> *"Domain Engineering is the activity of collecting, organizing, and storing past experience in building systems or parts of systems in a particular domain in the form of reusable assets (i.e. reusable work products), as well as providing an adequate means for reusing these assets (i.e. retrieval, qualification, dissemination, adaptation, assembly, etc.) when building new systems."* (Czarnecki & Eisenecker, 1999)

3.2.2.1 Domain and knowledge analysis

Domain analysis is one of the activities used in Knowledge Modelling, defined as

> *"the activity of identifying objects and operations of a class of similar systems in a particular problem domain."* (Neighbors, 1980)

It is also noted that:

> *"knowledge should be represented and analysed on several different levels simultaneously."* (Tansley & Hayball in Kingston, 1994)

One of the outcomes of this doctoral research is a set of guidelines and recommendations that target knowledge sharing problem at different levels, including the adoption of shared knowledge models, and knowledge objects, which have been modelled and codified using standard domain and knowledge modelling techniques already discussed.

3.3 KNOWLEDGE SHARING – A PERSPECTIVE ON POLICIES

Open Access Initiatives are being promoted internationally across research institutions for the purpose of increasing knowledge flows, and facilitating knowledge sharing and re-use (KS/R):

> *"The practice of self-archiving and open-access repositories - both aimed at enhancing knowledge sharing and free circulation of socially relevant information - is becoming an authoritative alternative, when not a substitute, for traditional journal publications."*
>
> (Jankowski, cited in Calise et al., 2010)

However, some of the scholarly literature exploring future directions for knowledge sharing research has failed to identify 'knowledge sharing policies', such as Open Access, as a research area that needs further investigation. This section addresses this knowledge gap, by identifying relevant studies and by making the case for knowledge sharing policies to be included in the 'future research directions' for this critical field of inquiry.

Enablers

Two key enablers for knowledge sharing (KS) across scientific and research communities are:

a) Open Access initiatives (Calise et al., 2010)

and

b) Knowledge sharing artefacts and models for web based and networked technologies (Fikes et al., 1991)

Given the multiple disciplinary and methodological perspectives in Knowledge Sharing, no single state of the art review can be dispensed from inherent limitations. An example is provided in a recent 'narrative' literature survey (Wang & Noe, 2010) which provides a basic categorisation of knowledge sharing literature namely organisational, interpersonal and motivational, summarised in the first three columns of Table 6. The review omits some other existing research perspectives, such as the regulatory/policy aspect of knowledge sharing, which is added as a contribution from this research as the fourth column.

It is a known fact that narrative reviews can result in a type of 'cognitive bias', also referred to as 'cognitive algebra': where the studies included in the review are selected to support sets of pre-constituted opinions, which inevitably end up being reflected in the conclusions (Valentine et al. 2010; Polkinghorne, 2007). An additional dimension in the column 'regulatory/policy' is differentiated from the 'organisational' aspects in that the latter refer to 'within the firm's boundary' (the organisation), while policies are applicable to entire sectors or institutions at national and even international level, and therefore exceed the organisational boundary.

ORGANISATIONAL	INTERPERSONAL	MOTIVATIONAL	REGULATORY/POLICY
culture and climate	team characteristics and processes	beliefs of knowledge ownership	mandates
management support	diversity	perceived benefits and costs	actual benefits and costs to the wider community (public funding)
rewards and incentive	social networks	i/p trust and justice	compliance of processes with mandates
structure		individual attitudes	norms

Table 6: Areas of Emphasis (Wand & Noe, 2010) modified by Di Maio, 2011

Knowledge sharing research can be tackled from many different angles, and no single approach can ever be fully comprehensive and exhaustive, yet limitations of any given approach should be acknowledged.

In a study that looks at the types of agreements, rules of engagement and managerial practices adopted by the parties, specific knowledge-sharing activities are identified, as well as five primary contexts that can affect such successful knowledge-sharing implementations (Cummings, 2003):

- relationship between the source and the recipient
- form and location of the knowledge
- recipient's learning predisposition
- source's knowledge-sharing capability
- broader environment in which the sharing occurs

However neither the policies (organisational as well as sector wide instruments) nor how these policies mandate the adoption of the technical implementations (artefacts such as vocabularies, diagrams, rules, specifications etc.) are included in the exploration:

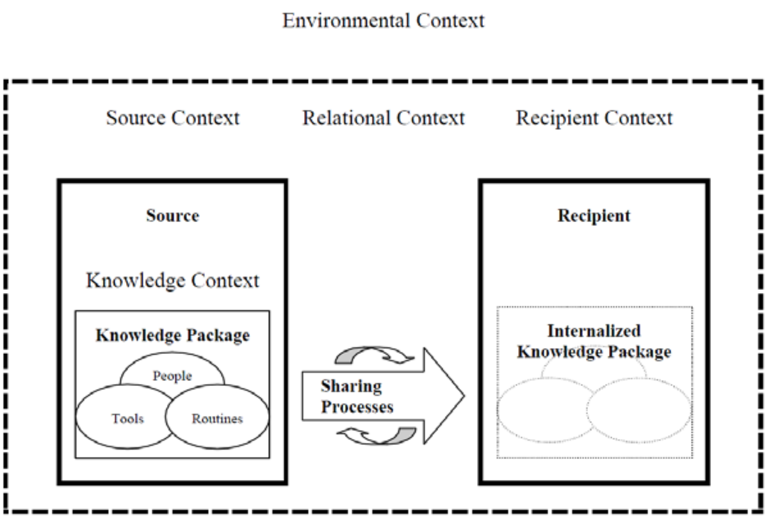


Fig. 17: Primary Knowledge Sharing Contexts (Cummings, 2003)

In related work, various areas which need policies need to be developed (Cohendet & Meyer-Kramer, 2001) namely:

· accessibility of knowledge: driving factors and conditions, processes and their impacts;
· communities and their dynamics;
· cognitive structures and content of knowledge;
· role of knowledge for macro-economic growth;
· the importance of technological breakthroughs vs. comprehensive diffusion of knowledge;
· the role of public infrastructures in a broad sense
(e.g. universities, R&D landscape, institutional connectivity, standards, etc.);
· policy rationales (e.g. co-ordination failure) and political governance.

Knowledge sharing policies and initiatives have been developed frantically in the last two decades, approximately coinciding with the increased adoption of the internet, however the adoption and implementation is largely driven by the 'practice', especially in the case of grassroots movements of pioneering academics and institutions from which much of the current efforts originate (Suber, 2009).

However, based on the gaps in literature identified above, a strong case can be made for policies to be included in the future research direction for knowledge sharing research.

Figure 18, from Wang and Noe, has therefore been modified accordingly.

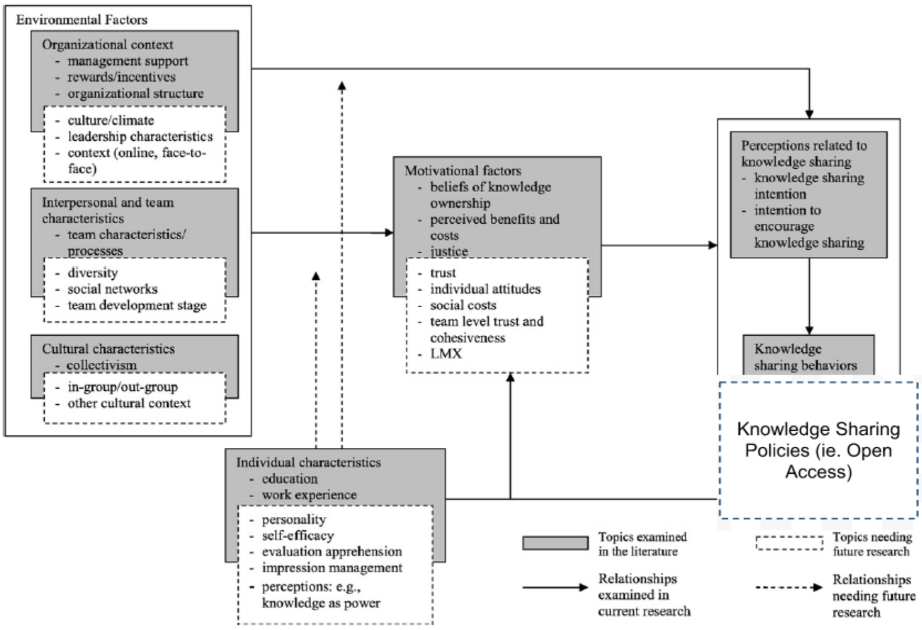


Fig. 18: Knowledge Sharing Future Research Directions (Wang & Noe, 2010), updated by Di Maio

3.3.1 Open Access, Pointers to Literature

It is not in the scope of this study to discuss 'what is Open Access', as extensive and comprehensive references exist, some of which are referenced and summarised in the following paragraphs.

In Chapter 4 a short discussion introduces the relevant question pertaining to Open Access policies in relation to the scope and context of this research.

Two relevant points are however reiterated:

1. Open Access policies, also known as 'mandates', are generally distinguished into two kinds: green (self-archiving) and gold (publishing in OA Journal) (Carr et al., 2003).

2. The existence of Open Access mandates declared by funders and by research institutions, does not, per se, constitute any warranty of their effectiveness (Besemer, 2006) nor that:

a) they are adhered to

b) they are effective in assisting knowledge seekers locate and access knowledge resources.

Several studies have been carried out over the past two decades to help understand and quantify the 'Open Access' phenomenon at global scale, one of the most recent (Laakso et al., 2011) contains an overview of key studies in the field, summarised in Fig. 19, as well as providing a statistical analysis of Open Access chronological development over recent decades which identifies roughly three distinct phases:

- Pioneering (1993-1999)
- Innovation (2000-2004)
- Consolidation (2005-2009).

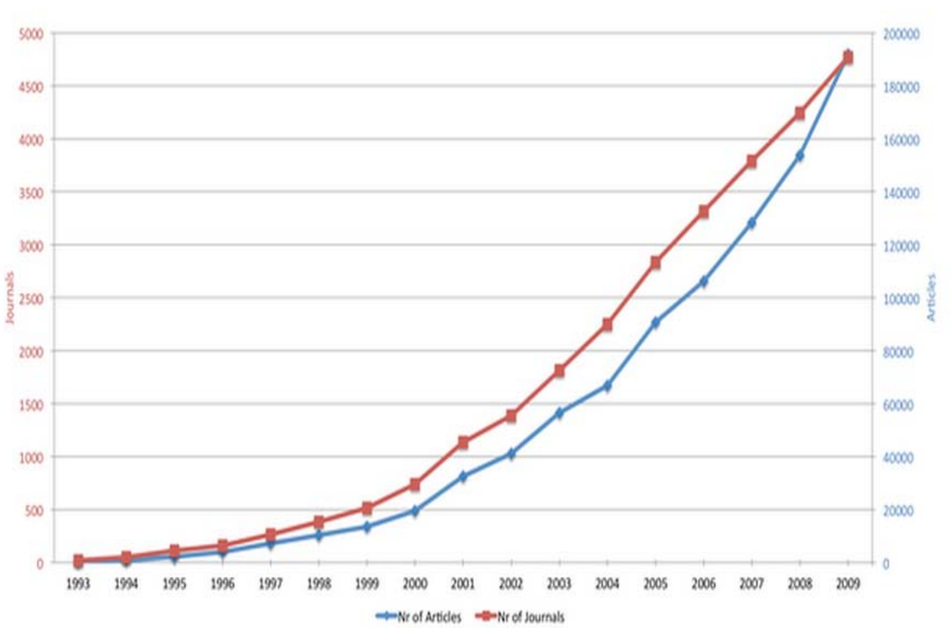


Fig. 19: The Development of Open Access 1993–2009 (Laakso et al., 2011)

An earlier related study (Björk et al., 2010) searched, using public search engines over the web, a sample of 1837 papers published in 2008, using the paper title as the search term, shows that 8.5% of published articles were freely available at the publishers' sites, and for 11.9% free versions could be retrieved using search engines, making the overall OA percentage 20.4%. The two studies above in combination show a steady increase of Open Access publication in the last two decades, however they also show the lack of qualitative, evidence-based research studies.

This doctoral research, taking into account the relevant literature cited in Table 7, positions itself to fill clear knowledge gaps in this important, albeit novel, field of investigation.

Source	Main Findings	Commentary
"Ware M, Mabe M, (2009) - An overview of scientific and scholarly journals publishing. International Association of Scientific, Technical and Medical Publishers."	"Number of scholarly journals has increased by 3.5% annually over the last 3 centuries, while growth in the total number of articles during the same time period has increased at the slightly slower pace of about 3% annually."	"Important trend when studying developments in growth of OA publications."
"Crawford W (2002) - Free electronic refereed journals: getting past the arc of enthusiasm. Learned Publishing"	"Study, conducted in 2001, to chart the 1995 OA landscape found evidence of 86 journals publishing in 1995 which fulfilled the criteria of free, refereed, and scholarly; it also investigated the status and activity 6 years later (in 2001). Only 49 journals, or 57%, were still actively publishing."	"Given the speed of change, measure or reconstruct, the open availability of journal articles prior to 1988 is a challenging task. Study includes review of OA journal developments between 1995 and 2001."
"Wells A (1999) -Exploring the development of the independent, electronic, scholarly journal. M.Sc.	"List of scholarly OA journals by combining data from several e-journal lists, verifying found journals by visiting their websites resulting in a list of 387 journals, publishing	"Compiled with some limiting criteria, the list can still be considered an important snapshot of OA publication in early years."

Thesis, The University of Sheffield.	an average of 18 articles per year."	
"Gustafsson T (2002) -Open Access - En empirisk undersökning om fritt tillgängliga vetenskapliga journaler på Internet, M.Sc. 92 p".	"Revisited Crawford list of journals to check their continued activity and to expand their new entries resulting in a total of 317 journals. Around half of the journals Wells originally documented had become inactive ; only 193 still publishing."	"This result is in line with the mortality rate noted by Crawford."
Hedlund T, Gustafsson T, Björk B-C (2004) The Open Access scientific journal an empirical study.	"Sent out a web survey to the editors of each of the journals in Gustafsson's updated list of 317 OA journals for which contact details were available, 300 in total, and received 60 responses, a response rate of 20%.	"They found out that during the year 2002, these journals published on average 20 articles each."
"Morris S (2006) Personal View: When is a journal not a journal - a closer look at the DOAJ. Learned Publishing"	"Study analysed the results of a labour-intensive data collection process where volunteers manually went through journal websites collecting publication metrics from 1213 of the total of 1443 OA journals listed in Directory of Open Access Journals (DOAJ). One key result, on average, 42 articles were published annually per journal."	"Study did not separate between journals which had been born OA and those that had later converted to OA, and focused on the article volume without regard for retrospective archival or conversions from subscription-based to open. So while the data is comprehensive, the results are not suited to represent the availability of OA article volume retrospectively for a given year."
"Björk B-C, Roos A, Lauri M (2009) Scientific journal publishing: yearly volume and Open Access availability."	"Sampled 100 of 1485 OA journals from Ulrichsweb Periodicals Database; found to publish on average 34.6 articles."	"Study is not directly comparable to studies based on DOAJ sampling as it excluded titles from large publishers which were known to charge author fees."
"Dallmeier-Tiessen S, Goerner B, et al.. (2010) Open Access Publishing - Models and Attributes. Max Planck Digital Library."	"Analysed data for all active English language journals listed in the DOAJ, (2838 in 2009) average journal published 43 articles in its most recent active year, which was either 2007 or 2008."	"Due to its focus on the current status of OA and an extensive publisher-level analysis, the study did not attempt to separate between converted subscription journals and born OA journals."
Edgar BD, Willinsky J (2010) A survey of scholarly journals using Open Journal Systems. Scholarly and Research	"Surveyed journals which use the popular Open Journal Systems (OJS) publishing platform. Study based on 998 survey responses (2748 questionnaires, a 36%	"Only a small number, approx. 7%, had uploaded back-issues to their archives. The self-reported annual average number of published articles among the

Communication"	response rate). Majority were founded as OA journals directly on the platform, but there were also many journals which had migrated to the OJS platform either from print-only or from other means of electronic publishing."	responding journals for 2008 was 31."

Table 7: Open Access Studies (from Laakso et al., 2011)

3.4 THE SOCIO-TECHNICAL CONTEXT

This research studies the relation between two distinct disciplinary dimensions of knowledge sharing in the context of 'networked organisations': on the one hand the policies that regulate knowledge sharing 'in theory', because it is largely through policies that processes and behaviours are shaped in an organisational context (Beer, 1998), and on the other hand the adoption of knowledge sharing artefacts and techniques, as these are the mechanisms that enable knowledge sharing in the context of 'networked organisations', namely, on the web or other electronic means of data exchange over an open public communication network.

One of the goals of this research is to measure the distance between these two dimensions, that can also be referred to as a **pragmatic gap**, and to recommend measures to reduce this gap. The multiplicity and complexity of the research dimensions can be framed in the context of 'socio-technical systems' (STS) research:

> *"Socio-technical systems involve a complex interaction between humans, machines and the environmental aspects of the work system—nowadays, this interaction is true of most systems in the workplace."*
>
> (Emery & Trist, 1960)

The socio-technical dimensions of networked capability research have already been explored in related work:

> *"Our ultimate purpose is not to send data through technical networks but to share understanding through social ones."* (Sillitto, 2006)

Social	Create shared mental models to change and align values, behaviours and actions across the enterprise
Process	End to end business processes to create the desired operational, strategic and political effects
Information	The right information available to the right people at the right time - understood and trusted
Technical	Compatible, available, evolvable and reconfigurable resources to support the enterprise

Fig. 20: Adapted by Sillitto (2006) from "Engineering Networked Enabled Capabilities"

Figure 21 identifies policy/organisation as social aspects of a network-enabled capabilities architectural framework.

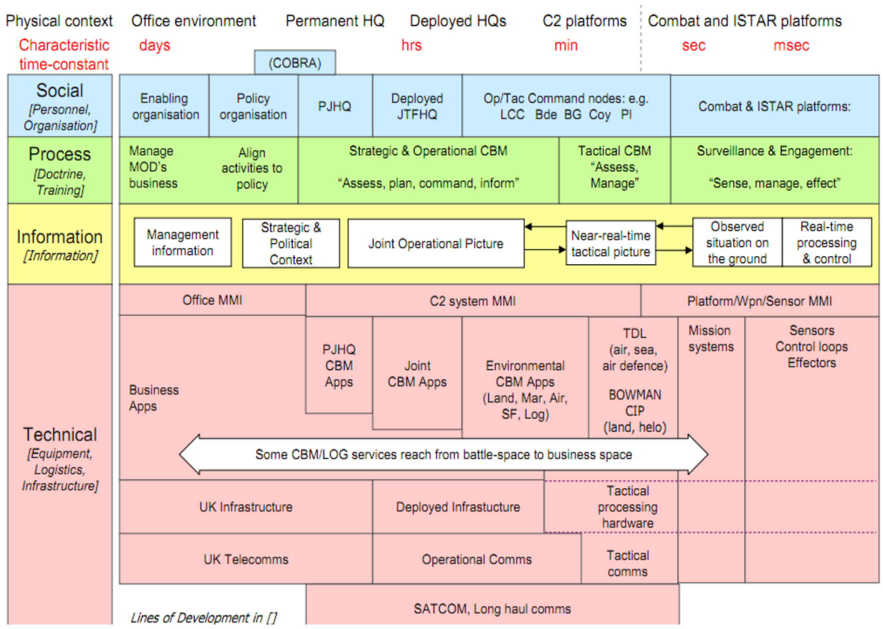


Fig. 21: Policy and Organisational Issues in Socio-Technical Systems Research Agenda (Sillitto, 2006)

Two of the basic principles of socio-technical systems development that concern this research are (Cherns 1987; Clegg 2000):

- Boundary control: boundaries should not be drawn so as to impede sharing of information, knowledge or learning.
- Information flow: information should be provided to those who require it when they require it.

At organisational level, 'information policies' governs knowledge and information flows:

> *"Socio Technical Systems include technical systems but also operational processes and people who use and interact with the technical system. Socio-technical systems are governed by organisational policies and rules."* (Sommerville, 2004)

Thus this research, which studies in part what policies govern the flow of scholarly knowledge and how they are adopted in practice, addresses directly open questions in socio-technical systems research. Among the techniques used to support the integration of disparate angles of the field work, a morphological box (aka Zwicky box) is being developed, shown in Table 8.

> *"Morphological analysis is a group of methods that breaks down a system, product or process into its essential sub-concepts, each concept representing a dimension in a multi-dimensional matrix. Thus, every product is considered as a bundle of attributes. New ideas are found by searching the matrix for new combination of attributes that do not yet exist."* (Ritchey, 2005)

General Morphological Analysis (GMA) consists of a set of techniques for structuring and investigating 'all' the possible relationships contained in multi-dimensional, non-quantifiable, problem complexes. It was first developed in the late sixties by Fritz Zwicky, a scientist based at Caltech, and is currently applied to the study of policy analysis.

Morphology - from the Greek 'morphe', the study of shape or form - is concerned with the structure and arrangement of parts of an object, and how these fit together to create a whole, either by studying the physical, social or conceptual (Ritchey, 1996).

The morphological field developed to explore the research field for this doctoral study includes Theory, Methodology and Artefacts with Cognitive, Organisational and Technical aspects:

Cognitive

Language, level of skill required to use/adopt the knowledge

Organisational

Policies and management practices that promote knowledge sharing and re-use

Technical

Choice of codification, knowledge representation techniques and standards adopted

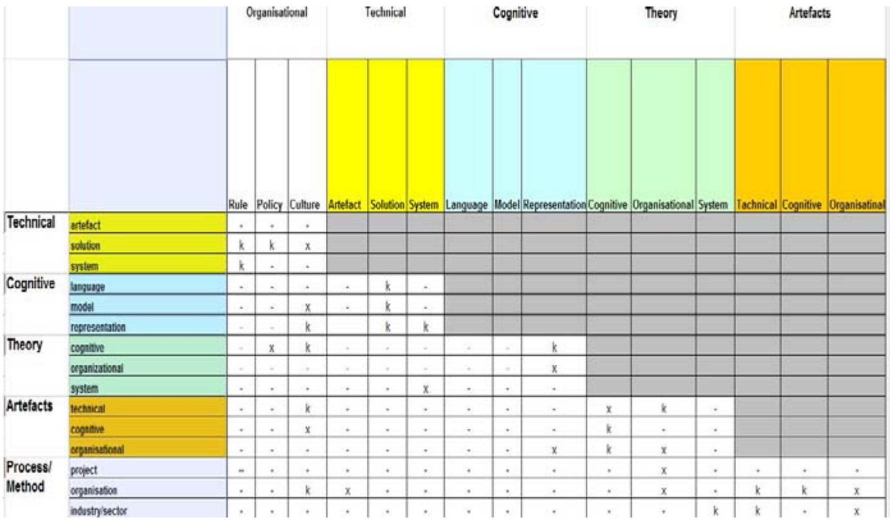
Selected elements from the morphological box are then used to navigate the research space. The complex mix of dependencies that contributes to the challenge is reduced to a combination of governance, process and structure, and mapped to the corresponding section of the morphological box:

	Cognitive	Organisational	Technical
Theory		GOVERNANCE	
Methodology			PROCESS
Artefact	STRUCTURE		

Table 8: Morphological Box

In related work a more complete morphological analysis of the research space is being carried out. The procedure of iterating analysis and synthesis to narrow the scope of a morphological field is called "cross-consistency assessment" (CCA) (Ritchey, 1996). A CCA matrix mapping the emerging research space is being developed alongside this doctoral dissertation, however it is open for ongoing refinement and future work.

Purely as a pointer, Fig. 22 is a CCA matrix currently being explored:[16]



		Organisational			Technical			Cognitive			Theory			Artefacts		
		Rule	Policy	Culture	Artefact	Solution	System	Language	Model	Representation	Cognitive	Organisational	System	Tachnical	Cognitive	Organisatinal
Technical	artefact	-	-	-												
	solution	k	k	x												
	system	k	-	-												
Cognitive	language	-	-	-	-	k	-									
	model	-	-	x	-	k	-									
	representation	-	-	k		k	k									
Theory	cognitive	-	x	k	-	-	-	-	-	k						
	organizational	-	-	-	-	-	-	-	-	x						
	system	-	-	-	-	-	x	-	-	-						
Artefacts	technical	-	-	k	-	-	-	-	-	-	x	k	-			
	cognitive	-	-	x	-	-	-	-	-	-	k	-	-			
	organisational	-	-	-	-	-	-	-	-	x	k	x	-			
Process/ Method	project	-	-	-	-	-	-	-	-	-	-	x	-	-	-	-
	organisation	-	-	k	x	-	-	-	-	-	-	x	-	k	k	x
	industry/sector	-	-	-	-	-	-	-	-	-	-	-	k	k	-	x

Fig. 22: CCA (experimental) for Emerging Field of Research

[16] Courtesy of T Ritchey, Swemorph, Sweden

3.5 CONCLUSION

One of the motivations behind this literature review is that as reported cursorily in Chapter 2, the outcome of some observations indicated that 'knowledge' is still a poorly understood construct in engineering research, and this is reflected in the gaps in state of the art. This chapter seizes the opportunity to capture insights from literature as ample evidence that 'knowledge', in particular 'explicit knowledge', requires adequate techniques, such as representation and codification, to enable sharing. In addition, two key knowledge gaps are identified:

- the importance and relevance of policies from 'future research directions' is missing in one of the recent overviews literature review in this academic field. In this chapter this gap is addressed by identifying the relevant arguments and sources provided, as well as with a recommendation in the final section of the thesis.
- existing research in Open Access consists mainly of large scale quantitative and statistical studies, which the rest of this study addresses by providing a complementary 'evidence-based' research angle.

Other literature reviewed demonstrates the relevance of knowledge and knowledge sharing to socio-technical systems research, and a morphological analysis approach is an example of a technique that helps the many dimensions of the problem space to be studied in correlation.

The rest of this dissertation discusses how this research makes a novel contribution toward both knowledge gaps identified.

"There comes a time when the mind takes a higher plane of knowledge but can never prove how it got there."

Albert Einstein

CHAPTER 4

RESEARCH INSTRUMENTS

OVERVIEW OF THE SECTION

This section provides an overview of the collection requirements, with references to existing auditing frameworks and environments, making the case for using URIs (web addresses) as evidence to answer the research questions, and provides the justification for the novel set of instruments developed for this research: essentially, no existing tool suits the requirement. Two methodologies (KAF and OAM) and corresponding implementation prototypes used as collection instruments are presented in this section. In the context of the study undertaken, KAM and OAM are the 'research protocols' used to carry out the systematic review of audited projects.

4.1 RECAP OF RESEARCH QUESTIONS AND HYPOTHESES

The main proposition is:

> There is a gap between theory (policies) and practice (their adoption) in OA
>
> * an equivalent formulation of the hypothesis is 'Open Access policies are not adhered to in practice'

To test the proposition above, this research asks the following question (Q1) and formulates a corresponding hypothesis (H1).

Q1. How can the gap between T and P be identified?

H1. By gathering evidence of the difference between T and P

Evidence of T is gathered initially by surveying the knowledge sharing policies *knowledge sharing 'in theory', and evidence of P is gathered by auditing publicly-funded systems engineering research projects 'in practice'. A 'collection' tool is developed to gather:

a) evidence of T : what are the Open Access policies? what do the policies prescribe?

b) evidence of P: to what extent are the policies identified in step a) adopted in practice?

Evidence of T, The Policies: The question 'what policies exist in support of knowledge sharing of publicly-funded research?' is answered initially by literature reviews (non-empirical research).The policy documents published on the websites of the respective funding bodies (Research Funding Councils), to 'open scholarship' are cross-referenced with evidence gathered among different sources, including exchanges with key funding bodies personnel and FOI requests, to see if the information located across multiple sources was consistent, as explained in Chapter 5. Although P for this study has been acquired via literature review, a template and online form to collect policy data have been developed for completeness, as explained further below.

Evidence of P – To gather evidence of how the policies above are implemented 'in the practice' an auditing methodology and tool are required.

4.2 COLLECTION REQUIREMENTS

The requirements for the auditing methodology and tools to support this research are:

- to enable the audits to be carried out 'remotely', to minimise the need for on-site visits and travel, given the limited resources available for this doctoral research
- capable of auditing 'knowledge resources' which are shared publicly on the web

- to provide support for domain specific knowledge (this study is auditing knowledge in a specific domain)

- to provide an online environment/freely available tool that can be used to support the auditing process

- to enable users without specialised knowledge of Open Access or knowledge policies to carry out a systematic audit of Open Access knowledge resources

- to be capable of locating documentary evidence in support of knowledge as fact using URIs

4.2.1 Linked Data for Fact Checking

The ability to verify facts via gathering evidence is essential to reason, make inferences, draw conclusions and, essentially, to make informed decisions. On the web, which is the largest open, large-scale distributed knowledge base, fact checking is particularly important to the accuracy of reasoning, which can be defined as the act or process of using one's reason to derive one statement or assertion (the conclusion) from a prior group of statements or assertions (the premises) by means of a given method. The validity of any given argument, relies on the premise of the argument being 'true'; any validity claim or assertion needs to be verified or verifiable, with some exceptions that may be satisfied with theoretical assumptions. Yet in practice, few mechanisms exist which are used in the real world to distinguish between fact, belief and opinion, and unverified assumptions are routinely passed as 'knowledge' (Fig. 23).

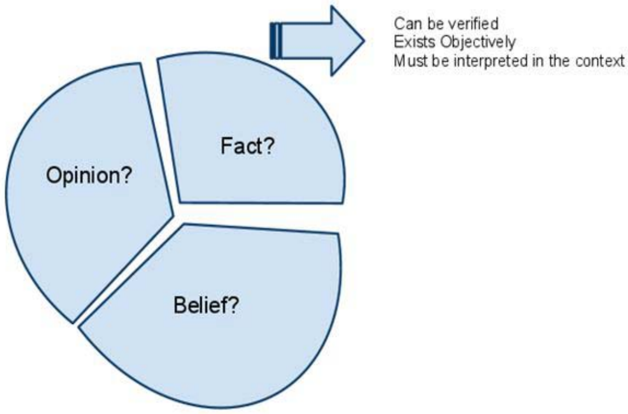


Fig. 23: Belief, Opinion or Fact? (Di Maio, 2011)

Fact checking is adopted routinely in investigations (research) in providing evidence (legal/making the case) and in decision making (to reduce over-reliance on assumptions). Tracking the source of evidence in the technical and systems sciences is called 'provenance' (from the Latin 'pro venire', which means 'to come from'.[17]), which is typically intended 'source' and 'origin of', but it can also be used as documentary evidence and justification, of some assertion or decision. Evidence can take the form of data, information, or a digital object, such as for example, a record. Provenance can be described in various terms; for example in the context of database systems, it is defined as the description of the *origins* of data and the process by which it arrived at the database (Buneman et al., 2001) but it can also be seen as *lineage*, as information describing materials and transformations applied to derive the data (Lanter, 1991). Provenance can also be distinguished between 'data' and 'process', the latter being applicable in our definition above of provenance as 'justification' (Simmhan et al., 2005). There are different kinds of provenance, and possibly definitions vary for each, some examples are provided in Table 9.

[17] Online Etymology Dictionary http://www.etymonline.com/index.php?term=provenance

Types of Provenance	Explanation	Example
verify/validate	provenance used to verify the validity of data/information	provide the source of information and double check/cross reference an assertion
track/monitor	provenance to evaluate the history of a fact/assertion/piece of evidence	establish the sequence of events or chronology of evolution of a digital object
demonstrate/prove	provenance data as evidence	support of an hypothesis or claim

Table 9: Examples of Types of Provenance

'Provenance' on the web can be used as evidence that a statement is true, false or unverified, hence serve as a fact checking mechanism. The 'linked data' model (Bizer et al., 2009) is a contemporary metaphor that illustrates a rather mature concept of 'hyper data' or 'hyperlinked data'[18], which has been at the heart of the web architecture since its inception. In related work[19] it is proposed that to be used reliably every assertion/claim should be verifiable via a hyperlink to supporting evidence for which there should be a verification mechanism to help answer the following questions, for example:

- is the evidence valid?
- what is the degree of confidence in the validity?
- what evidence supports the confidence in the validity?

Figure 24 (Levitt et al., 2010) provides some examples of 'sources of evidence' (evidence of knowledge as being a fact, or any evidence that is relied upon to make decisions).

[18] Idehen, K. BLOGSPOT https://plus.google.com/112399767740508618350/posts/hCmZKxA4CJY
[19] http://wiki.esi.ac.uk/Workshop:_Understanding_Provenance_and_Linked_Open_Data

SECONDARY SOURCES	
Secondary data	Data already available, previously collected by the organisation or by others.
Literature review	Looking at past research or commentary to assess what is already known.
Comparative studies	Looking at how similar activities are undertaken in other settings.
Scenarios	Formal speculation about future trends or events.
Document review	Reports or correspondence, either internal to the organisation or produced outside.
Self assessment	The organisation's own monitoring and evaluation of its performance.
PRIMARY SOURCES	
Consultation	Seeking views of the staff, service users or other stakeholders of the organisation under review.
Survey	Systematically gathering data from staff, service users or other stakeholders.
Focus group/ discussion	Usually with staff or service users.
Hearing	Where invited witnesses make oral presentations and are questioned, usually in public.
Interview	With main informants, usually undertaken one to one, either face to face or by phone.
Observation	Usually on visits to sites where services are provided; likely to be non participant observation.

Fig. 24: Examples of Sources of Evidence

On the web, each source, or set of sources should be located and accessed using URIs.

4.2.2 Existing Tools and Methods

Table 10 summarises a shortlist of the existing instruments evaluated against the requirements for this research. Existing and publicly available knowledge auditing methodologies and tools do not satisfy the auditing requirements for this study, therefore *ad hoc* auditing instruments are devised, presented in the paragraphs that follow.

Audit methodology	Provider	Remote evaluation	Auditing knowledge shared on the web	Domain specific	Free toolkit?
Knowledge Studio	Knowledge Studio	no	no	yes	no
DAF	Hatii Glasgow	no	no	no	yes
Knowledge Audit Cycle	Schuma Marr 2001	no	some	possible	no
Systematic Knowledge Audit with Application	Hong Kong Polytechnic	possible	no	possible	no
http://www.tlainc.com/articl97.htm	possible	possible	no	possible	no

Table 10: Knowledge Auditing Methodologies Surveyed

The first collection tool developed presented in this section is KAF, a 'knowledge auditing framework' modelled on DAF, an existing 'data auditing framework', as it was deemed to be the most recent published auditing instrument used in the sector. After piloting KAF 'in the field', it emerged that it targeted a rather detailed level of knowledge resources, while many (most, as it was confirmed from later findings) of publicly-funded projects in systems engineering research did not seem to have even a website as such.

A more generic level of auditing was required. The second collection tool that evolved from KAF - is called Open Access Monitor (OAM). It can be considered an adaptation of KAF, adopting a similar process and principles, but in a slightly more agile and efficient version designed to capture a lesser degree of detail, but enough to answer the research question and test the hypotheses for this study. The audits presented in the following chapter of this thesis have been carried out using OAM. Both instruments are described below, together with justification for the respective modelling choices, and links are provided to the respective online working prototypes.

4.3 KNOWLEDGE AUDIT FRAMEWORK (KAF)

Publicly-funded research generates a vast amount of 'new knowledge', at least that's how it justifies the disbursements to taxpayers, however no mechanisms exist to assess 'exactly' what knowledge, and what effort is necessary to be able to access and re-use this new scholarly knowledge. This research draws upon an established technique used in Knowledge Management, called 'knowledge auditing', and the instrument that emerges, named KAF (Knowledge Auditing Framework) has integrated essential aspects of an existing framework currently in use, DAF (Data Auditing Framework).

KAF can be applied in principle to audit any project, organisation or institution, although the current working version targets the system domain, given the current scope of work. The paragraphs that follow provide background information to Knowledge Audits in general, and illustrate in some detail the KAF architecture, as well as its limitations.

4.3.1 About Knowledge Audits

Knowledge Audits are advocated as essential activities for most knowledge management activities (Sharma & Chowdury, 2007) and consist typically of inventorying and mapping:

> *'A Knowledge Audit (K-Audit) is hence a systematic examination and evaluation of organizational knowledge health, which looks at whether knowledge is exploited when needed. More specifically, it is an analysis of the organization's knowledge needs, existing knowledge assets or resources, knowledge flows, future knowledge needs, knowledge gaps, and finally, the behavior of people in sharing and creating knowledge. In one way, a knowledge audit can reveal an organization's knowledge strengths, weaknesses, opportunities, threats and risks.'*
>
> (Cheung et al. 2007; Hylton 2002; Liebowitz et al. 2000;
> Schikkard & Toit 2004 in Sharma & Chowdhuri, 2007)

Knowledge audits can be viewed as planning documents, providing a structural overview of a designated section of an organisation's knowledge, as well as details of the qualitative and quantitative characteristics of the individual chunks of knowledge within that designated section (Liebowitz, 1999), where the document also identifies the knowledge repositories in which those chunks reside.

In knowledge management literature 'knowledge assets' are defined as 'stocks from which services are expected to flow' (Boisot, 1998), paraphrasing the definition of asset in financial accounting. By contrast, this research studies knowledge as 'public good', in particular knowledge generated using public funding. Therefore in this research knowledge assets are defined as 'the artefacts resulting from the process of knowledge 'codification', as explained in Chapter 1, and the preferred expression to point to knowledge artefacts is *knowledge resource*. The preferred terminology in this research is 'knowledge resources', although knowledge assets, artefacts and resources are sometimes used interchangeably, unless otherwise stated. Knowledge Assets can be grouped into four categories (Nonaka & Takeuchi, 1995), two tacit and two explicit, as illustrated in Fig.25:

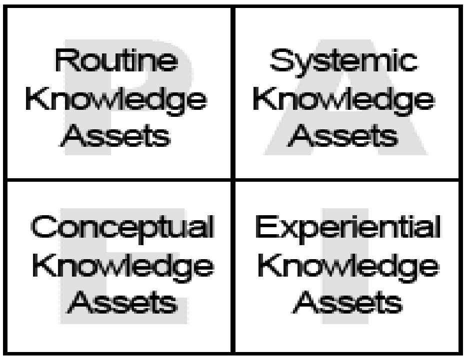


Fig. 25: Four Types of Knowledge Assets (Nonaka & Takeuchi, 1995)

The two 'explicit' categories of assets directly relevant to KAF (where they are called 'knowledge resources') and which are reflected in the design of the knowledge inventory templates are:

Systemic Resources, Combination, Systematising (Virtual collective): Explicit, codified, systematic, descriptive, complete, comparative, evaluative. Examples include documents, specifications and manuals.

Systemic resources consist of systematised and packaged explicit knowledge, such as explicitly stated technologies, product specifications, manuals and documented information about customers and suppliers, including patents and trademarks, visible and tangible, they be transferred easily (Chou & He, 2004).

Conceptual Resources, Externalisation, Originating (Face-to-face individual): Symbols, concepts, brands, styles, metaphors, analogies, emergent, developmental. Examples include product concepts and design. Conceptual resources consist of explicit knowledge articulated through images, symbols and languages. They are the assets based on the perceptions held by customers and employees of the organisation. For example, concepts or designs which are perceived by the members of the organisation. Conceptual knowledge assets usually have tangible forms and are easier to articulate. According to the distinction referenced above, conceptual resources contain information required to product development and specific design features, while conceptual assets/resources can be further distinguished (Chourabi et al., 2010) into:

-**Domain facet:** contains basic concepts and relations for describing the content of engineering assets on a high semantic level. (Domain ontology)
-**Product facet:** contains concepts and relations representing artefact types as well as their information model. In the systems engineering domain, a system is described with several views such as, contextual, dynamic, static, functional or organic.
-**Process facet**: contains concepts and relations that formally describe engineering activities, tasks, actors, and design rationales concepts (intentions, alternatives, argumentations and justification for engineering decisions).

This research integrates the two descriptions, and postulates that to be shared, for every conceptual knowledge resource there should be a corresponding systemic one, Fig. 26:

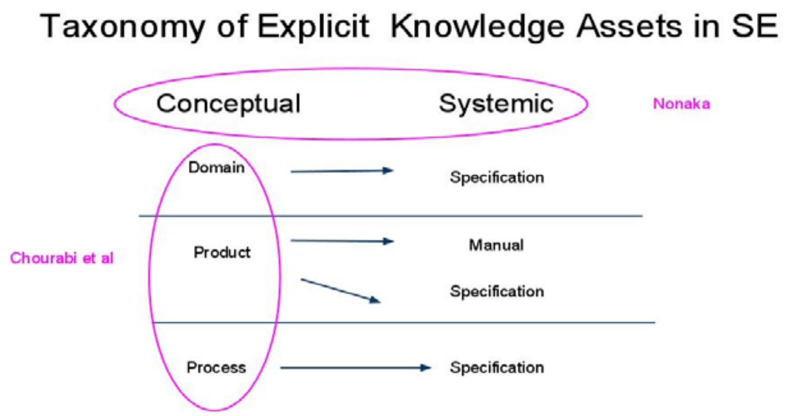


Fig. 26: Integrated Explicit Knowledge Assets (Di Maio, 2011)

4.3.2 KAF Overview

The Knowledge Audit Framework (KAF) is a technical instrument that supports the systematic evaluation of the adoption of knowledge sharing practices, by providing a methodology and a set of templates for the auditing of 'explicit' knowledge resources. The current version of KAF includes a template and a process to perform Knowledge Inventorying with emphasis on knowledge sharing artefacts. Publicly-funded research generates 'knowledge', sometimes referred to as 'scholarly knowledge'. In other sections, this thesis discusses in some detail how several initiatives exist to promote the 'Open Access' to scholarly knowledge.

However no mechanism exists to assess exactly 'what project knowledge' generated with public funding is openly accessible via the web. KAF is developed to provide guidance for the planning and execution of systematic knowledge audits that can take place remotely and via web searches, without requiring the cooperation and consent of the owners of the resources, or their permission to physically enter their premises, although communication and participation are explicitly encouraged and designed in the auditing process starting with phase one. The primary goal of KAF is to enable an objective evaluation of what knowledge resources are shared, where they are located and who is (or is not) responsible for their creation and maintenance, with particular emphasis on auditing knowledge sharing and re-use.

KAF currently comes in two versions: KAF-G, a generic, domain independent version and KAF-SE, targeted at the systems engineering domain. The two versions share the same auditing process, (Table 10) but each adopts different templates: KAF-G targets the 'generic knowledge', which is non domain specific, while KAF-SE uses templates which have been structured to model 'system knowledge', which is the domain of systems engineering.

4.3.3 KAF vs. DAF

KAF is designed to assist the evaluation of knowledge resources and sharing practices in systems engineering, while DAF (Data Audit Framework) was published by HWAII at the University of Glasgow, and is proposed as a methodology and toolkit to audit data assets. DAF was not meeting the requirements for our study, in that:

a) it audits 'data' and not 'knowledge', the respective definition and differences presented earlier in this dissertation

b) it requires the cooperation of the organisations whose datasets are audited, employee time, visits to the premises, and the analysis of archival documentation.

This research however, is explicitly seeking to audit 'knowledge resources', and not solely 'data repositories'. KAF however, inherits and shares some of DAF core components, namely:

- general methodological approach
- structure (the KAF process is articulated in four stages)
- some of the original arguments are retained, and where suitable,
- portions of the original methodology are adopted

Specifically, KAF differs from DAF as follows:

Type of resources: KAF targets knowledge resources, which is more high-level, contextualised information, as opposed to data resources.

Scope: DAF is aimed primarily at Higher Education institutions, while the current version of KAF is intended to support the auditing of 'any' knowledge.

Process: like DAF, KAF is articulated around four stages, however each step in the corresponding KAF stage is different from DAF, for example, KAF does not require on-site visits and full access to project documentation, but relies on knowledge that is shared primarily on the web and by open electronic communication, and that can be accessed (or not accessed) via inspection of the project website, or remotely via electronic communication exchanges with the project team (mainly emails).

Audit mode: KAF is designed to support remote audits, however templates can also be used during on-site visits if desired.

Taxonomy: KAF rationale is based on a different conceptual/taxonomic set of relations than the original DAF, as briefly explained below.

Emphasis: KAF emphasises the audit of Knowledge Sharing artefacts and procedures in relation to their *reusability*, rather than simply to inventory their existence.

	DAF	KAF
Type of resources	Data	Knowledge
Scope	Higher Education	Any organization
Process	4 stages, in situ	4 stages, remote
Audit mode	In situ	Remote + in situ
Taxonomy	Data	Knowledge
Emphasis	Data in repository	Data shared on the web

Table 11: Main Differences between KAF and DAF

KAF contains a template and a process to carry out 'knowledge resources audits' that can be applied in principle to any project, organisation or institution, although the current working version targets the systems engineering domain.

The original DAF methodology upon which KAF is based, relies upon a concept map that uses two taxonomic groups for data assets ('by origin' and 'by nature'), as indicated in the original DAF document[20]. KAF by contrast is shaped taking into account the organisational knowledge taxonomy (Vasconcelos et al., 2000) which distinguishes between different types of knowledge, such as tacit and explicit, and splits the latter into declarative and procedural, associating it with tangible assets/resources which are the object of study in this doctoral research, as shown in Fig. 27. Purely for completeness, the original organisational taxonomy diagram is extended to include 'causal' knowledge[21], where evidence of 'provenance', for example knowledge which answers the question 'why', is also considered where possible.

[20] Page 7, DAF Methodology http://www.data-audit.eu/DAF_Methodology.pdf

[21] Private Correspondence with Vasconcelos

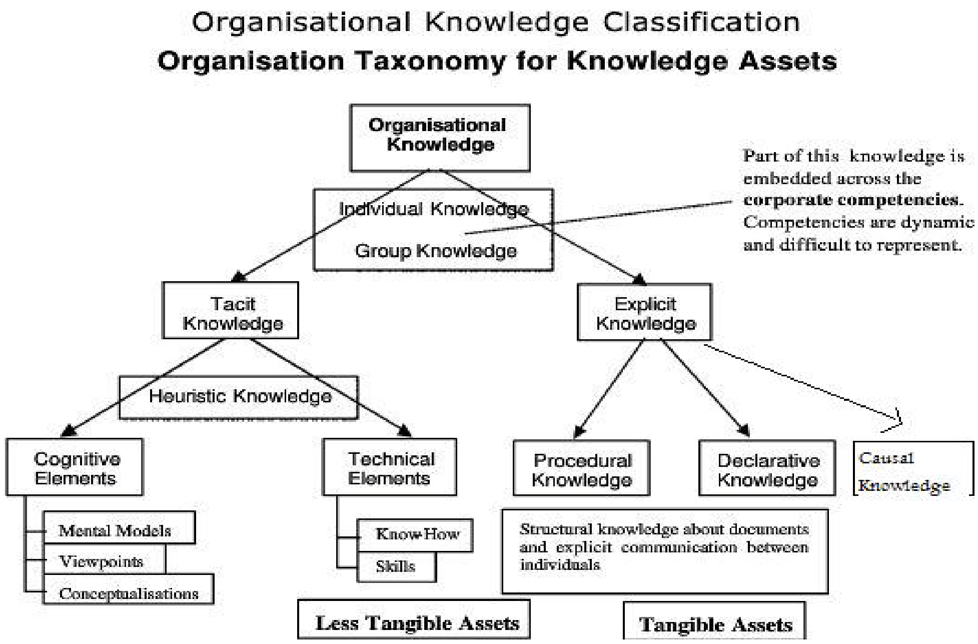


Fig. 27: Modified Organisational Taxonomy (Vasconcelos et al., 2000) modified by Di Maio

4.3.4 KAF Process and Templates

To effectively manage and facilitate the re-use of explicit knowledge assets, an organisation must adopt knowledge sharing artefacts and behaviours (Weber). Conducting an audit enables a systematic mapping of knowledge re-use artefacts, highlighting gaps and weaknesses in the overall Knowledge Management strategy for the research projects under investigation, and for the organisations that initiate and deliver them. The framework is designed to support audits both at project and organisational level, and without dedicated or specialist staff, since carrying an audit follows a simple, diagrammatic process and does not require any specific training, nor any particular resources. The current version of the audit templates is designed to collect information required to evaluate the level of adoption of knowledge re-use artefacts and practices, that is, it audits specifically resources via their availability using URIs, discussed in Chapter 3.

The audit framework addresses five core questions:

1. What knowledge resources are published on the web using a URI?

2. Where are these resources located? (What is the URI for each resource?)

3. What knowledge sharing mechanisms (notation, formalisms) are adopted (or not)?

4. Who is responsible in the organisation for making knowledge shareable?

5. Is the organisation adopting/following a knowledge sharing and re-use policy?

KAF consists of:

a) An audit process

b) An audit template

c) An online repository to collect and analyse the data of the audited project, a prototype of which is provided as proof of principle

d) Examples of forms for communicating with project teams and their organisations

4.3.4.1 The process

KAF is designed to support audits carried out remotely, and via email and other means of remote communication including telephone. The length of each audit depends on a number of variables, but each audit, based on our experience, can take anything between 1-3 days (4 - 24 work hours) from start to finish, assuming there will be time gaps between each step of the audit. The formula for carrying out Information and Knowledge Audits activities can vary, however auditing revolves around a set of standard steps (Liebowitz et al., 2000).

An auditing process needs to be 'standardized', among other reasons, because its results need to be 'replicable' and 'verifiable'. There are four main stages to the KAF process illustrated in Fig. 28, described in the paragraphs that follow.

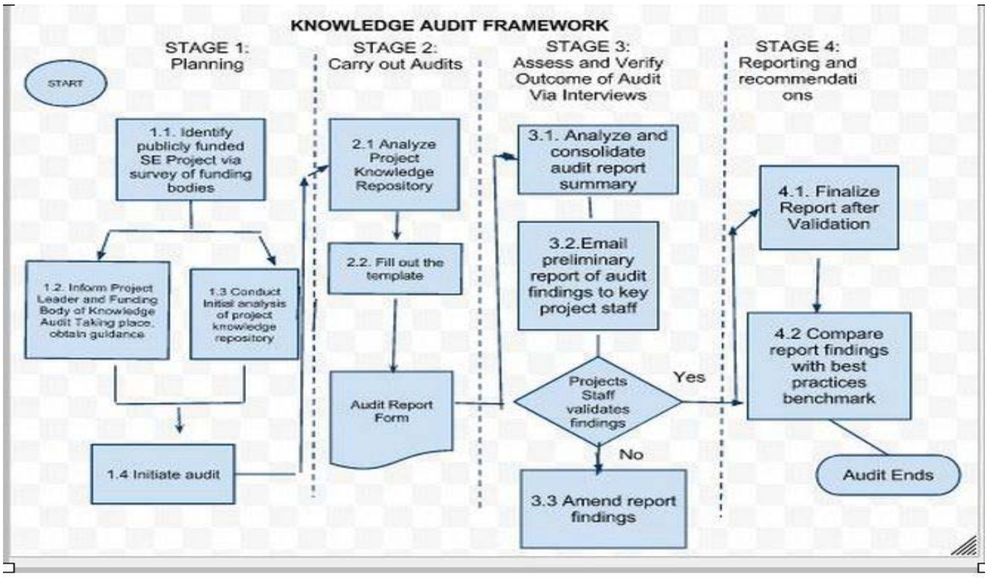


Fig. 28: The Four Stages of the KAF Process

STAGE 1: Planning (identify project and key team members)

Stage 1.1 Identify what to audit

The first step in a KAF audit consists of identifying the project, or organisation, to be audited. Knowledge Audits can be targeted to organisations, institutions, departments, public authorities. What to audit depends on why an audit is carried out. In this research, audits are carried out to try to establish to what extent systems engineering research projects share research output on the web as 'Open Access' resources.

Audits can be carried out for a variety of different reasons. When identifying 'what to audit', the person or office responsible for creating, publishing and maintaining knowledge resources should also be identified, so that it can be contacted in the next step.

Stage 1.2 Make contact with the people in charge

After identifying what to audit (an organisation or project for example) and the person in charge (a knowledge manager or Principal Investigator, for example), establish an initial contact informing them that a Knowledge Audit is about to be carried out, and to provide you with any information that they consider relevant or useful to carry out the audit. See an example in Appendix E.

Stage 1.3 Conduct initial analysis

The purpose of a KAF audit is to identify with minimal input (only the input that has been provided as a response to step 1.2) what knowledge resources can be located and accessed via web searches. A number of possible outcomes can be expected. An initial analysis of 'what to audit' will help locate basic information such as website address, who is the owner of the resources, any copyright or other access/re-use limitations etc.

Stage 1.4 Initiate the audit

Initiating the audit consists of assigning a unique identifier to the audit, and to create a document where the audit information can be recorded and annotated manually, before being entered in the online audit repository/database.

STAGE 2: Execution (carry out the audit)

Stage 2.1 Analyse (identify) knowledge resources for the project

Based on the goal of the audit (why is this audit being carried out?), and on the preliminary information gathered in step 1. 3, and after having made initial contact with the owner of the knowledge resource, one can start identifying which knowledge resources can be located, and what information is available for each resource. Make a note for each knowledge resource.

Stage 2.2 Fill out the slots in the audit form template

Fill out the audit template. This can initially be done manually, and then transferred to the electronic repository. Repeat for each knowledge resource.

STAGE 3: Verification

Stage 3.1 Consolidate results of the audit in a report

If more than one 'knowledge resource' is available, consolidate the inventory into one document that summarises how many resources have been identified, and what information has been located for each. This could be, for example: knowledge resources for project x are: one diagram, one vocabulary, one rule set.

Stage 3.2 Send a copy of the initial report to the project team for validation

After making a summary of the inventoried resources, send the preliminary report to the person responsible for the knowledge resource, and ask them to validate it, that is, to confirm that the information on the report is correct, else, to provide correct information.

This step is particularly important not only to ensure that the audit is correct, but also to establish a 'rapport' with the owners of the knowledge resource, and to engage them with the audit process.

Stage 3.3 (optional) Carry out an interview and amend report

(Repeat 3.3 until 3.2 is validated OK). If necessary, arrange for a phone call with the knowledge resource owner to gather information and amend the report until it is accurate.

STAGE 4: Report Findings

Stage 4.1 Finalise report

After inventory summary is approved by the project lead, finalise the report. Produce a final report incorporating any updates provided in Stage 3.

Stage 4.2 Compare report findings with good practices, issue recommendations

If the knowledge resources are not shared using good practices, (see recommendation section of this thesis), issue recommendations accordingly.

Stage 4.3 Get feedback from the team on the KAF process

A feedback form is provided for knowledge resources owners who have been involved in the audit process, to gather input to be incorporated in future iterations.

STAGE 5: Audit Ends

4.3.4.2 The templates

This section presents some of the background rationale and structure of KAF's generic template and the corresponding conceptual knowledge schema, for the inventorying of knowledge resources that can fit non domain specific knowledge auditing requirements, as well as a more targeted template designed to audit knowledge in the systems engineering domain, and includes a discussion about some of the design choices.

Additional templates could be designed to suit different requirements. The templates have been developed adopting principles of knowledge modelling techniques discussed and justified in literature in an earlier part of this dissertation, and can be composed and presented to the user adopting different styles and interfaces. The following three main template components of the KAF template are described below. These are, however suggested as 'in principle' template design structures, and can be adopted flexibly by different implementations, and ideally evolved and refined with usage. KAF users are encouraged to engage and contribute to the KAF resources via the project website.

1. About the audit – the first portion is designed to contain 'metadata':

AUDIT NAME	Acronym, easy to remember
UNIQUE IDENTIFIER	If the acronym is not unique, other unique ID
AUDIT BEGIN DATE	Date Phase 1 is initiated
AUDIT END DATE	Date audit is completed
AUDITORS NAME AND CONTACT	Name of person/persons carrying out the audit
AUDIT PURPOSE	Details of motivation
OTHER INFO	Catch all slot

2. About what is being audited – this portion of the template targets knowledge about the project/institution or other entity being audited. Note that if the template is split into different documents, the first field should report the unique identifier for this audit across all corresponding portions of the template. If the template is implemented as single document/form, then the repetition is not necessary:

(AUDIT NAME/UNIQUE AUDIT ID)	
CONTACT PERSON	Person appointed to liaise in relation to the audit, or other responsible person
WEBSITE	Web page
OTHER REPOSITORY	Other online address where information about what is being audited may be located and accessed
IS THERE A KNOWLEDGE SHARING POLICY?	Some projects declare this information explicitly, in other cases it may be necessary to ask the contant person
IS THERE A KNOWLEDGE SHARING PROCESS?	Some organisations may have standardized procedures in relation to knowledge sharing to reflect the organisational culture

3. Knowledge resources – the auditing template structure leverages standard knowledge elicitation questions, such as what, who, where, how, why (and why not), and in the case of temporal knowledge is sought, when. These can be mapped to capture various kinds of knowledge, such as declarative, procedural and causal knowledge, as introduced in an earlier section.

QUESTION TYPE	EXAMPLE	KNOWLEDGE TYPE
WHAT (factual, declarative)	What knowledge resources are available? What ...definition of the knowledge object	declarative
WHO (person or role associated with this knowledge)	Who knows.... Targeting the person or role associated with the target knowledge	declarative
WHEN (temporal dimension)	Any temporal information in relation to the knowledge, including date the knowledge resource is created/updated/edited/deleted	declarative
WHERE (location dimension)	Location of the knowledge, if web based, this should be a URI	
HOW (process)	Is this knowledge associated to a process/procedure? If so, specify it here. Typically, the resource is invoked by some workflow, rule or other policy	procedural
WHY (cause, provenance, justification)	Provenance, justification for this knowledge resource. Could be required by some process	causal

4.3.4.3 KAF- SE:

Based on the notion of 'domain' introduced in Chapter 3, and adopting techniques from Knowledge Engineering methodologies such as CommonKADS, which consist of capturing knowledge using a series of layers/facets via different models, for the example given, the knowledge model, a domain model and a template structure for the auditing of systems engineering knowledge is provided schematically below.

This extends and should be used in combination with the generic elements of the template. The knowledge model used for the auditing template (descriptive) is also used to devise the knowledge sharing model for sharing systems knowledge (prescriptive) presented in the recommendation section of this thesis. The 'domain' model for systems engineering knowledge can be derived from a number of sources. For the purpose of this research, the following two are adopted to model the template: a representation of the system lifecycle, and the system specification document.

4.3.4.4 The 'system lifecycle':

A system development lifecycle is typically envisaged as 'phases', undertaken sequentially or iteratively as needed (until the requirements are met) throughout the life of the system, a typical example is provided in Fig. 29.

KAF does not take into account, at this stage, different lifecycle models (waterfall vs. iterative for example) but only typical phases, say for example:

- Requirements Analysis
- Specification
- System Analysis and Design
- Implementation
- Testing
- Maintenance

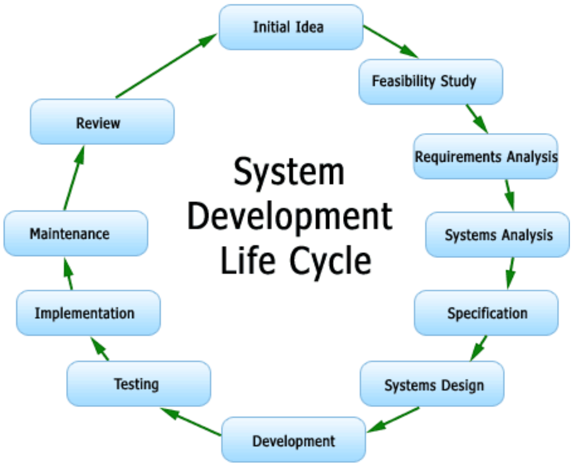


Fig. 29: A Typical System Lifecycle (Courtesy of Sharper Tutorials)

A system specification is a document which includes, at a minimum, a high level overview of the system and a description of the system goals. A summary example of a system specification used in industry is reproduced in the next two pages, of which the working version of KAF only adopts the top level categories.

EXAMPLE OF SYSTEM SPECIFICATION [22]

1.0 Introduction

This section provides an overview of the entire system or product. This document describes all subsystems including hardware, software, human activities, documents, and process.

1.1 Goals and objectives

Overall business s goals and project objectives are described.

1.2 System statement of scope

A description of the entire system is presented. Major inputs, processing functionality and outputs are described without regard to implementation detail.

1.3 System context

The system is placed in a business or product line context. Strategic issues relevant to context are discussed. The intent is for the reader to understand the "big picture."

1.4 Major constraints

Any business or product line constraints that will impact the manner in which the system is to be specified, designed, implemented or tested are noted here.

2.0 Functional and Data Description

This section describes overall system function and the information domain in which it operates.

2.1 System architecture

A context-level model of the system architecture is presented.

2.1.1 Architecture model (e.g., ACD)

A context-level model of the system architecture is presented.

2.1.2 Subsystem overview

Each subsystem noted in the architecture model is described briefly.

2.2 Data Description

Top-level data objects that will be managed/manipulated by the system or product are described in this section.

2.2.1 Major data objects

Data objects and their major attributes are described.

2.2.2 Relationships

Relationships among data objects are described using an ERD- like form. No attempt is made to provide detail at this stage.

2.2.3 System level data model

An ERD for the system is developed (this section may be omitted).

2.3 Interface Description

The system's interface(s) to the outside world are described.

2.3.1 Machine interfaces

Interfaces to other machines (computers or devices) are described.

2.3.2 External system interfaces

Interfaces to other systems, products or networks are described.

2.3.3 Human interface

An overview of any human interfaces to be designed for the system/product is presented.

3.0 Subsystem Description

A description of each subsystem is presented.

3.1 Description for Subsystem n

A detailed description of each subsystem is presented. Section 3.1 is repeated for each of n subsystems.

---

[22] Used with the permission of R.S. Pressman & Associates, Inc., http://www.rspa.com/docs/Systemspec.html

3.1.1 Subsystem scope

A statement of scope for subsystem n is presented.

3.1.2 Subsystem flow diagram

A diagram showing the flow of information through the subsystem and the transformation that it undergoes is presented.

3.1.3 Subsystem n components

A detailed description for each component of subsystem n is presented. Section 3.1.3 is repeated for each of k components.

3.1.3.1 Component k description (processing narrative)

3.1.3.2 Component k interface description

3.1.4 Performance Issues

Special performance required for the subsystem is specified.

3.1.5 Design Constraints

Any design constraints that will impact the subsystem are noted.

3.1.6 Allocation for Subsystem n

The allocation for implementation (e.g., will the subsystem be implemented in hardware, software, by a human, etc.) is described.

3.2 Diagrammatic model for Subsystem n

A diagrammatic model for each subsystem is presented. Section 3.2 is repeated for each of n subsystems.

4.0 System Modelling and Simulation Results

If system modelling and simulation and/or prototyping is conducted, these are specified here.

4.1 Description of system modelling approach (if used)

The system modelling approach (including tools and/or mathematical models) is described.

4.2 Simulation results

The results of any system simulation are presented with specific emphasis on data throughput, timing, performance, and/or system behaviour.

4.3 Special performance issues

Special performance issues are identified.

4.4 Prototyping requirements

If a system prototyping is to be built, its specification and implementation environment are described here.

5.0 Project Issues

An overview of the overall system/product project plan is presented.

5.1 Projected development costs

The results of system-level cost estimates are presented.

5.2 Project schedule

A top-level schedule for the development project is proposed.

6.0 Appendices

Presents information that supplements the System Specification.

6.1 Business Process Descriptions

If the specification is developed for a business system, a description of relevant business processes is presented here.

6.2 Product Strategies

If the specification is developed for a product, a description of relevant product strategy is presented here.

6.3 Supplementary information (as required)

Both lifecycle phases and system specification are useful archetypical 'knowledge models'. The KAF-SE template is constructed by integrating 'core elements' of the two models identified above, schematically captured in Table 12:

Requirements Analysis	System Scope, Goals Context Constraints Axioms/Rules
System Analysis And Design	Architecture Components/Subcomponents Functional Descriptions Data Inputs/Outputs Interfaces Performance Resources
Implementation	Code, Materials, Software and Hardware
Testing	Test Schemas, Outputs, Simulations Benchmarks
Maintenance	Documentation New Requirements

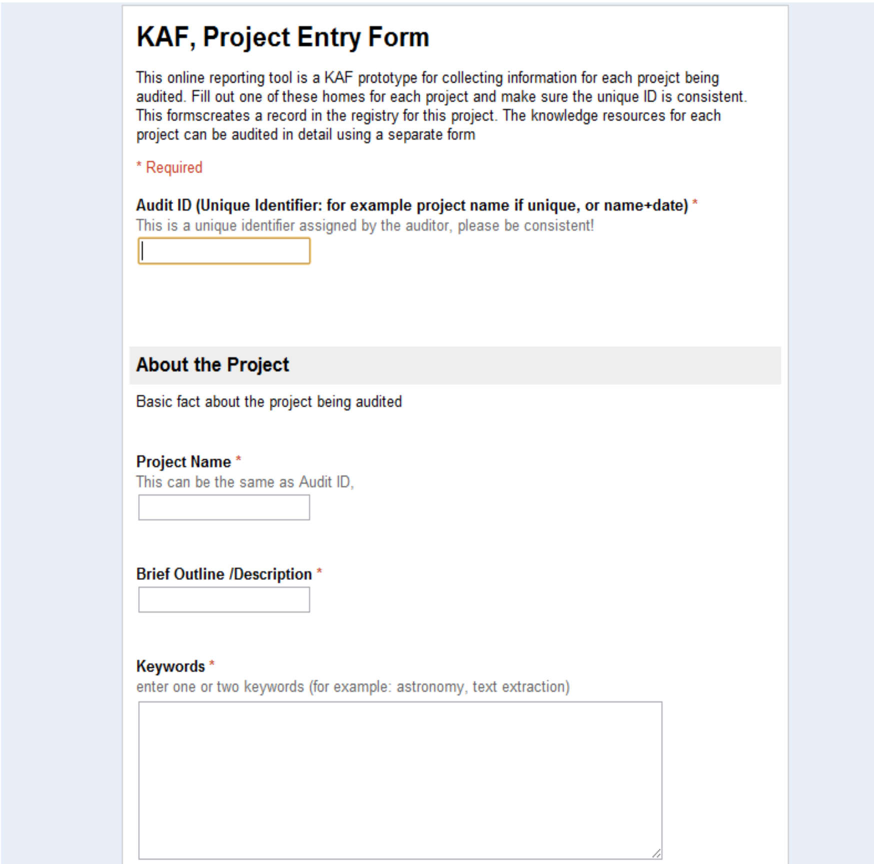
Table 12: Core SE Knowledge

The template schemas outlined above are provided as examples and 'in principle' guidance, however the actual collection form structure design and GUI (graphical user interface) is flexible and can built around and adapted.

4.3.4.5 The online collection tool

A prototype online collection form, is provided as part of this deliverable for the purpose of this study, and can be accessed using the following URI http://tinyurl.com/5rg4wyj. The prototype is implemented using a free instance of 'Google Apps' spreadsheet and form, and it uses the template above. A project website is developed with additional documentation that incorporates the collection tools above.

KAF website: Fig. 30 is a screenshot of the first portion of the auditing template of the KAF prototype implemented for this research.[23]


KAF, Project Entry Form

This online reporting tool is a KAF prototype for collecting information for each proejct being audited. Fill out one of these homes for each project and make sure the unique ID is consistent. This formscreates a record in the registry for this project. The knowledge resources for each project can be audited in detail using a separate form

* Required

Audit ID (Unique Identifier: for example project name if unique, or name+date) *
This is a unique identifier assigned by the auditor, please be consistent!

About the Project

Basic fact about the project being audited

Project Name *
This can be the same as Audit ID,

Brief Outline /Description *

Keywords *
enter one or two keywords (for example: astronomy, text extraction)

Fig. 30: Screendump of a KAF Template Prototype

[23] https://sites.google.com/site/kaframework/

4.3.4.6 Examples of communication

Appendix E provides examples of communication with the project team:

a) a note to inform the funding body of audits taking place, describing the purpose, scope and process of the audit,

b) a note to inform the project team/leader and obtain input and guidance,

c) an email to the project team leader to verify/validate the correctness of the report containing preliminary findings,

d) a note to communicate the final findings and issuing recommendations.

4.4 OPEN ACCESS MONITOR (OAM)

Open Access Monitor is a customisation of KAF, since KAF was initially being implemented with DAF as a blueprint – it inherited a certain degree of inherent granularity, typical of data oriented information structures. Three factors influenced the decision to morph KAF into OAM:

1. KAF targets the auditing of knowledge resources in some detail, with a corresponding level of effort, but the outcome of the initial pilots and cursory assessment of the field, indicated that many publicly-funded projects do not even have a website. Therefore many of the questions in the KAF template are irrelevant to most of the projects being audited. Some questions were hard to answer, and even more difficult to analyse statistically, and would have required a high amount of qualitative evaluation and interpretation of the result, which would have resulted in this research exceeding its allocated resources.

KAF is useful and fulfils its purpose of auditing knowledge resources, but was not a particularly good fit to serve the research requirements of the particular domain being audited. The knowledge model developed as a guideline for the KAF auditing template however, serves well as an example used in the final recommendation section of this thesis.

2. Given the limited amount of resources available for this doctoral research, and given there is no need (nor the time) to pitch the audits at such granular level, a more lightweight instrument was devised, to capture the essence of 'knowledge sharing' practices based on the availability of 'Open Access' resources, with less effort than KAF.

3. OAM is equipped with a star ranking system, a form of heuristic evaluation to rank the audited projects according to their level of knowledge *sharedness*. The rationale of the star ranking system is presented in Section 4.4.2.

4. OAM includes an additional template (not provided in KAF) to survey the knowledge sharing and Open Access policies.

4.4.1 OAM Process and Template

OAM is modelled on KAF, from which it inherits the main structure, which includes a process, a template, an online collection tool, and some templates of communication and correspondence to serve as supporting guidelines for auditors who may wish to adopt them. OAM supports two processes:

OAM Process A: the auditing of Open Access or other knowledge sharing policies

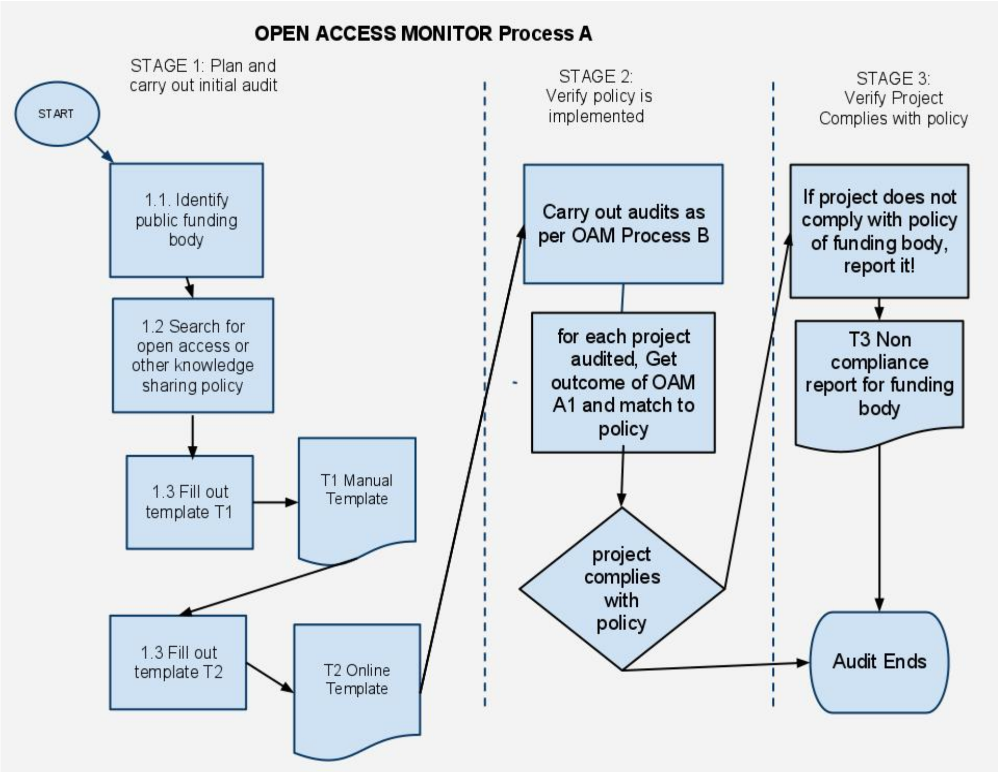


Fig. 31: OAM Process A

Open Access policies of publicly-funded bodies, in case these are not already known, are reported 'non-compliant', when an auditor finds that a particular project does not offer Open Access resources as prescribed by the policy. This process uses two templates: T1a and T2a (two versions), and T3 described below.

NOTE: in the case of our study, the Open Access policy was already known from literature review, so this process and template were not used; they are however provided in the public website to facilitate future uses of the OAM framework.

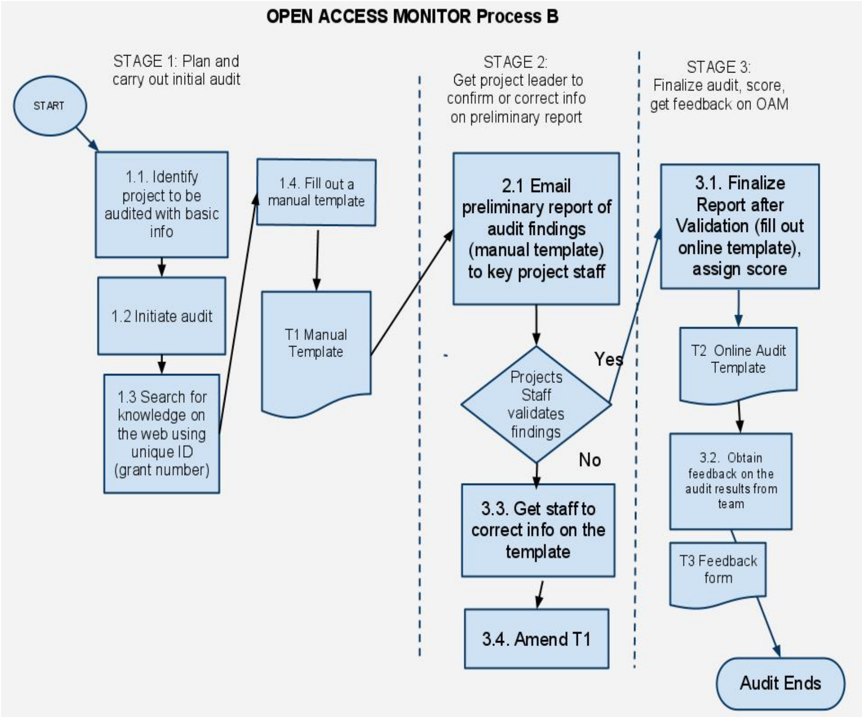


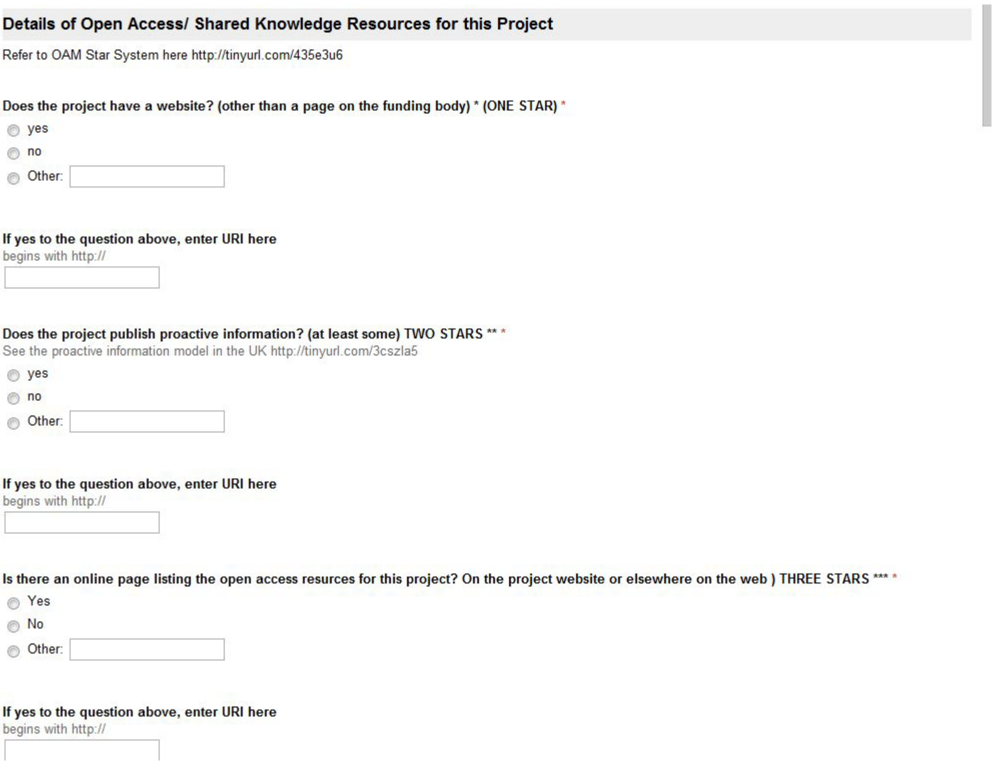
Fig. 32: OAM Process B

OAM Process B: the auditing of publicly-funded projects

The auditing process of OAM is essentially exactly the same as KAF, but some of the phases have been merged to save time, without detracting thoroughness. So OAM audits are carried out across three steps, instead of four (see diagrams in Figs. 31 and 32).

OAM Templates

The templates associated with Process A are T1a and T2a, and T3a whereby T1 is the manual version of T2, essentially the same template structure applies, but the first version is 'manual', either on paper or on any other document format that the auditor wishes to keep a copy of, while T2 is represented by the online form currently available in the prototype.


Details of Open Access/ Shared Knowledge Resources for this Project

Refer to OAM Star System here http://tinyurl.com/435e3u6

Does the project have a website? (other than a page on the funding body) * (ONE STAR) *

- yes
- no
- Other:

If yes to the question above, enter URI here

begins with http://

Does the project publish proactive information? (at least some) TWO STARS ** *

See the proactive information model in the UK http://tinyurl.com/3cszla5

- yes
- no
- Other:

If yes to the question above, enter URI here

begins with http://

Is there an online page listing the open access resurces for this project? On the project website or elsewhere on the web) THREE STARS * ***

- Yes
- No
- Other:

If yes to the question above, enter URI here

begins with http://

Fig. 33: Screendump of OAM Template

It should be noted that for consistency, the template for Process A has been loosely modelled using a similar table format to the Digital Curation Centre[24] but later modified to overcome one of its limitations, by adding a column to 'monitor' (Fig. 34).

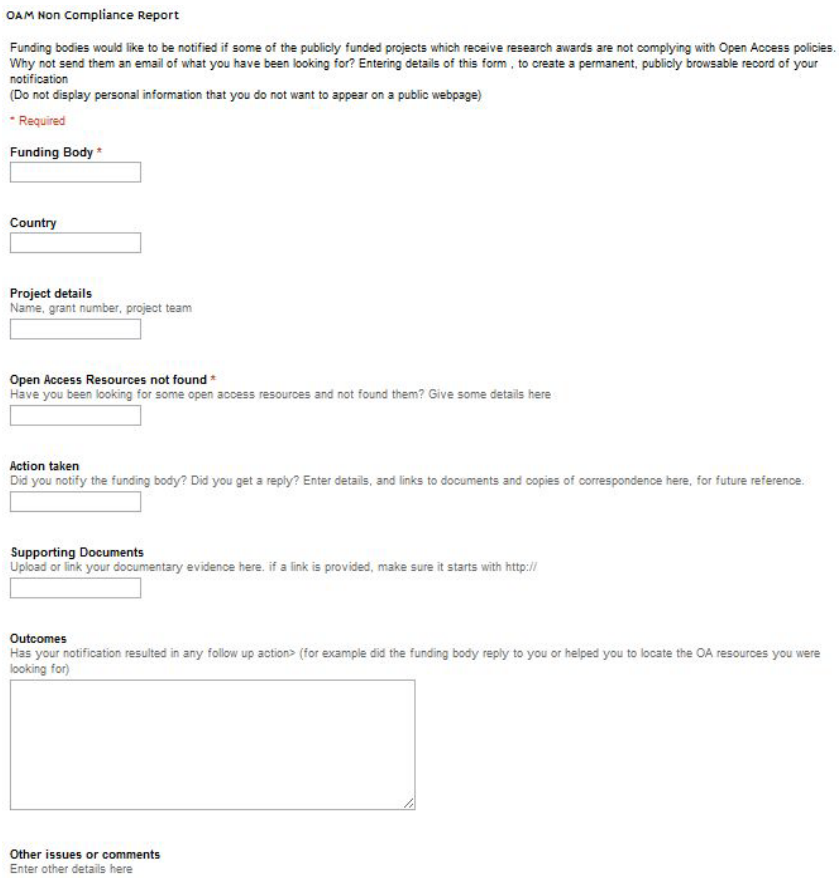
One of the empirical findings from triangulating the evidence gathered via the DCC report with an FOI request, revealed a discrepancy in the information, possibly caused by the fact that the DCC table does not distinguish between 'monitoring of the scholarly outputs' and 'monitoring of the data'. In the next chapter more documentation of this empirical finding is provided.

[24] http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies

	Policy coverage		Policy stipulations				
Research Funders	Published outputs	Data	Time limits	Data plan	Access / sharing	Long-term curation	Monitoring
Arts and Humanities Research Council	●	●	●	●	●	◐	○
Biotechnology & Biological Sciences Research Council	●	●	●	●	●	●	●
Cancer Research UK	●	●	●	●	●	●	●
Engineering and Physical Sciences Research Council	●	●	●	○	●	●	●
Economic and Social Research Council	●	●	●	●	●	●	●
Medical Research Council	●	●	●	●	●	●	○
Natural Environment Research Council	●	●	●	●	●	●	●
Science and Technology Facilities Council	●	○	●	○	●	○	○
Wellcome Trust	●	●	●	●	●	●	●

Fig. 34: Modified DCC Table

In process A, Template 3 (T3a) consists of a 'non compliance report'; in this case some Open Access resources expected for a given project are not found.


OAM Non Compliance Report

Funding bodies would like to be notified if some of the publicly funded projects which receive research awards are not complying with Open Access policies. Why not send them an email of what you have been looking for? Entering details of this form , to create a permanent, publicly browsable record of your notification
(Do not display personal information that you do not want to appear on a public webpage)

* Required

Funding Body *

Country

Project details
Name, grant number, project team

Open Access Resources not found *
Have you been looking for some open access resources and not found them? Give some details here

Action taken
Did you notify the funding body? Did you get a reply? Enter details, and links to documents and copies of correspondence here, for future reference.

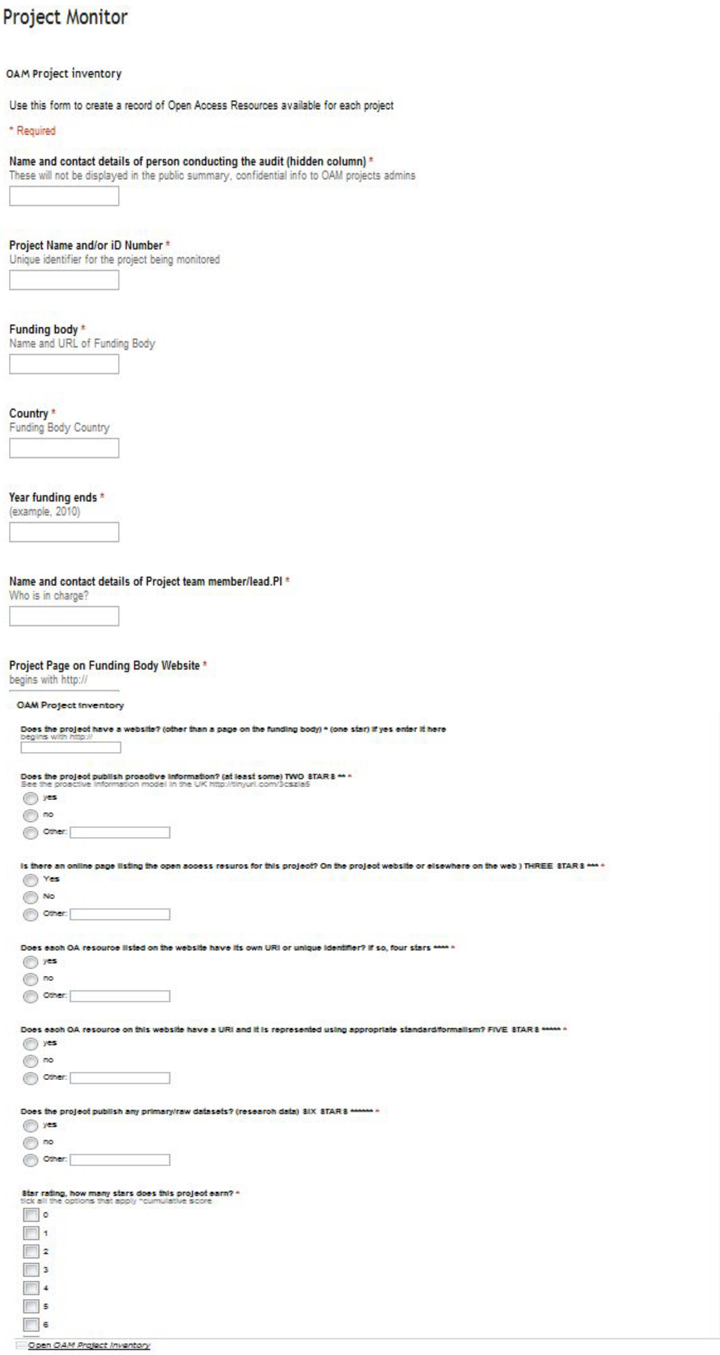
Supporting Documents
Upload or link your documentary evidence here. if a link is provided, make sure it starts with http://

Outcomes
Has your notification resulted in any follow up action> (for example did the funding body reply to you or helped you to locate the OA resources you were looking for)

Other issues or comments
Enter other details here

Fig. 35: OAM Non Compliance Template

The same applies to the templates associated with Process B, T1b and T2b and T3. In process B T3 is the feedback form, essentially allowing any user of OAM, whether an auditor or a general member of the public, to provide feedback on the OAM process and outcomes.


Project Monitor

OAM Project inventory

Use this form to create a record of Open Access Resources available for each project

* Required

Name and contact details of person conducting the audit (hidden column) *
These will not be displayed in the public summary, confidential info to OAM projects admins

Project Name and/or iD Number *
Unique identifier for the project being monitored

Funding body *
Name and URL of Funding Body

Country *
Funding Body Country

Year funding ends *
(example, 2010)

Name and contact details of Project team member/lead.PI *
Who is in charge?

Project Page on Funding Body Website *
begins with http://

OAM Project inventory

Does the project have a website? (other than a page on the funding body) * (one star) if yes enter it here
begins with http://

Does the project publish proactive information? (at least some) TWO STARS ** *
See the proactive information model in the UK http://tinyurl.com/3cszla5

- yes
- no
- Other:

Is there an online page listing the open access resuros for this project? On the project website or elsewhere on the web) THREE STARS * ***

- Yes
- No
- Other:

Does each OA resource listed on the website have its own URI or unique identifier? If so, four stars ** ***

- yes
- no
- Other:

Does each OA resource on this website have a URI and it is represented using appropriate standard/formalism? FIVE STARS *** ***

- yes
- no
- Other:

Does the project publish any primary/raw datasets? (research data) SIX STARS **** ***

- yes
- no
- Other:

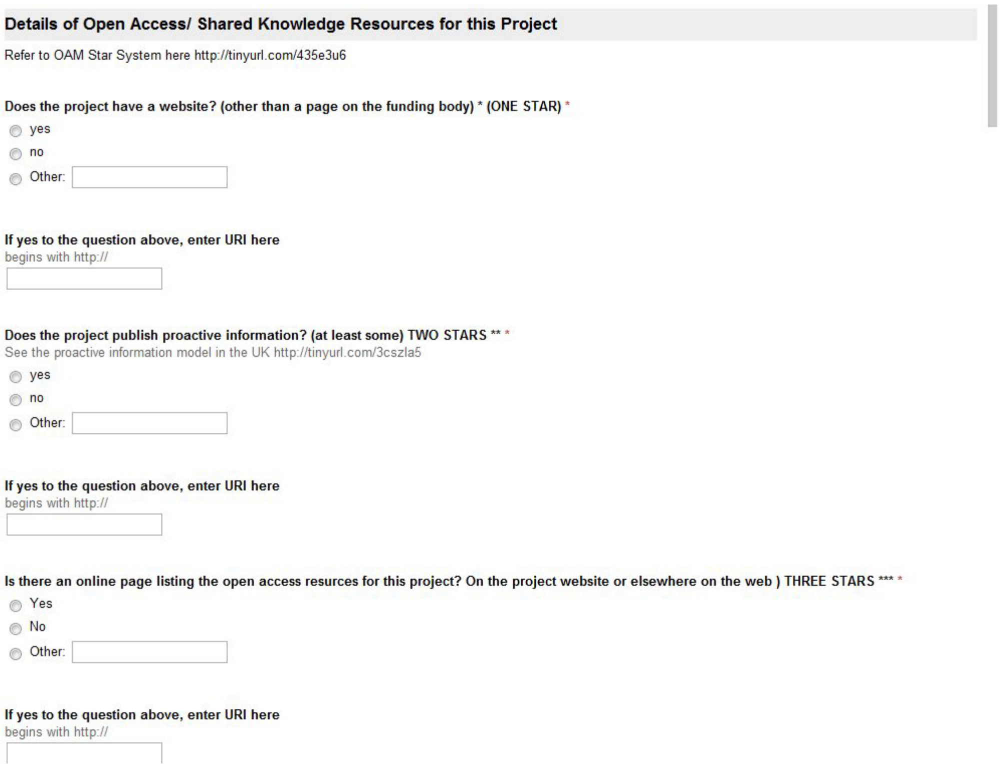
Star rating, how many stars does this project earn? *
tick all the options that apply *cumulative score

- 0
- 1
- 2
- 3
- 4
- 5
- 6

Open OAM Project Inventory

Fig. 36a: Screendump of Process B template

The most recent and updated version of the OAM funded project template conforms to the requirement that every knowledge resource should correspond to a URI, and has been modelled to elicit URIs accordingly; see Fig.36b.


Details of Open Access/ Shared Knowledge Resources for this Project

Refer to OAM Star System here http://tinyurl.com/435e3u6

Does the project have a website? (other than a page on the funding body) * (ONE STAR) *

yes

no

Other:

If yes to the question above, enter URI here

begins with http://

Does the project publish proactive information? (at least some) TWO STARS ** *

See the proactive information model in the UK http://tinyurl.com/3cszla5

yes

no

Other:

If yes to the question above, enter URI here

begins with http://

Is there an online page listing the open access resurces for this project? On the project website or elsewhere on the web) THREE STARS * ***

Yes

No

Other:

If yes to the question above, enter URI here

begins with http://

Fig. 36b: Screendump OAM Template Final

4.4.2 OAM Indicators: Star ranking for heuristic evaluation

The typical analysis of Open Access usage historically has revolved around performance data (hit, view and visit) providing insights into traffic, ranking of entry pages (Jamali et al., 2005) or a statistical evaluation of *green* over *gold* access (Gómez et al., 2009). A more detailed set of indicators looks into whether Open Access resources are reached via search engines, back links and direct access (Mayr, 2006).

Here a set of pragmatic Open Access indicators are developed based on a set of heuristics, unspoken rules of thumb: an expert chef will not measure numerically or mathematically how many grains or grams of salt a dish requires, and has a sense of what a 'pinch' means, in relation to the size of the pan, the size of the grains of salt, the type of salt etc. and can even distinguish meaningfully between a pinch and a *good pinch* even (a sense of measure that is hard to interpret for next generation of robot chefs). It's the kind of knowledge that is not easily codified, however often the only one that works.

Heuristic-based methods have historically been used in usability engineering (Nielsen, 1994), but have since been adopted by the system related professions, to capture and evaluate human factors and behaviours (Rizzo et al., 2005), as well as in system engineering research (Valerdi, 2010).

In this research heuristics are used in the competence questions to evaluate KAF described in the section above, and a 'set of heuristics' developed to evaluate the level of sharing of artefacts, implemented via a scoring schema that here we called OA Star System, that can also serve as a prototype for 'Open Access' indicators. The model (and inspiration) for a heuristic Open Access 'star system' is TBL 'star system' for linked data[25]. The OAM star system provides some rules of thumb to help 'benchmark' and enable an 'in principle' quantitative measure to assess the level of compliance with Open Access policies, as well as good practices in web-based knowledge sharing, as prescribed in literature, and discussed in Chapter 3 of this thesis, such as using URI and shared, explicit knowledge schemas.

[25] http://www.w3.org/DesignIssues/LinkedData.html

The set of heuristics used in OAM is illustrated in Fig.37, and justified in the following paragraph. Each heuristic is described briefly below.

★ Project funded with taxpayers money has a website

★★ The project website proactively publishes project information, including how is funding allocated, how decisions are made , or at least some of it (see the proactive information model by ICO UK)

★★★ The project has a webpage listing all its Open Access knowledge resources delivered with public funding (papers, vocabularies, diagrams, specifications, rules, code etc)

★★★★ Each resource has a unique web address/URI

★★★★★ Each resource has a unique URI and adopts open standards and appropriate knowledge sharing representation formalism and notation*

★★★★★★ The project shares primary/raw datasets

Fig. 37: Screendump of heuristic OAM indicators

1. Project/Organisation has a website.

Sounds trivial, but it is surprising how many projects and organisations still do not have a website, despite the availability of free and intuitive (foolproof) website-making tools. Websites, or online project pages, are considered one of the best mechanisms for knowledge sharing on the web; this is later proven empirically by establishing a positive correlation in the audited portfolio between the existence of a website, and a high OAM score, as shown in Fig 44b, Section 5.2.2.5.

2. Proactive Information.

Publishing proactive information has been recommended by progressive governmental institutions for years, as a cost effective way for the tax payer to provide access to knowledge of public initiatives and organisations.

One of the early models of proactive information is published by the UK government, referenced by OAM.[26] Proactive information is considered good practice in that it saves time to the knowledge seeker, who can grasp the essential with one glance, instead of having to search, guess and become frustrated, and to the taxpayer, because it reduces the level of FOI requests public authorities have to answer.

3. Project /Organisation has a web page listing all the Open Access resources.

In addition to publishing information about the project/organisations, a page lists all the publicly-accessible knowledge resources. Considered good practice because it saves time and facilitates access, retrieval and re-use.

4. Every resource has a unique URI.

The uniqueness of resource identifiers is essential to the web functioning as a knowledge base, as discussed in the early pages of this dissertation.

5. Unique URI and appropriate representation/formalism.

A unique web address for each resource is good practice, but well-structured, appropriately represented, expressed and formalised resources reduces the cognitive load in humans, and the computational load in machines (such as semantic technologies and search engines). OAM does not prescribe what level of formalisation is appropriate, and leaves this choice to the knowledge publisher.

6. Share raw datasets.

This is something relatively new, under debate at the time of writing. The sixth star anticipates that raw datasets will also be mandated in future Open Access policies. NOTE: in this study the sixth star was not applied.

---

26 http://www.ico.gov.uk/upload/documents/library/freedom_of_information/detailed_specialist_guides/generic_scheme_v1.0.pdf

4.5 EVALUATION

One of the main criteria for evaluation in systems engineering is 'fitness for purpose', which is referred to in literature as a type of 'heuristic' evaluation.[27] The KAF methodology is considered valid if, after performing the knowledge asset audit, it becomes possible to answer one or more of the following questions:

1. What knowledge resources are shared, therefore publicly reusable?
2. Where (on the web) are these resources located?
3. What knowledge sharing mechanism/techniques are adopted (or not)?
4. Who is responsible in the organisation for making knowledge shareable?
5. Is the organisation adopting/following a knowledge sharing and re-use policy?

Four pilot cases were carried out to evaluate the 'fitness for purpose' of the auditing instrument, and in each of the four cases one or more of the questions above became answerable. The results of the pilot audits are reported in a later section. Similarly, OAM is evaluated heuristically against competence questions – an OAM audit is considered valid if after performing the audit it becomes possible to answer one or more of the following questions:

- Does the project have a website?
- Does it publish proactive information?
- Does it have a webpage listing all resources?

[27] Defence Test & Evaluation Strategy http://www.mod.uk/NR/rdonlyres/4252C64B-6340-4168-92D4-23FCC6177BFE/0/te_strategy.pdf

- Does each OA resource have a URI?
- Does each resource have a unique URI and adopt an appropriate formalism?
- Does it publish raw data?

4.6 CONCLUSION

Chapter 4 of this thesis discusses the rationale and justification for the development of *ad hoc* collection instruments in support of this research, KAF and OAM, and presents their architecture and components such as their core process, templates and supporting materials and links to online prototypes.

A novel set of heuristics is devised (the OAM star system) to enable parametric 'ranking' of different knowledge sharing behaviours in publicly-funded projects, as well as their comparative and statistical analysis, which are provided in forthcoming chapters. Literature references that justify the use of heuristics are provided.

The next part of the thesis presents non-empirical as well as empirical analysis of RC1 and RC2, and an analysis of the respective findings.

"Some problems are so complex that you
have to be highly intelligent and well informed
just to be undecided about them."

Laurence J. Peter

CHAPTER 5

FIELDWORK

OVERVIEW OF THE SECTION

This section presents research components 1 and 2, respectively, as introduced in the methodology chapter.

Component 1 consists of a study of the policy landscape in relation to 'knowledge sharing'. Many policies exist 'on paper' to support and encourage knowledge sharing, but they have different names, scope and reach.

Goal: To identify which policies exist that support 'knowledge sharing' in relation to scholarly outputs, and how these policies reflect on grant holders' behaviours in practice, specifically in the domain under investigation.

Method: A combination of literature review, data base searches, and ongoing exchanges with experts and peers who specialise in either policy evaluation or Open Access. In some cases, ongoing feedback was provided by experts as 'peer review' of the draft versions of this section of the thesis. The section contains a discussion, the main findings and a qualitative analysis of the findings.

Component 2 presents the background and rationale to the audits being carried out, together with findings, and an analysis of the same.

Goal: The audits have a twofold function: first of all, to evaluate the suitability of the collection instruments presented in Chapter 4. The instruments and the prototype implementation have been developed iteratively, thanks to the feedback received from various contributors including researchers, principal investigators and independent auditors recruited for this part of the study.

Secondly, the audits constitute the main collection mechanism to gather empirical data of Open Access adoption in the target population segment, to help measure the gap between theory and practice.

Method: The auditing tools, consisting of processes and templates, supported by fully working online prototypes have been developed as described in Chapter 4. A target population segment is identified, and basic data gathered in a document called the projects portfolio, attached in Appendix G. Each project in the segment is 'audited' using the OAM process and template, and scored to be able to perform some analytics. This section includes a discussion, a justification for the selection of the audits and corresponding inclusion criteria, and presents the aggregate findings of all the audits, and an analysis of the same.

5. 1 Research Component 1: Knowledge Sharing Policies

5.1.1 Exploratory Evaluation

5.1.2 Impact Evaluation

5.1.3 Analysis of the Findings

Conclusion

5. 2 Research Component 2: The Audits

5.2.1 The Pilots (carried out with KAF)

5.2.2 The Full Audits Portfolio (OAM)

5.2.3 Analysis of the Findings

Conclusion

5.3 Overall Conclusion

5.1 RESEARCH COMPONENT 1: KNOWLEDGE SHARING POLICIES

The primary intended goal of this research component is to understand what regulatory apparatus exists that governs knowledge sharing and level of access to publicly-funded scholarly research. As it emerges from the analysis reported further below in this section, in the UK alone each individual Research Council has its own version of an Open Access policy, emanating from a UK-wide (RCUK) policy. In turn, the RCUK policy has been devised to comply with international and global trends towards making research outputs more accessible, originating from a decade-long of key related initiatives in the UK and abroad. This part of the thesis shows that what emerges from this comparative evaluation of the different regulatory initiatives is a fragmented regulatory landscape, where systemic discrepancies become evident when the various policies are considered 'as a whole'. It is not within the scope of this paper to analyse in detail each single aspect and implication of the intricate and uneven legal regulations, since this would require a team of experts, and exceed the resources available for this research at this stage. It is also not the scope of this section to enumerate every single piece of literature and article or study carried out on the subject of 'Scholarly Publications', an overview of which is included in Chapter 3 of this thesis.

This component of the study aims to capture the 'systemic' aspects of knowledge sharing policies, as discussed in the research challenges section at the beginning of this dissertation, to try to understand 'what works'. The main terms of reference for this enquiry are the core guiding principles established by the 'Policy Evaluation Framework' (Davies et al., 2005) adopted in the UK public sector to answer the following questions:

Exploratory evaluation: what knowledge sharing policies are there?

Impact evaluation: do people know about these policies? How are these policies adopted in the field? (surveys, audits)

The initial finding of this research component, carried out via the analysis of relevant literature (Digital Curation Centre report, 2009) is sufficient to answer 'in principle' the research question in relation to the current focus of the study (what knowledge sharing policies exist that support access to scholarly outputs in systems engineering research in the UK): the funding council for systems engineering research, EPSRC, has had an Open Access policy since 2005 (Appendix F), however this does not guarantee that 'knowledge' is adequately represented, shared, nor accessible to a knowledge seeker. Furthermore, when probing further the content of literature with FOI requests yields unexpectedly contradictory outcomes: different dates, different interpretations of the scope of the policy and its monitoring are provided by different individuals in the research council itself, and related institutions (Appendix F).

Overall, the findings of RC2 are to be considered 'formative', in that they constitute an ongoing effort carried out at a time when policies and legislation are still being discussed by the larger community of academics, researchers, policy makers and members of the public, and the rules of the game are subject to various currents of political influence, and change by the day. As such, this section of the thesis serves also as a working document to support ongoing exchanges and further analysis among different stakeholders involved in the process.

Chapter 3 of this thesis explains why pragmatic, technical knowledge is best represented via explicit 'technical' artefacts, such as explicit knowledge models, which are inherently codified, and why to be shared these should adopt appropriate formalisms and notation, and that technical knowledge sharing may extend beyond pure Open Access to scholarly publications, which per se, does not guarantee 'access to knowledge'. This section considers what social, organisational behavioural mechanisms (codified norms, such as policies) exist to promote knowledge sharing. In particular this research analyses Open Access policies, which are devised to increase access to publicly-funded research outputs, without necessarily taking into account the expectations and behaviours of knowledge seekers in the respective communities. In the last two decades, to coincide with the increased adoption of the internet, knowledge sharing policies and initiatives have proliferated, aimed at encouraging web-based dissemination of knowledge, especially in the domain of public sector information (PSI), which in some cases - depending on the definitions adopted by the respective governing bodies - overlap with the scholarly domain, where knowledge is generated by publicly-funded research. Many of these initiatives have been triggered by high-level strategic international agreements promoted by a diverse set of individuals and organisations, such as, for example OECD.

5.1.1 Exploratory Evaluation

Initiatives aimed at increasing 'access to knowledge' of scholarly research outputs have proliferated in recent years, and some have resulted in policies and legislation, however as of today, few of these initiatives have translated into legislation, and their effectiveness, in terms of supporting access of knowledge seekers to scholarly outputs, is still limited and not fully understood. The section that follows provides an overview of these initiatives, with an initial analysis and evaluation.

5.1.1.1 International declarations (OECD, Budapest, Berlin)

5.1.1.2 International directives (EU PSI 2003)

5.1.1.3 National legislations (that apply in a single member state to all governing bodies, such as the FOI Act 2000)

5.1.1.4 National policies of each governing body (for example, each individual research council in the UK have their own policy to govern Open Access)

5.1.1.5 Other general initiatives, such as A2K

5.1.1.6 Knowledge transfer policies

5.1.1.1 International declarations

The time line of early Open Access initiatives can be traced back to the sixties (Suber, 2009) however the Budapest Initiative in 2003 is flagged as a key milestone (Budapest Open Access Declaration) which initiated a major progressive contemporary global movement fostering the promotion of free access to scholarly knowledge. The Budapest Declaration says:

"By 'Open Access' we mean free availability [of scholarly journal publications] on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited." (Budapest Open Access Initiative)

Other initiatives followed, notably the Bethesda Statement on Open Access Publishing in June 2003, and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (Berlin Declaration) which was agreed in a 'landmark meeting organised by the Max Plank Society and the European Cultural Heritage Online project, bringing together international experts with the aim of developing a new web-based research environment using the Open Access paradigm as a mechanism for having scientific knowledge and cultural heritage accessible worldwide.' According to the Bethesda and to the Berlin Declaration, Open Access contributions must satisfy two conditions:

1) The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use.

2) A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in an appropriate standard electronic format is deposited (and thus published) in at least one online repository using suitable technical standards (such as the Open Archive definitions) that is supported and maintained by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable Open Access, unrestricted distribution, inter operability, and long-term archiving.

Other high level international initiatives contribute to the mission of widening access to knowledge, for example the OECD Declaration on International Science and technology Cooperation For Sustainable Development published in 2004 (OECD Declaration) declared to support the exchange of information about effective methods to promote sustainable development through science and technology. This set the direction, at least in theory, for further developments. As of today, knowledge sharing is reported to be one of the pillars of the OECD strategy for economy development:

> *"The G-20 recognizes that knowledge sharing is a distinct tool for promoting growth and development that is complementary to finance and technical cooperation."* (G20 DWG, 2011)

However it should be noted that the initiatives above are not reflected in legislation, and at the time of writing, no law exists (yet) to mandate or monitor the extent of Open Access publishing.

5.1.1.2 EU PSI directive

Public Sector Information has always been one of the main sources of primary data for many research activities and data-centric services in modern economies, but thanks to the current explosion of web-based technology applications and infrastructure many more opportunities are opening up for a variety of stakeholders. The Council and the European Parliament adopted a Directive on the re-use of public sector information which deals with the way public sector bodies should enhance re-use of their information resources which, the EU says, is based on two key pillars of the internal market: transparency and fair competition. The directive establishes minimum rules for the re-use of PSI throughout the European Union, but encourages member states to go beyond minimum rules and to adopt open data policies, allowing a broad use of documents held by public sector bodies. (EU PSI Re-use Directive, 2003).

Individual member states have adopted the Directive with different legislative instruments and local variations (Implementation of the PSI directive). Interestingly, research institutions are excluded from the EU PSI directive with a comma in its Article 1:

> *"Art 1. - The Directive shall not apply to [...]*
> *e) documents held by educational and research establishments, such as schools, universities, archives, libraries and research facilities including, where relevant, organisations established for the transfer of research results;"*

Additionally, the following information was provided by the Office of the National Archive:

> *"Neither the EU PSI Directive of 2003 nor the UK Re-use of Public Sector Information Regulations 2005 specifically define what public sector information is. However, both the EU Directive and the UK Regulations make clear it covers information and content that is held by public sector bodies that fall within the scope of the Directive Regulations and where the rights are held by the public sector body."* (Office of The National Archive, 2011, personal correspondence)

For the purpose of FOI legislation however (at least in the UK) universities are considered 'public authorities' and must comply with FOI legislation.

5.1.1.3 National legislation (UK)

The most notable example of legislation aimed at making accessible and transparent public sector information, is the FOI Act.

In the UK publicly-funded Research Institutions and Universities are considered 'public authorities', and therefore PSI legislation applies (FOI Act 2000):

"53, 1 The governing body of:

an institution within the further education sector,

(b) a university receiving financial support under section 65 of the Further and Higher Education Act 1992,

(c) an institution conducted by a higher education corporation,

(d) a designated institution for the purposes of Part II of the Further and Higher Education Act 1992 as defined by section 72(3) of that Act, or

(e) any college, school, hall or other institution of a university which falls within paragraph (b)."

However, under the transposed Regulation 1515, comma 3, b (derived from the EU Directive, 2003) PSI regulations do not apply to universities and research institutions as these, among others, are 'exempt' (Regulation 1515, 2005). In the UK, the FOI Act and the Regulation 1515 both aim at increasing 'access to knowledge', however they seem to be conflicting in their definition of *public authority.*

5.1.1.4 National policies of individual governing bodies (UK)

At national level, each governing body responsible for a public sector may adopt a different version of the relevant policy. For example in the UK, each of the five major research councils has a different position in relation to:

a) Open Access and b) data sharing.

In further work, we reserve to undertake a more detailed comparative analysis of the same.

A brief outline of these policies, as examples of such differences, is summarised in Table 13:

AHRC	http://tinyurl.com/69eo5rk
BBSRC	http://tinyurl.com/63qzbfq
EPSRC	http://tinyurl.com/6jpg7x5
ESRC	http://tinyurl.com/6xfywjw
MRC	http://tinyurl.com/6e4wmgb
NERC	http://tinyurl.com/5scd86y
STFC	http://tinyurl.com/6jj52ly

Table 13: Brief Outlines of Policies of Individual Governing Bodies (UK)

The differences in wording, scope and reach of the various policies of each Research Council in the UK contribute to the challenge of understanding, managing and even monitoring the policy adoption, especially when a research project is funded by more than one funding council. A comparative analysis of these policies is provided in Table 14, acquired from a report that attempts to summarise schematically similarities and differences (Jones, 2009) courtesy of the Digital Curation Centre (DCC), and serves as the baseline outcome of RC1 for this study.

● Full Coverage ◐ Partial Coverage ○ No Coverage

	Policy Coverage		Policy Stipulations					Support Provided			
Research Funders	Published outputs	Data	Time limits	Data plan	Access/ sharing	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	◐	○	●	○	◐	○
BBSRC	●	●	●	●	●	●	●	●	●	◐	●
CRUK	●	●	●	●	●	●	●	◐	●	○	○
EPSRC	●	●	●	○	●	●	●	◐	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	◐
MRC	●	●	●	●	●	●	○	◐	●	○	◐
NERC	●	●	●	●	●	●	●	●	●	●	◐
STFC	●	●	●	●	●	●	●	◐	●	◐	○
Wellcome Trust	●	●	●	●	●	●	●	●	●	◐	●

Terminology Clarifications

- **Published outputs**: a policy on published outputs e.g. journal articles and conference papers
- **Data**: a datasets policy or statement on access to and maintenance of electronic resources
- **Time limits**: set timeframes for making content accessible or preserving research outputs
- **Data plan**: requirement to consider data creation, management or sharing in the grant application
- **Access/sharing**: promotion of OA journals, deposit in repositories, data sharing or reuse
- **Long-term curation**: stipulations on long-term maintenance and preservation of research outputs
- **Monitoring**: whether compliance is monitored or action taken such as withholding funds
- **Guidance**: provision of FAQs, best practice guides, toolkits, and support staff
- **Repository**: provision of a repository to make published research outputs accessible
- **Data centre**: provision of a data centre to curate unpublished electronic resources or data
- **Costs**: a willingness to meet publication fees and data management / sharing costs

Table 14: Overview of Funders Policies, DCC Report 2009[28]

The funding Council whose policy applies to the portfolio of audits under study in RC2 is the EPSRC, which is the funding body for Systems Engineering research in the UK as of 2010.

According to the table in Jones's report, and the corresponding legends, EPSRC 'monitors' the compliance to their access policy.

However, the outcome of an FOI request seems to contradict the information in the DCC report:

[28] http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies

26 August 2011

Dear Paola,

FREEDOM OF INFORMATION ACT (FOIA) 2000 REQUEST FOR INFORMATION

Thank you for your request for information dated 10 August 2011. EPSRC confirms it holds some of the information you have requested.

In your request you asked for the following information:

a. How has EPSRC monitored the uptake of Open Access policies since it first published its policy?

EPSRC recently announced two policy developments relating to access to research data – the first, announced in early May 2011, set out a general policy framework on the access to and sharing of EPSRC-funded research data; the second, announced in early June 2011, concerns the specific case of access to published peer-reviewed research articles. Full details of both are available on our website.

The policy framework sets out EPSRC's expectations concerning how the institutions and researchers it supports will manage and provide access to EPSRC-funded research data. EPSRC recognises that the infrastructure and resources allocated to managing and sharing research data vary widely among the research organisations it funds. EPSRC also acknowledges that to comply with the framework research organisations will need to review (and may need to change) their research data management practices and capabilities. EPSRC recognises that not all the identified changes may be achievable in the short term, but may require action in the medium to long term.

To allow time for this EPSRC expects all those it funds to have developed a clear roadmap to align their policies and processes with EPSRC's expectations by 1st May 2012, and to be fully compliant with these expectations by 1st May 2015. EPSRC will monitor progress and compliance on a case by case basis, but, given the stated timescales, has not monitored progress to-date. EPSRC will investigate identified cases of non-compliance, and if in due course it appears that proper sharing of research data is being obstructed EPSRC reserves the right to impose appropriate sanctions.

Fig. 38: Screendump Reporting Excerpt of Reply from EPSRC to FOI Request

Although this study focuses on EPSRC, further FOI requests were made to other UK Research Councils, to verify whether the discrepancy between literature and FOI outputs were applicable to other funding councils. Also in the case of BBSRC, which according to the DCC table above also monitors compliance, the FOI request states that this is not the case, see the excerpts from correspondence below:

19 September 2011

Reference: FoI-132

Dear Dr Di Maio,

FREEDOM OF INFORMATION ACT 2000 – REQUEST FOR INFORMATION

I am writing in response to the request you made to BBSRC under the Freedom of Information Act 2000.

and

Can I see some example of documentation/data supporting the outcomes of monitoring activities?

As we do not yet have formal monitoring in place, we do not have any related documentation or data.

The discrepancy between the table in the DCC report and the outcome of FOI request may be caused, at least in part, by the table layout, whereby a single column indicates 'monitoring activities' for both the scholarly outputs, and the research data.[29] However the FOI responses seem to diametrically contradict some of the information provided *prima facie* by corresponding reference literature. To study and fully understand and manage the causes and the extent of such contradictions is beyond the immediate scope of this thesis, and is set aside as a key pointer for work to be done. At the time of writing the issue is under review and, possibly, if the discrepancy is confirmed, the recommendation that follows will include an additional, separate column added to the table to report the monitoring of the scholarly outputs (journal publications) and data separately, and a suggestion that a hyper-link to corresponding documentary 'evidence' to support every claim made in the table, and other references.

[29] Private email exchange with Angus Whyte, DCC and Sarah Jones, HATII Glasgow, October 2011

5.1.1.5 Access 2 Knowledge (A2K)

Another class of initiatives known collectively as Access to Knowledge consists of a *"set of principles emerged from a loose collection of different social movements, which in turn resulted from to changes in economy and society produced by new information technologies."* (Balkin, 2006). A2K initiatives have contributed to generate awareness of Open Access practices, and to the principles that have become reflected to some extent in generic 'policies' however, they have not resulted in practical arrangements or legislation to support their adoption in as far as this enquiry has been able to establish so far[30].

5.1.1.6 Knowledge Transfer (KT)

Knowledge transfer (KT) can be used to describe the knowledge flows between different units, divisions, or organisations rather than individuals (Szulanski et al., 2004), the emphasis of KT is on generating income from knowledge transfer activities, rather than maximising access to knowledge. KT is also defined as

> *"the process through which one unit (e.g., group, department, or division) is affected by the experience of another"* (Argote & Ingram, 2000).

A definition of KT commonly adopted in EU and UK is:

> *"The process by which the knowledge, expertise and intellectually linked assets of Higher Education Institutions are constructively applied beyond Higher Education for the wider benefit of the economy and society, through two-way engagement with business, the public sector, cultural and community partners."*
>
> (Unico Report, 2008)

[30] Private Correspondence with Denise Nicholson and colleagues @Witwatersrand University

At international level, the EU Commission says that member states should:

> *"ensure that all public research organisations define knowledge transfer as a strategic mission" and that they should "support the development of knowledge transfer capacity and skills in public research organisations, as well as measures to raise the awareness and skills of students – in particular in the area of science and technology – regarding intellectual property, knowledge transfer and entrepreneurship"*
>
> (EU Commission Recommendation, 2008)

and that

> *"The EU Commission also states that it wants to move towards a position in which " knowledge transfer between universities and industry is made a permanent political and operational priority for all public research funding bodies within a Member State, at both national and regional level."*
>
> (EU Commission Recommendation, 2008)

Figure 39 (Unico Report, 2008) shows a diagrammatic rendition of a model of Knowledge Transfer, emphasizing the 'economic activity' aspect of knowledge flows:

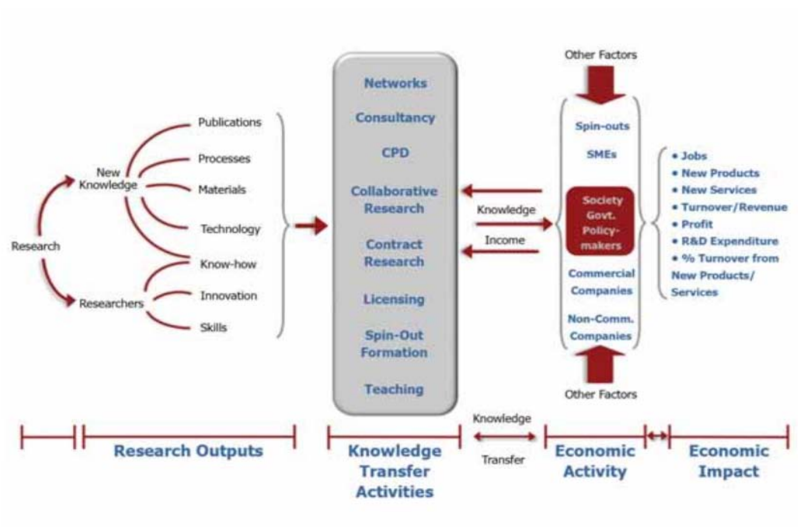


Fig. 39: Model of Knowledge Transfer within the Innovation Ecosystem (Source: University of Glasgow) in: Metrics for the Evaluation of Knowledge Transfer Activities at Universities

Knowledge Transfer principles consist of restricting access to knowledge, to allow for the commercial exploitation of knowledge resources, and generate income streams via the sale of educational materials such as teaching, consulting services and licensing mechanisms - essentially, without further clarification, KT appears to be in direct contrast with Open Access principles, which ensure that access to knowledge is free and unrestricted. Furthermore, Knowledge Transfer policies shape and mandate the knowledge exchange perceptions and behaviours at praxis level (Hauser, 2010), and Intellectual Property (IP) clauses of commercial contracts part of 'Knowledge transfer programmes' restrict and constrain knowledge flows between academia and industry, effectively pre-empting Open Access policies to take hold (Gardner et al., 2010). In the summer of 2011, The Global Congress on Intellectual Property and the Public Interest, August 25-27, 2011, convened over 180 experts from 32 countries and six continents to help re-articulate the public interest dimension in intellectual property law and policy, known as ' The Washington Declaration'[31] which addresses two main points:

1) International intellectual property policy affects a broad range of interests within society, not just those of rights holders. Thus, intellectual property policy making should be conducted through mechanisms of transparency and openness that encourage broad public participation. The declaration advocates new rules.

2) Markets alone cannot be relied upon to achieve a just allocation of information goods — that is, one that promotes the full range of human values at stake in intellectual property systems.

One of the recommendations of the declaration is to encourage evidence-based policy making:

[31] http://infojustice.org/washington-declaration-html

"Most would agree that research used in policy making should meet basic standards of transparency. Yet the intellectual property policy debates of the past two decades have not done so. Industry-funded research dominates the intellectual property policy conversation, yet virtually none of the major industry-sponsored studies document their methods, assumptions, or underlying data in any detail. The institutions responsible for intellectual property policy making have failed to exert enough pressure for either transparency or quality—and in many cases have relied on discredited statistics in their own statements."

5.1.2 Impact Evaluation

Having established in the preceding section that several initiatives and different kinds of policies and regulations exist (exploratory evaluation), the question 'do people know about these policies?' falls under 'impact evaluation' (Davies et al., 2005). One of the conclusions of a UK-wide study carried out in 2008 says:

"Researchers remain poorly informed about Open Access. Awareness is growing but still only slowly and there remain many misconceptions. Researchers are eager to maximise their own impact and reputation but do not understand what means and opportunities are available to them." (Swan, 2008)

Unfortunately, the raw data used to generate the conclusion above was not available for further analysis since JISC, despite being a promoter of openness and transparency in research, does not have a policy that requires grant holders to retain and publish research data.[32]

[32] Private Correspondence with A. Swan and JISC Officials

In a series of ethnographic observations, as introduced in the methodology chapter and carried out in various occasions throughout the duration of the study, diverse communities of academics and practitioners in the engineering and systems engineering fields of practice were observed in relation to knowledge sharing behaviour and awareness of Open Access. Information was gathered from the participants observed around the four issues given below, in the context of conversations and professional exchanges, at the national and international academic and professional events listed in Appendix D:

a. what is Open Access?

b. what is the Budapest declaration?

c. would you agree that engineering knowledge on the web is shared using technical artefacts (diagrams, vocabularies etc), to journal publications?

d. would examples of knowledge sharing models and schemas be useful to provide guidance on how to best share knowledge on the web?

Answers were recorded manually on various pieces of papers (and napkins, when the exchanges took place over a coffee or a meal), then reported by hand in a log book, and summarized for legibility in Table 15.

ETHNOGRAPHIC SETTING	DATE	Individuals consulted and observed (approx)	Q1 YES	Q1 DONT KNOW	Q2 YES	Q2N DONT KNOW	Q3 NO	Q3 YES	Q4 Yes	Q4 No
Walk in engineering campus (6 weeks, local)	2009	30	0	30	0	30	0	30	30	0
Systems engineering networking meeting (1 day, national)	2010	24	0	24	0	24	0	24	24	0
Space symposium (local)	2010	14	0	14	0	14	0	14	14	0
Systems Engineering conferences (4 days, international)	2010	30	4	26	0	30	5	25	30	0
Systems Engineering Conference (2 days, international)	2011	22	3	19	2	20	4	18	22	0
Total		120	7	113	2	118	9	111	120	0

Table 15: Summary Log of Ethnographic Observations

The impact evaluation carried out via direct observations and informal conversations with approximately 120 individuals over a period of 3 years, confirms the finding of the earlier study referenced above, which had a focus on the UK. In conclusion, the majority of academics and practitioners in systems engineering research observed, have limited or no awareness of Open Access policies, (1 and 2), and agree that technical knowledge is generally shared using artefacts such as diagrams, and every participant would find useful to have examples or models of knowledge sharing schemas.

5.1.3 Analysis of the Findings of Research Component 1

When considering the regulatory landscape for knowledge sharing 'as a whole', the lack of alignment and cohesion between the many policies, initiatives and relevant legislative instruments listed above becomes evident, revealing some discrepancies and contradictions:

1. The policy landscape is fragmented across different levels, for example, different policies address loosely different layers of the information management chain such as Data, Information and Knowledge.

2. Policies target roughly the same 'knowledge sharing' space, but they are not harmonised, and are not evenly supported by legislation, so that some, for example, knowledge transfer initiatives under contract law are legally binding, while it is unclear what sanctions and provisions exist for the lack of compliance with Open Access policies.

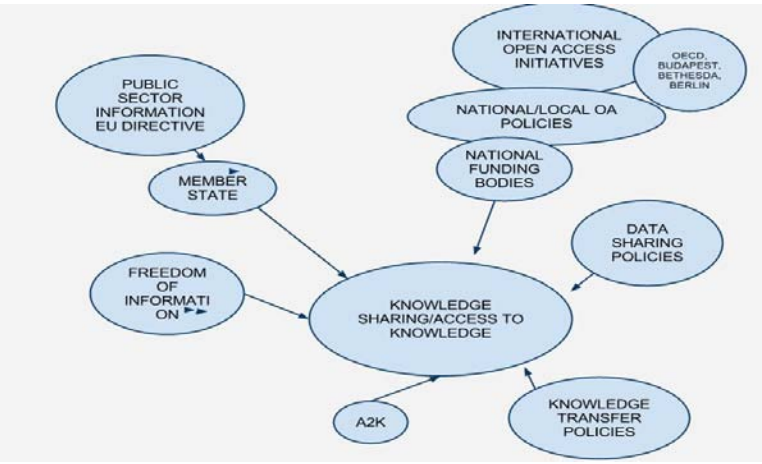


Fig. 40: Fragmented Landscape of Policies and Initiatives that Regulate Knowledge Sharing

3. Some of the current legal provisions for the protection of Intellectual Property, and programmes such as 'Knowledge transfer' that restrict knowledge flows between academia and industry for the purpose of commercialisation, in the absence of any guidelines and specification, could be in conflict with Open Access policies.

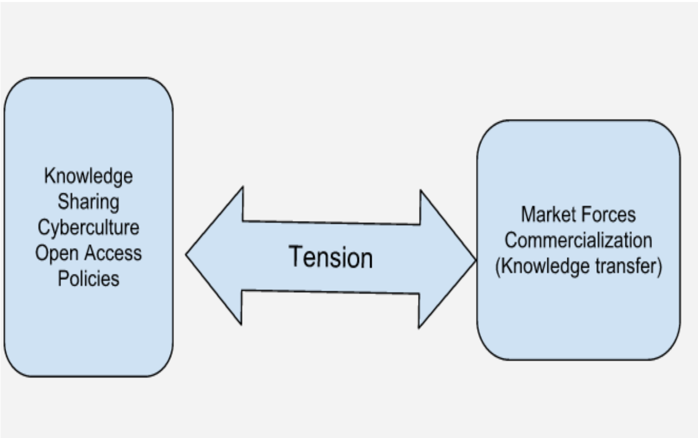


Fig. 41: Tension between Knowledge Sharing Initiatives and Knowledge Transfer

4. There is a lack of consistent, shared and commonly adopted definitions and concepts. For example in relation to what is 'public authority', according to the sources reviewed, in the UK the PSI directive does not apply to universities and research institutions, however the FOI Act does refer to Universities as Public Authorities. Open Access policies do apply to publicly-funded research through awards granted to research and educational institutions, such as Universities.

5. There is a lack of a clear specification of what content Open Access publications should contain (what portion of the research outputs, for example) and what technical artefacts (notation, formalism, content types) should be adopted to optimise access to knowledge sharing. No formal or informal 'Knowledge Object Model' is defined for Open Access.

6. There is a lack of documented, evidence-based monitoring mechanisms for policy implementation across the board. Monitoring initiatives seem to be notional, and not supported by documentation.

7. There is also a lack of integration between the social aspects of knowledge sharing practices, such as social norms and policies, and the technical aspects of knowledge sharing, i.e. the adoption of the appropriate artefacts and processes.

8. Some of the artefacts used in key references, such as Table 13 reported in Section 5.1.1.4, may require further optimisation to reduce possible ambiguity.

Conclusion

The intended goal of this research component is to answer the following question:

What knowledge sharing policies exist to support access to scholarly research outputs?

In relation to systems engineering research in the UK, for which the funding council is EPSRC, it is established that an Open Access policy applies, however this policy as such does not grant access to project 'knowledge' in the form of technical artefacts. No technical specification or minimum knowledge sharing requirements are associated with the policy to prescribe or recommend good practices, as they emerge from other research disciplines such as web sciences. An unintended outcome of this research component is the mapping and comparative evaluation of a variety of knowledge sharing initiatives and policies, which is to be considered formative and ongoing.

The second question this research component is designed to answer is:

Do people know about Open Access policies?

Previous studies concluded, among other things, that there is a poor level of awareness of Open Access, based on a generic sample of the UK academic population (Swan, 2008). These conclusions are confirmed and further validated by ongoing empirical ethnographic observations in a specific academic field, engineering, reported in Table 14.

5.2 RESEARCH COMPONENT 2: THE AUDITS

A systematic review of 100 audits of publicly-funded projects was undertaken, and the outcomes investigated via meta-analysis, that is, using a quantitative, statistical evaluation of the findings. Pilot audits were carried out at different stages of development of the auditing tools, KAF and DAF.

As mentioned in the introduction of Chapter 5, the audits have two functions, one to evaluate whether the auditing tools (the method, the templates, the prototypes) are 'fit for purpose' and the other is to gather empirical data about the hypothesised adoption of Open Access policies of the sample of projects under investigation (RQ2, H2). This section presents and justifies the rationale behind inclusion criteria, describes the pilots as well as the fully audited portfolio, and presents and analyses the main findings and conclusions. Validation is remanded to Chapter 6. The research methodology and collection tools generated in this research are domain independent, which means they can be used to research and audit any field, however the focus of this study is constrained to the systems engineering domain in the UK, to be consistent with the initial motivation for the work, as discussed in the methodology chapter, as well as to ensure the feasibility of the study within the resources.

5.2.1 Piloting the Audits

Pilot cases have been selected from a portfolio of research projects funded by Research Councils, according to the following broad characteristics:

- Projects are related to the 'systems engineering' domain either directly - funded as 'systems engineering' projects, or indirectly - the project deliverable is a system, even if it is not categorised explicitly under system engineering by their funding body.
- They are UK-based projects publicly-funded through one or more UK research councils.

Several pilots were carried out at various stages of KAF development illustrated in Chapter 4 of this thesis, four of which, selected semi-randomly, have been documented and the outcome of which is summarised in Tables 16-19:

- Assist – BBSRC
- Pathtext – BBSRC
- Nectise – EPSRC
- Widefield Camera (WFCAM) – STFC

ASSIST: This project was funded to investigate text mining using two case studies in the social sciences. An online project page was promptly located via searches, included hyperlinks to project information and ample documentation, and made use of multimedia and diagrams. However, no system specification was located.

PROJECT DETAILS	INFO	URL	NOTATION	FORMAT
AUDIT ID	Assist	http://www.nactem.ac.uk/assist/		
Funding Council	BBSRC			
What knowledge resources generated by this project are publicly available?	High level Diagram, Presentation		Diagram, narrative	Html PDF
What knowledge resources are not available?	No Glossary, No Detailed system diagram, No Code available			
Where are they located?	On the web, on project page			
What KS mechanisms are adopted?	n/a			
Who is responsible for making the KS decision?	Project Lead			
Is the organisation conforming to a knowledge sharing policy? (including IP contractual constraints)	Broadly to BBSRC policy			

Table 16: KAF Development Outcome Summary - ASSIST Pilot

PATHTEXT: A project part of the NACTEM portfolio, (same as ASSIST) PATHTEXT serves as a bridge between the use of graphical notations used to represent 'collective knowledge' that the researchers call 'pathway visualizations', and it is believed that they could serve to support the development of user interfaces for knowledge discovery if they can be linked to the text in publications. As in the case of ASSIST, the project has a well documented online page with some technical content, although no system specification was located.

PROJECT DETAILS	INFO	URL	NOTATION	FORMAT
AUDIT ID	Pathtext	http://www.nactem.ac.uk/pathtext/		
Funding Council	BBSRC			
What knowledge resources generated by this project are publicly available?	papers		narrative	Html PDF
What knowledge resources are not available	system specs			
Where are they located?	website			
What KS mechanisms are adopted?	links to papers of project website			
Who is responsible for making the KS decision?	project leader			
Is the organisation conforming to a knowledge sharing policy? (including IP contractual constraints)	Open Access policy of funding body			

Table 17: KAF Development Outcome Summary - PATHTEXT Pilot

NECTISE: A research project designed to investigate the network capability readiness involving ten top UK universities, some of which host Systems Engineering Centres of Excellence (the Universities of Bath, Cambridge, Cranefield, Leeds, Leicester, Loughborough, Manchester, Queens, Strathclyde and York) in partnership with BAE, a leading industrial partner in the defence sector.

The online project page contains a static (not hyperlinked) list to papers and publications, however only a few of these could be retrieved via online searches. Further analysis of the website using DMOZ indexing, shows that some papers are available or hyperlinked as PDFs on the website, but not accessible from the home page, or from any of the website pages linked to the home pages. Some publications in PDF were located on secondary pages of the website using online searches, but not accessible from the website navigation menu.

PROJECT DETAILS	INFO	URL	NOTATION	FORMAT
AUDIT ID	Nectise	http://www.nectise.com		
Funding Council	EPSRC			
What knowledge resources generated by this project are publicly available?	Some papers are listed on the project page			
What knowledge resources are not available	Glossary, System Diagrams, Specification		narrative	html pdf
Where are they located?	Project Page on the web			
Who is responsible for making the KS decision?	Industry Partner			
Is the organisation conforming to a knowledge sharing policy? (including IP contractual constraints)	Contractual policy IP restrictions due to industrial partnership with BAE			

Table 18: KAF Development Outcome Summary - NECTISE Pilot

WFCAM: The goal of this research project is to develop an ambitious instrument, an infrared wide field camera, consisting of a cryogenic camera and associated equipment, reportedly the most capable IR imaging survey instrument in the world.

The online page for this project has the most detailed technical documentation encountered during the piloting phase.

PROJECT DETAILS	INFO	URL	NOTATION	FORMAT
AUDIT ID	WIDE FIELD CAMERA	http://www.jach.hawaii.edu/UKIRT/instruments/wfcam/		
Funding Council	STFC			
What knowledge resources generated by this project are publicly available?	detailed system diagrams for various components and processes	http://www.jach.hawaii.edu/UKIRT/instruments/wfcam/user_guide/description.html	diagram	html JPEG/ GIF
What knowledge resources are not available	Glossary			
Where are they located?	On the project page			
Who is responsible for making the KS decision?	Project lead			
Is the organisation conforming to a knowledge sharing policy? (including IP contractual constraints)	No, the funding body does not seem to have a policy			

Table 19: KAF Development Outcome Summary - WFCAM Pilot

5.2.1.1 Main findings of the pilots

Exploratory pilot audits were carried out on projects funded by different research councils to test the suitability of KAF as a collection tool, and refine the auditing instruments accordingly, as well as to take the initial pulse of possible outcome, four of which are reported above.

In relation to the first goal of the audits:

Is the auditing tool KAF adequate to capture and record project knowledge?

Performing audits using KAF has successfully enabled answering the competence questions, summarised above, therefore it can be said that KAF fulfils its purpose, and as such it is a valid collection instrument. However some limitations became evident during piloting:

- The KAF process is 'too rigidly structured', as sometimes an auditor (a knowledge seeker) needs to go back and forth through steps to actually get to locate the knowledge resources they are after, and the number and order of steps they follow needs to be more flexible.
- Templates are designed to capture a high level of detail for each knowledge resource, which may be desirable in some cases, however in the case of this study they may be overkill, in that an evaluation of the compliance with Open Access policies should be possible with lesser detail, resulting in a more feasible effort.
- The same KAF process can be completed with less iteration, therefore leveraging the principle of parsimony; the auditing process should be revised and optimised further.

In relation to the second goal of the audits:

Is the KAF tool useful to gather field data in relation to the adoption of Open Access policies?

KAF helps to gather data, and analysing the outcome of the pilots shows that the quality and detail of the knowledge sharing formalisms adopted by each project varies greatly, consisting generally of high level information with limited presence of accessible and reusable system knowledge artefacts, with some notable exceptions, as in the case of the WFCAM project.

Unexpectedly, despite the fact that the STFC is the Research Council which at the time the funding was awarded, had the least explicit commitment to Open Access, the online pages of the WFCAM project offered the highest level of technical detail of knowledge sharing resources found within the pilots. Although all projects tend to have a web page, this is generally notionally used, and no consistent standard exists as to what is should contain.

Lessons The pilots provided an opportunity to test run the first version of KAF, the *ad hoc* collection instrument devised to answer the research question in this doctoral thesis, however:

- More compact, less granular auditing templates could result in more targeted questions, and reduce the effort required for each audit (the effort criteria is essential for the successful completion of this research within allocated time and budget).
- While KAF is, in principle, useful to audit project knowledge, it is not designed with a built-in mechanism to allow the comparison between different projects, since it measures qualitative data, which is different for each project. A more objective, flexible, and more computable set of criteria needs to be developed and applied to the audits.
- A benchmarking/scoring mechanism against set criteria would be helpful to rank the projects and facilitate the analysis of the outcomes.
- The inclusion criteria for the portfolio of pilot audits were too broad, as they covered a range of different projects funded by more than one funding council. Given the limited resources, to help answer the research question with some degree of certainty, the study should be narrowed further.

The above lessons contributed to the development of Open Access Monitor (OAM), that adopts a modified, more refined version of the KAF process and template, as illustrated in Chapter 4 of this thesis, as well as narrower inclusion criteria for the full study, which has been limited to EPSRC-funded projects in systems engineering research. The rest of this section discusses the audits carried out using OAM.

5.2.2 The Full Portfolio and Outline of the Systematic Review Criteria

Testing KAF with the pilots resulted in a number of considerations and tacit intuitions as to the nature of knowledge auditing, summarily captured in the paragraphs above. To increase the meaningfulness and reach of a study, a more refined collection instrument was developed, also described in detail in Chapter 4 of the thesis, and the inclusion criteria for the full study were narrowed as follows:

- Only projects funded by EPSRC under 'systems engineering research' were audited in this part of the study (additional projects funded by other Research Councils were later audited for external validation, as reported in Chapter 6). A full portfolio of funded System Engineering research was sourced from the funding council itself (Appendix G). In the process, it was noted that funding councils across the UK do not adopt consistent terminology or categorisation criteria for what is considered 'systems engineering' research. The study of the lack of consistent terminology and categorisation criteria is remanded to related and further work. To reduce the bias that could have arisen from auditing a cross category sample, a single category sample was deemed more appropriate, therefore:

- Only projects which were already concluded at the time the audits were carried out were included in the sample.

One of the recurring 'excuses' of principal investigators for not publishing any research outputs online is "we haven't got around to doing that yet." It is understandable that during research different priorities compete, and sharing the findings may go the bottom of the list.

The audited sample therefore includes 'all' the EPSRC systems engineering research funded projects that ended in 2009 and 2010, resulting in a census of approximately 100 audits, which was a nice round figure that would simplify some statistical analysis of the finding.

The audited sample, together with the data sheet containing the findings is enclosed in Appendix H. The core criteria for the systematic review undertaken are summarised in the following table, adapted from PRISMA[33], a recent interdisciplinary initiative aimed at establishing high level evaluation parameters for systematic reviews:

	SYSTEMATIC REVIEW CRITERIA (ADAPTED FROM PRISMA)
	Review protocol: OPEN ACCESS MONITOR (OAM) Protocol URL : openaccessmonitor.org
	Study characteristics (100 publicly-funded projects were audited using OAM protocol, to evaluate whether the Intervention Open Access Policy as specified in literature and by the mandate of the respective public funding body is adhered to by grant holders. The projects were compared against each other and ranked according to OAM heuristics) Report characteristics (audited all EPSRC funded projects in Systems Engineering ended in 2009'/10, reported in English, outcomes published in IJCSI Journal October 2011) Inclusion criteria: publicly-funded projects in systems engineering research in the UK
	Information sources: all sources that could be located via online searches.
	Search Strategy: search the open web using any search engine chosen by the auditor (the investigator) using the following criteria: for grant number, project name and name of principal investigator to see if (three different auditors were used across the portfolio to reduce bias, and all the audits were carried out twice)
	Selection criteria: census mode, all projects matching the criteria were included in the review
	Data extraction : data compiled from a templates provided in the OAM protocol
	Risk of bias : 3 sources of bias are identified and discussed, outcomes verified and validated independently by different investigators, and only the validated dataset used to draw conclusions

[33] http://www.prisma-statement.org/

The design for the systematic review of the audits is represented in Fig. 42:

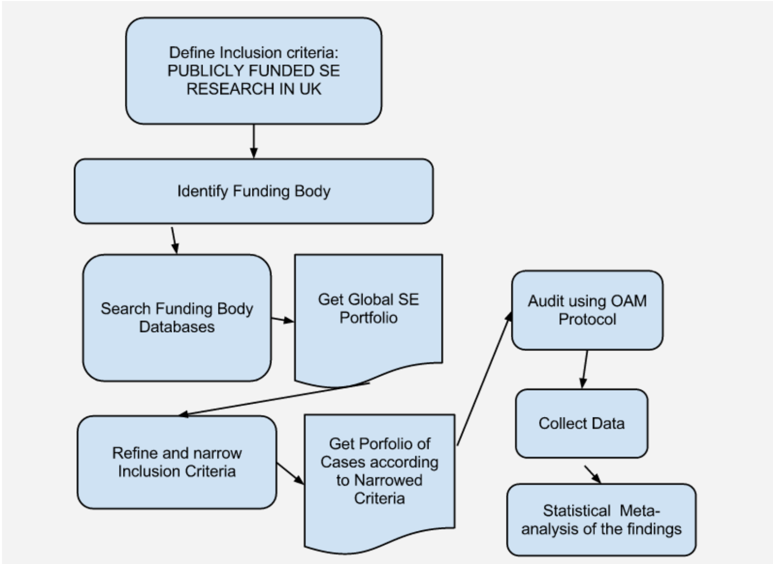


Fig. 42: Design of the Systematic Review

5.2.2.1 The auditors

One of the criteria for the validity of systematic reviews is the transparency and ability of replication of the research outcomes. Auditing frameworks such as KAF and OAM are designed to standardise the auditing process. This however does not mean that the results of the audits are uniform if performed by different auditors. A pool of three (paid) volunteer auditors was recruited to assist with the audits, in particular with cross validation and reliability testing, reported in the Chapter 6 (Quality and Evaluation of the findings). The auditors were university students in their early-mid twenties, from three different academic disciplines: engineering, bioinformatics and psycholinguistics, conversant with the use of electronic communication and the internet.

They were given an introductory training session on Open Access and auditing using OAM, and were then tasked with carrying out various sets of audits remotely, working from their own premises, although communication with the principal research and author of this thesis was provided as ongoing support.

5.2.2.2 The search task

A 'search task' (Bystrom & Hansen, cited in Poddar & Ruthven, 2010) can range from an unstructured statement of an information need from instructions to carry out a series of interactions, through to a description of an information problem, to a set of predefined actions. The search task was devised taking into account the fact that the auditor role is equivalent in this study to that of a 'knowledge seeker' who may not have any knowledge of what paper title to look for, however it does not at this stage consider personal information needs, or the operational context. It can be said that the task is artificially created and simulating an archetypal, albeit simplified model of knowledge seeker need. In the first round of audits, auditors were given a list of projects (Appendix J), and were asked to audit a subset of their choice from those ending in 2010 and 2009.

They could choose any key term from the project sheet to carry out their search, on Google, or any other search engine. In the first instance, they were asked to use any combination of search terms for no more than 10 minutes, which was a scheduled length to comply with allocated research timeframe and resources. In the second round (the re-audits) to increase the homogeneity of the results across different auditors, each was asked to perform the search using a standard algorithm as follows:

1) search for project name
2) then project name and grant number
3) then project name, grant number and name of PI

A further set of audits were performed to remove any bias that could be the cause of variation of search results due to causes other than different search criteria.

For example, a set of audits was carried out without writing to the PI to inform them that the audits were taking place, to eliminate the possibility that the email would prompt a researcher to publish its resources purely as the result of being made aware of the audit taking place. The test audits confirmed that the difference of search results between auditors is <3 scores, when the search terms are standardised and the PI not solicited to update the Open Access publications.

Limitations to future replication of the results

Since the web is dynamic and changes occur regularly, it can be expected that some of the results may not be exactly replicable, and that the longer time elapses since the study was carried out (Summer 2011), the more variations between outcomes can occur.

5.2.2.3 Meta-analysis of the systematic review of the findings

In the paragraphs that follow the scores of the audited projects are statistically analysed to search for patterns of behaviours: how many projects score high, meaning they make appropriate use of knowledge sharing artefacts and techniques? How many publicly-funded projects do not produce Open Access resources that can be located via web searches using standard search criteria?

The first round of OAM audits was carried out across 103 SE projects funded by EPSRC, using the OAM process (basically searching project name and/or grant number using open search engines such as Google), with the following results:

For the majority of the projects, no Open Access, or other resources were retrieved (the search yielded no results), (57 projects, score 0). The second largest group scored very low (15 projects, score 1). The third largest group however, scored the highest number of points (11 projects, score 15).

The findings are summarised in Fig. 43.

Score	Number of Audits
0	57
15	11
14	2
10	6
6	5
3	5
2	1
1	15

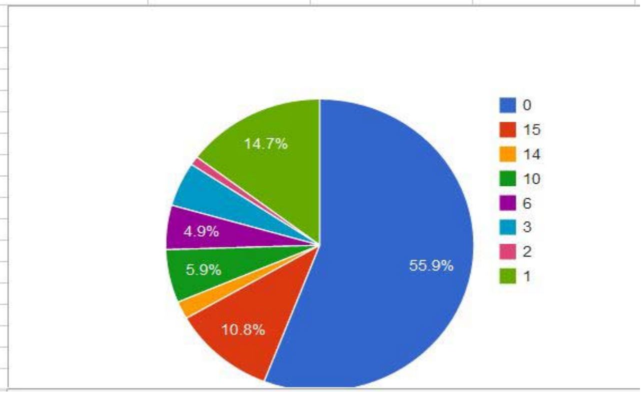


Fig. 43: Summary of Audit Findings

The bad news: more than fifty percent of the projects audited did not share any knowledge resources online that were identifiable via the project name or project audit. This does not mean that the resources are 'not shared', but that they could not be found, retrieved and accessed via standard web searches when the first audit was initially performed. The large majority of the projects audited (73) yielded less than a score 3, which corresponds to minimal knowledge sharing. If this finding is validated, it fully confirms the hypothesis that 'Open Access policies are not adopted' by the majority of publicly-funded research projects in systems engineering in the UK.

The good news: The third largest group in this research sample, (11) actually meets all the knowledge sharing criteria established by the OAM star rating benchmark, and scores the highest rank (15).

This outcome, if confirmed after verification, can be interpreted as a small success story in itself; while the large majority of research projects do not share knowledge according to commonly agreed good practices, approximately 10% of the audited sample does so in full and in compliance with good practices.

5.2.2.4 Qualitative analysis of the findings

The 11 projects which scored the highest are identified as 'knowledge sharing champions' and are being studied qualitatively with a close up individual evaluation, and targeted with case studies to acquire qualitative understanding of the circumstances that lead to such shining knowledge sharing performance. These may be written up and reported in future work.

5.2.2.5 Further hypotheses

The dataset obtained via the audits allows the exploration of a further hypothesis:

H4. - There is a correlation between the project having a website and the high overall ranking in the OAM benchmark

By selecting the relevant columns in the data sheet of the audited portfolio (does the project have a website, where the answer is yes, and the score column) then sorting and clustering them, as shown in Fig. 44a, a positive correlation can be observed, confirming the hypothesis is true. Therefore it can be concluded that although Open Access policies do not prescribe the use of project websites as a mechanism for knowledge dissemination, and although shared knowledge resources and research outcomes can be retrieved via web searches independently from whether the project has a website or not, the overall score tends to be higher where the project has a web page.

SCORE		PROJECTS HAVE WEBSITE
15	11	YES
14	1	NO
10	6	YES
6	5	YES
1	13	YES
0	58	NO

Fig. 44a: Numeric Representation of Positive Correlation between High Score and Project Having a Website

The correlation shown numerically in Fig.44a is rendered graphically in Fig.44b, where the pink area of the diagram shows the projects with a high score and the blue area of the diagram shows the projects that have a website that could be retrieved via searches.

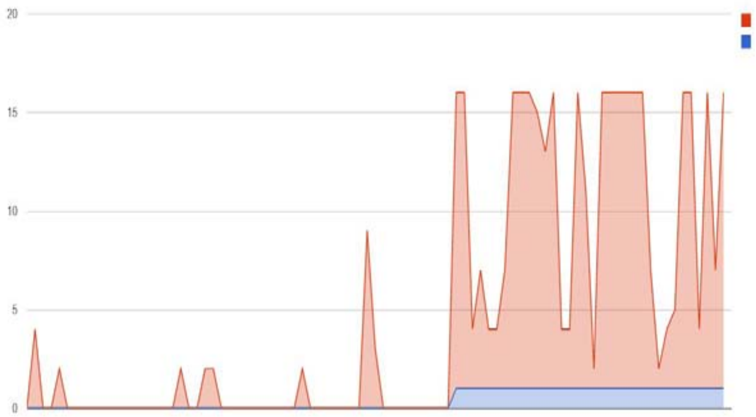


Fig. 44b: Graphic Representation of Positive Correlation between High Score and Project Having a Website

H5. Is there a correlation between the existence of a partner and the score? This hypothesis could also be phrased as 'does the participation of a project partner (as indicated on the grant page of the funding body) have an impact the Open Access outputs of a research project? '

To test this hypothesis, further investigation is undertaken on all the projects that score 0 and 15 from the main audited portfolio, for a total of 60 projects. A manual search is undertaken to verify which of these have a project partner, the summary table enclosed is further summarised in Fig.45 (see Appendix K):

	A	B	C	D
1	SCORE	HAS PARTNER	DOES NOT HAVE PARTNER	
2	15	5	6	
3	0	16	33	
4				
5				
6				

Fig. 45: Screendump of Correlation Score/Partner

A meta-analysis of these findings suggests that:

no partner = 0

partner = 1

Where The Project Has A Partner,

The Score Is 15 = 5 Projects

Where The Project Does Not Have A Partner

The Score Is 15 = 6 Projects

Where The Project Has A Partner

The Score Is Zero = 16 Projects

Where The Project Does Not Have A Partner

The Score Is Zero = 33 Projects

No significant correlation is detected between projects having a partner and their OAM score, in the sample under study. H5 is therefore false, according to these findings. No further investigation is undertaken to answer this hypothesis.

5.3 OVERALL CONCLUSION

This section of the thesis presents the work carried out under the two main research components, RC1 and RC2. The goal of RC1 is to determine what policies exist in support of knowledge sharing, and how they apply to the target domain (knowledge sharing in systems engineering) so that the gap between policy and practice can be measured and evaluated. The goal of RC2 is to evaluate the target population of publicly-funded projects selected according to the inclusion criteria described, against the heuristics developed in Chapter 5 of the thesis.

RC1 establishes that:

- Although the overall policy landscape is fragmentary, the governing policy of the funding body under study (EPSRC) operates an 'Open Access policy' according to which at least some knowledge resources, for example journal publications, should be openly accessible on the internet. Open Access policies however do not promote unrestricted access to more technical 'knowledge artefacts' such as diagrams, system requirements, specifications, process flows and data dictionaries.
- There are discrepancies between different sources of information, and evidence gathered via FOI requests contradicts, at least in part, the evidence gathered via the literature review.

RC2 establishes that a limited percentage of the projects audited generate Open Access resources that can be accessed via keyword searches using public search engines. According to the statistical analysis of the empirical data generated by this study, in relation to the sample being investigated, the main proposition and hypothesis H2 of this study is **true** in over 50 percent of the cases.

Further hypotheses, H4 and H5, are formulated and tested as being true and false respectively.

"You can know the name of a bird in all the languages of the world,
yet know absolutely nothing whatever about the bird…
So let's look at the bird and see what it's doing -- that's what counts.
I learned very early the difference between knowing the name of something and knowing something."

Richard Feynman

CHAPTER 6

QUALITY AND EVALUATION

OVERVIEW OF THE SECTION

Much is said and written about research quality and evaluation. The challenges to ensure and evaluate the quality of this research are many fold – given the variety and diversity of research methods adopted, no single qualitative approach is comprehensive enough, or fully applicable.

> *"Qualitative research methodologies vary considerably in their aims and epistemological assumptions and these, in turn, fundamentally shape the methods or procedures employed and evaluation criteria used."* (Finlay)

As the methodology consists of an *ad hoc* combination of research techniques, similarly, different evaluation techniques apply for each component. This section introduces general considerations in relation to quality assurance, discusses standard quantitative and qualitative evaluation criteria according to different literature and epistemological stances, and points to novel standards applied in evidence-based research, used to evaluate systemic reviews and meta as well as meta-analyses. It then lists the evaluation carried out in relation to each individual component of the study, and concludes with a self assessment of the doctoral thesis.

6.1 A Word about Quality Assurance

6.2 Quality in Mixed Method Research

6.3 Quality and Evaluation for this Research

6.3.1 Evaluation of the Methodology

6.3.2 Evaluation of the Instruments

6.3.3 Evaluation of the Results

6.4 Overall Self-Assessment of the Thesis

6.1 A WORD ABOUT QUALITY ASSURANCE

'Quality', similarly to 'knowledge', can be easily defined by perceived and subjective criteria, rather than by measurable hard facts. The standards for quality assurance vary from sector to sector, and often from country to country, however:

> *"Academic quality involves adherence to key principles such as intellectual rigour, accurate recording and honest reporting of results, and integrity in recognising the work of other researchers. But there may be legitimate differences of view as to the quality of, for example, notably innovative or groundbreaking work; and notions of quality may in some cases be related to specific contexts such as the scope or remit of a research funder or publisher."*[34]

An example of the complex multi-dimensionality of quality, is provided in Fig. 46, adapted from literature in Quality Management (Garvin, 1988).

[34] RIN http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/quality-assurance-and-assessment-scholarly-researc

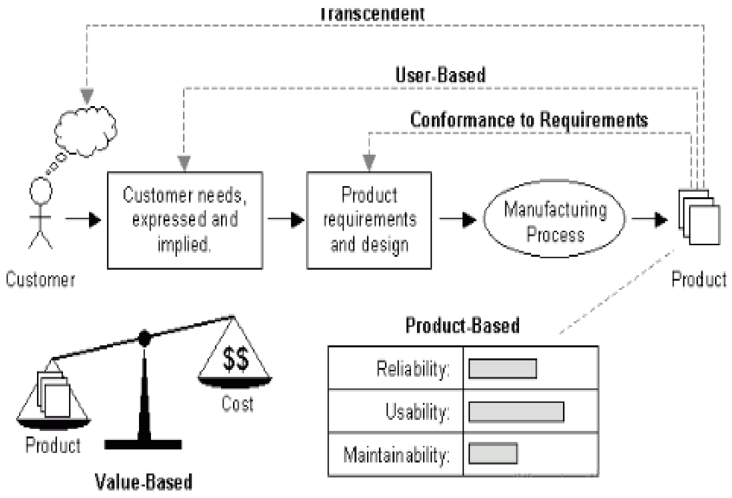


Fig. 46: Dimensions of Quality, adapted from Garvin (1988)

Figure 46 shows how different definitions and parameters exist for defining quality, which depends on the views and expectations of different stakeholders. Although the dimensions in Fig. 46 apply to quality management of products, rather than to research, they are a good example to illustrate the complexities associated with ensuring, managing and evaluating quality.

6.2 QUALITY IN MIXED METHOD RESEARCH

In quantitative research, quality is estimated in relation to validity, which is typically articulated in four stages (Cook & Campbell, 1979):

Conclusion Validity: whether a relationship exists between two variables

Internal Validity: to determine if this relationship is causal

Construct Validity: if the theoretical model is well depicted by the means through which it was operationalised

External Validity: if and to what extent findings can be generalised

In qualitative research, these validity criteria can be transposed to equivalent, broader categories, as shown in Table 20 below, to provide a better fit for what is described as a naturalistic/interpretive paradigm of enquiry, which this research follows, at least in part. It is noted however, that the validity of the conclusion criterion that towers over the others in Cook and Campbell (1979), is not directly translated into Lincoln and Guba's table:

Conventional inquiry	**Naturalistic inquiry**	**Methods to ensure quality**
Internal validity	Credibility	Member checks; prolonged engagement in the field; data triangulation
External validity	Transferability	Thick description of setting and/or participants
Reliability	Dependability	Audit – researcher's documentation of data, methods and decisions; researcher triangulation
Objectivity	Confirmability	Audit and reflexivity

Table 20: Qualitative Validity Terms (Guba & Lincoln, 1988)

> *"There is general agreement among qualitative researchers about the importance of critically evaluating research through the application of criteria. However, when it comes to choosing criteria, there is considerable divergence of opinion."* (Willig, 2001)

Another view suggests that:

> *"Whether quantitative or qualitative, research should rely on objectivity, internal validity, external validity, reliability, rigor, open-mindedness, and honest and thorough reporting."* (Ragin et al. 2003; Shavelson & Towne 2002; Wooding & Grant 2003, in Focus Technical Brief 2005)

This research adopts a mix methods design, that makes use of both qualitative and quantitative techniques, however validity and reliability which are considered evaluation criteria primarily in quantitative research, are adopted in this study, in particular in relation to the quantitative components, as discussed further down this section. Much emphasis, and some pressure, is placed to publish in 'quality journals' to demonstrate that research is of a certain quality, however, research should not be judged solely by whether or not it is published in the leading journals, and it is wrong to assume that *"research that is published in journals or cited by others is accurate, reliable, valid, free of bias, nonfraudulent, or of sufficient quality"* (Boaz & Ashby, 2003). Nonetheless, portions of this thesis have already been published in refereed conference proceedings and journals, as shown in Appendix A.

> *"Qualitative researchers have a responsibility to make their epistemological position clear, conduct their research in a manner consistent with that position, and present their findings in a way that allows them to be evaluated properly"* (Madill, 2000)

Research 'honesty' (Savin-Badin & Fisher, 2002), epistemological clarity and accurate reporting are also important quality factors in evaluation of research emphasised in this dissertation. In recent years several benchmarks have been developed to ensure that despite the heterogeneity of mixed methods adopted in a study, especially systematic reviews and meta-analyses, standard criteria can be used for their evaluation.[35]

---

[35] National Center for the Dissemination of Disability Research TECHNICAL BRIEF NO. 9 2005 http://www.ncddr.org/kt/products/focus/focus9/

To this effect this research adopts a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist, a benchmark established in 2009 to help authors report a wide array of systematic reviews which provides as standardised mechanism to ensure transparent and complete reporting. The PRISMA checklist for this study is attached in Appendix I.

6.3 EVALUATION PARAMETERS FOR THIS RESEARCH

The following parameters are evaluated below:

6.3.1 Correct problem formulation

6.3.2 Adequate methodology

6.3.3 'Fit for purpose' of the collection instruments

6.3.4 Validity of the data outcomes

6.3.1 Evaluation of the Methodology

The research problem formulation has been a real challenge, since it was difficult to define exactly what was the problem given the systemic perspectives this research has attempted to tackle. Eventually, the problem was formulated pragmatically:

'Despite the existence of a wealth of knowledge sharing initiatives and policies, such as Open Access mandates, it is still difficult for researchers to retrieve and access project knowledge even when publicly-funded.'

The following criteria constitute a guideline to evaluate whether the research problem, and to some extent the research questions, are well formulated (Cooper, 1984) and are tackled in various parts of the thesis, especially in Chapter 3:

1. The variables are defined, lexically and conceptually

2. The problem is stated so that the research designs and evidence needed can be easily identified and specified

3. The problem is placed in a meaningful theoretical, historical, and/or practical context, with direct references to relevant scholarly literature.

Methodological Fit

The notion of methodological fit was first established in organisational research (Bouchard 1976; Campbell et al. 1982; Lee et al. 1999). It can be said that 'methodological fit' is a type of heuristic evaluation, equivalent to what in systems practice is referred to as *fit for purpose*. Another important criterion that applies both to the methodology and to the outcomes of the research itself, is 'utility' (Miller, 1986), intended as 'is it useful in relation to its purpose?' In choosing and combining the different methods that make up this research, validated methodology design criteria (Creswell, 2006) have been adhered to, especially in relation to how the various research components relate to each other. These have been presented and justified in the methodology section. Additionally, ongoing critical appraisal was carried out throughout this research, and has shaped the iterative development throughout. This has resulted in various methods and instruments being scrapped, for example, a pilot survey carried out in 2010 never made it into the final study (although some of the questions were partially used in formulating the auditing templates), and the first auditing instrument, KAF, was not used for the final stages of the data collection, but rather modified and incorporated in OAM, as discussed in the relevant section. The iterative and cyclic nature of research, adopted in part in the development of this doctoral study, is illustrated in Fig.47.

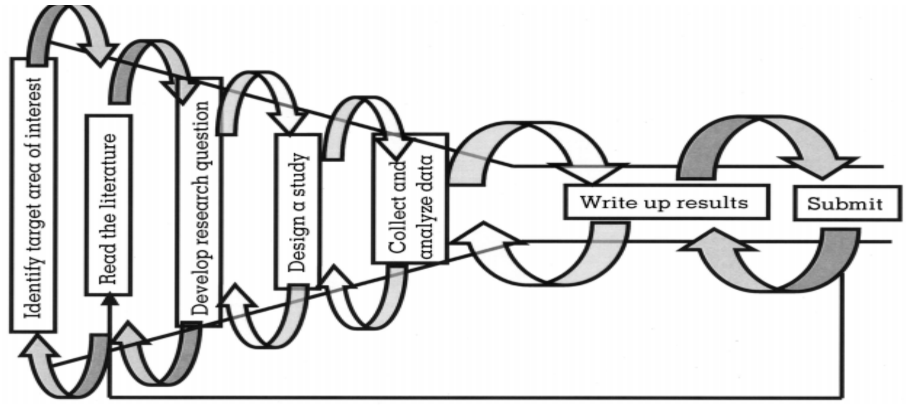


Fig. 47: Field Research as an Iterative, Cyclic Learning Journey (Edmonson McManus)

In summary:

Is the methodology a good fit for this research problem and questions?

Results obtained fall under two canonical categories of intended and unintended, and this research yields both, respectively identified in the following paragraphs. The methodology evolved dynamically from 'negative outcomes' of early research experiences and observations, so that these could also be analysed and contribute to advance the inquiry. There is no single way to address broad, complex, socio-technical challenges such as those associated with the knowledge sharing problem space, as tackled by this research. The justification for the choice of MMR is provided in the methodology section. The combination of methods adopted is the result of a balancing act between using the widest possible mix of techniques and paths, necessary to capture and cover as many possible research dimensions of the problem space, and leveraging pragmatism and parsimony, to ensure feasibility and manageability, and achieve at least some tangible results.

Taking a pragmatic approach to evaluation, a key criterion is whether a valid conclusion can be drawn from the research outcomes. In simple words, did the research yield results? Each research component yielded different results, RC1 qualitative, RC2 quantitative, from which conclusions could be drawn and recommendations formulated. Research methods were used in combination to validate and extend each other (Green et al.), as follows:

Triangulation: consists of more than just using multiple measurements of the same phenomenon; in addition to the use of diverse data, it involves combining different methods and theories, as well as perspectives of different investigators.
Four types of triangulation are used in research validation:

> "***Data triangulation*** *– the use of a variety of data sources and data sets in a study. Data may be both qualitative and quantitative, gathered by different methods or by the same method from different sources or at different times.*
> ***Investigator triangulation*** *– the use of several different researchers. Here the importance of partnership and teamwork is underlined as the way of bringing in different perspectives.*
> ***Theory triangulation*** *– the use of different theoretical viewpoints for determining competing hypotheses as well as for interpreting the single set of data.*
> ***Methodological triangulation*** *– the use of multiple methods to study a single problem or phenomenon. It may also include the use of the same method on different occasions and situations."* (Denzin, 1978)

One of the benefits of triangulation is that by blending sources of data, techniques and even researchers, the limitations and bias inherent to each respective method can be minimised (Jick, 1983).

The consistency of each finding, both in relation to RC1 and RC2, has been cross validated by triangulating methods, data and researchers, in the case of this study, auditors and key informants. For example, policy data collected via online searches and literature review, was validated by conversations and email exchanges with research councils' spokespersons, to confirm the findings of the analysis of literature, and then, where appropriate, by issuing corresponding FOI requests to access documentary evidence that could support the findings from literature and the conversations. Conclusions are drawn taking into account the entire spectrum of the validated outcomes.

Complementarity: Where the data collected using one method was partial or incomplete, other methods were adopted to validate and increment the findings, for example, in RC1, the Open Access policy impact evaluation was initially carried out by literature review and leveraging the findings of a previous study, however given that raw data was not available for further *ad-hoc* analysis and manipulation, existing findings and conclusions were integrated by doing ethnographic observations. The results of our observations carried out fully confirm and further validate, with empirical data, the conclusions of the previous study.

Development: Takes place when results from one method shape subsequent methods or steps in the research process, as reflected in the research design, and in the 'chain of evidence'. For example, after the outcome of the audits were systematically reviewed and analysed, the meaningful portion of the results that corresponded to good practice is going to be further investigated using qualitative methods.

6.3.2 Evaluation of the Instruments

The first criterion for evaluating KAF and OAM is 'Do they fulfil the need they were designed for?' In other words, do the collection instruments help to answer the questions the audits seek to answer? In both cases the answer is yes. A series of pilot audits carried out in KAF, as well as the full study carried out using OAM, prove that after audits have been carried out, it becomes possible to answer the key competence questions the audits were designed to answer, as explained in the relevant section. After some iterations and refinement the templates have become more robust, and the current working version is 'not so bad' (call it another heuristic). It should be noted that although the collection instruments did enable the satisfactory collection of the data required for the study, that is, they did fulfil their purpose, are still at prototype stage with limited functionality, nonetheless these have been made publicly available online. Occasionally the data entered in the online audit forms did not show up in the final sheet, and the audits had to be repeated, possibly due to some technical (browser synchronisation) issue, and beyond the scope and technical control of this research. As a result of this hiccup, an additional step was introduced in the auditing process that prescribed filling out a manual template a local copy of which auditors should store for safekeeping. This step is currently reflected in the KAF and OAM process diagrams. Resulting from various iterations, additional fields have been added to the template, and some of the initial questions which bunched up two different sets of values in the same field (integers and literals) were split into separate questions. These iterations are explained and documented in Chapter 4. The templates have fulfilled the initial purpose of providing empirical data for this study, however they are likely to continue to evolve and improve with usage.

6.3.2.1 Reliability of the auditing instruments

The reliability criteria are central to audits (RC2), as the aggregate dataset constitutes the quantitative aspect of this study. The reliability question can be formulated as follows:

Would two or more auditors following the same audit process and using the same audit template, come up with the same score?

KAF and OAM have been designed to allow for a degree of subjective judgement of different auditors, so a minimal level of variation of scores is acceptable. This variation has been estimated at a threshold level of 3: a difference of more than 3 points out of 16 (the total number of points allowed for each audit) between auditors, is established as an indicator of the score not being 'reliable'. The estimated threshold value for the difference in score between two auditors has been confirmed by a further test carried out and reported below. The whole idea of developing an auditing framework is to use a systematic instrument that should produce results, at least to some extent, consistent and replicable. The issue of replication is addressed in the section Evaluation of the Data Outcome.

6.3.2.2 Built-in verification mechanisms of the auditing process

Step 3 of the KAF and OAM auditing process consists of 'built-in verification'. The auditor writes to the grant holder asking to validate, or correct, the preliminary outcome of the audit. The reply of the PI is considered as 'inherent verification', so that every audit where the PI replied, either confirming or correcting the outcome of the audit, is considered valid and not validated further. However not all PI reply to correspondence, therefore not all audit outcomes can be validated using this built in validation mechanism.

An additional verification step was therefore required for those audits where the PI did not answer their emails. These were re-audited by an independent auditor, the results of the re-audits presented below.

6.3.2.3 Overall evaluation of the findings

The first quality criteria for the evaluation of outcomes are their validity and their reliability.

Validity is considered as the strength of our conclusions, inferences or propositions:

> *"Best available approximation to the truth or falsity of a given inference, proposition or conclusion."* (Cook & Campbell, 1979)

Reliability is the:

> *"Consistency of a measurement, or the degree to which an instrument measures the same way each time it is used under the same condition with the same subjects."* (Colosi & Dunifon, 2006)

How valid and consistent are the outcomes of the policy evaluation? (RC1)

The immediate goal of this research component is to understand what policies exist in support of 'knowledge sharing' in relation to the domain being studied (systems engineering research in the UK), taking into account the specific mandate that the corresponding funding research council is committed to, in the case of this study, EPSRC. This has been evaluated by analysing the documents (the policies and their references), and the documents associated with the policies, such as the UKRC policy the EPSRC policy remands to, as well as the DCC report that analyses and summarises the mandates by cross-validating the findings with:

a) correspondence with members of the funding councils

b) interviews with international policy experts.

The outcome of this research component is straightforward: the research council has an Open Access policy, which however does not grant access to 'technical knowledge artefacts' as typically adopted in knowledge sharing behaviours, because it does not specify what knowledge should be contained in the Open Access resources, or what technical artefacts should be adopted to make these Open Access resources 'findable' therefore 'accessible'. No further validation is required necessary for this outcome.

RC1 has also a wider secondary goal, that is, to understand the broader context and the 'Gestalt' in which the specific Open Access policy of a single research Council is developed. The findings obtained so far are to be considered 'summative', and constitute the basis for ongoing dialogues and further work.
The findings, while still being collected, are currently undergoing 'joint evaluation' (Beck & Buchanan-Smith, 2010) with selected groups of peers and experts:

> "*One of the frequently cited benefits of joint evaluations, and therefore one of the reasons for doing them, is that they are able to look at the big picture and evaluate collective action within the wider context. They can also tackle questions that cannot be addressed by any one agency, for example on coordination and coherence of the response: how agencies relate to each other and also to government authorities.*"

No conclusions have been drawn that rely solely on the findings of the comparative evaluation of the various initiatives, however the issues that emerged are addressed in the recommendations.

Limitation: The evaluation of the knowledge sharing and Open Access policy landscape, beyond the initial scope of finding out what policies apply to the population under study, has been deliberately exploratory, and without taking into account geographical or jurisdictional boundaries, as this was the only way to identify 'structural' gaps when the regulatory landscape is considered as a whole. However the evaluation of the outcomes is partial; still this portion of the study is considered open-ended, and ongoing. It is also relative to the limited expertise, worldview and sometimes, personal opinion and epistemological stance of the individuals consulted. No absolute validation criteria that this researcher is aware of can be applied for evaluating the outcomes of this research component at this stage. Therefore the study (the audits), as well as its outcomes, have been constrained on projects funded by a single council (EPSRC), which corresponds to the outcome of the first portion of RC1.

6.3.2.4 How reliable are the outcomes of the audits?

The question of 'reliability' can also be formulated as:

Would different auditors come up with the same score when auditing a project?

Although the OAM process and template are designed to guide and support a certain level of standardisation of the audits, some amount of subjective judgement is possible. It is the nature of heuristic evaluations to allow some degree of personal interpretation, which could result in slightly different outcomes. There are two possible ways in which the outcome of the audits can be different when performed by different auditors:

- error/omission: it is possible that auditors make a mistake, and do not find knowledge resources as the result of an oversight, despite these being available and findable on the web or
- assign a different value/score to the project being audited, due to a different interpretation of the findings.

Built-in validation and retests

As mentioned earlier in this section, the OAM process includes a built-in validation mechanism in step 3, which consists of contacting the PI asking to confirm or amend the outcome of the audit and/or the score. However, the majority of PIs did not reply when contacted via email, therefore all the projects where the PI did not reply to emails when asked to confirm, validate or correct the findings were re-audited independently, using a canonical test-retest approach. Initially, a random subset of 25 audits (approximately one quarter of the total) of the original global audit portfolio was first re-audited independently. Only approximately 50% of the audits were matched, with significant variations between scores assigned by two auditors in the other 50%. An allowance of two points for subjective factors was made, in that only variation between the two sets of scores greater than 3 points out of 16 was considered 'not valid'. Out of 25 'reliability tested' audits, one third were 100% consistent (same result) and approximately half were consistent within a 2 points variation (where the difference in score is <3. The other half resulted in variations between auditors of >3. See Table 21 for the figures.

PROJECT NUMBER		AUDITOR 1 SCORE	AUDITOR 2 SCORE	MATCH? (YES/NO)	IF N, DIFFERENCE IN INTEGERS
EP/E027024/1		14	0	N	14
EP/D063965/1		15	2	N	13
EP/E003257/1		0	9	N	9
EP/E018122/1		3	10	N	7
EP/E061982/1		0	6	N	6
EP/E057012/1		0	7	N	6
EP/D079365/1		1	6	N	5
EP/F031858/1		15	10	N	5
EP/D056268/1		6	1	N	5
EP/E05708X/1		15			4
EP/E02727X/1		0	3	N	3
EP/E044662/1		6	8	N	2
EP/F004117/1		6	4	N	2
EP/D50399X/1		2	0	N	2
EP/E057535/1		0	1	N	1
EP/F026781/1		0	1	N	1
EP/D035759/1		0	0	Y	0
EP/E057241/1		0	0	Y	0
EP/D032741/1		0	0	Y	0
EP/E057136/1		0	0	N	0
EP/D056268/1		6	6	Y	0
EP/C533186/1		0	0	Y	0
EP/E017304/1		0	0	Y	0

Table 21: Re-audited Projects Where the Difference in Score is >3

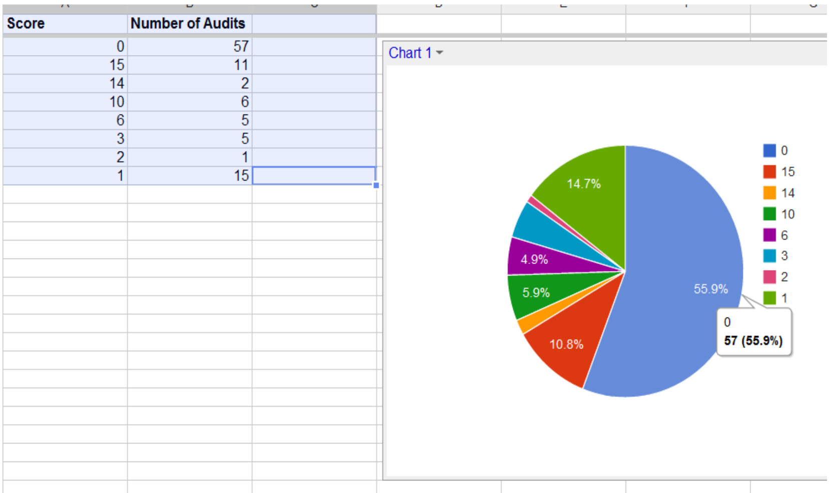
Fifty percent reliability was not considered satisfactory, however the variations could have been due to several factors, and the most obvious was the different search criteria between auditors. Therefore the project portfolio, where the PI has not validated or corrected the score, was completely re-audited, with the search algorithm defined for all auditors more narrowly, as explained in Section 5.2.2.2.

Some variation of the search results is expected between search engines, and within the same search engine in different dates, due to changes in the indexing. However, if none of the key terms produced results in terms of locating knowledge resources at any time, the study concludes that corresponding knowledge resources are not 'findable' via searches on the public internet, therefore they do not match the 'Open Access' definition, since accessibility of resources depends on their 'findability'(White, 2003).

To improve the reliability of the auditing instruments, a second round of audits was carried out to verify the outcome of the first round of audits, refine and harmonise the search criteria, the choice of which was initially left to each auditor (within a broad search logic left to the individual). The re-audits have been defined more tightly and used more consistently than in the first round. A more defined search algorithm however, helped to deliver more consistent results. Although there were some differences between the first and second sets of audits they seem to be evenly distributed, in the sense that marginal variations in scores between auditors, whether due to errors, omissions, different search criteria and/or different interpretation of the scoring system, were occurring similarly in both test and re-test, so that the final outcome of both audits is remarkably similar in statistical percentage, as shown in the tables and diagrams below.

Outcomes of the first round of audits

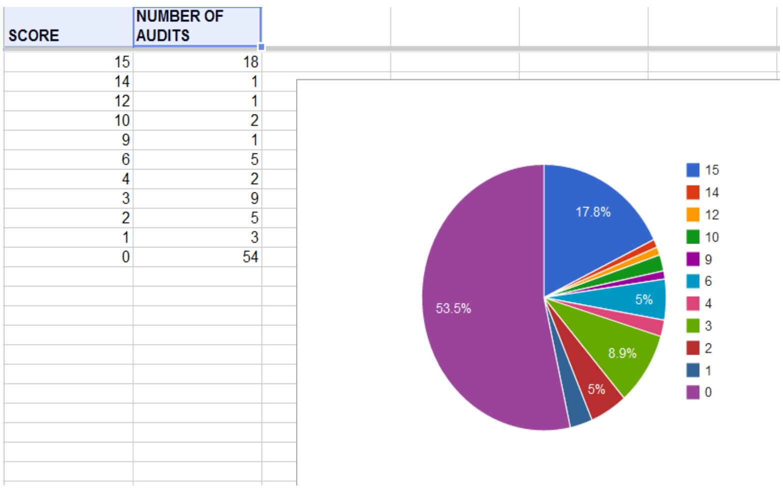
Just above 50% of the audited projects score zero (no online resources found), less than 15% score 1, which corresponds to minimal Open Access resources found, and just above 10% score 15, the highest value in OAM.



Score	Number of Audits
0	57
15	11
14	2
10	6
6	5
3	5
2	1
1	15

Fig. 48: First Round Audit Outcomes

Similar percentages correspond to the outcomes of the re-audited projects, although it should be noted that a higher percentage, just above 17% of projects, gets 15, the highest score.



SCORE	NUMBER OF AUDITS
15	18
14	1
12	1
10	2
9	1
6	5
4	2
3	9
2	5
1	3
0	54

Fig. 49: Re-Audit Outcomes

It is interesting to analyse differences in the scores a bit further, to understand if there are other factors, in addition to error and oversight or different search criteria, that can result in a different score. One of the hypotheses is that, after being contacted via email (in step 3 of the auditing process) some researchers would be prompted to actually publish some knowledge resources, so that maybe what was not found in the first round of audits, would be found in the second. An analysis of the comparative scores between the two sets of audits in Tables 22 and 23 shows that only a small number of projects re-audited resulted in a negative difference, where the second re-audit would yield a lesser score than the first. The majority of the projects seemed to have inherently increased the score the second time that they were audited. It was, however, impossible from this dataset to determine whether the increase in the score in the second round of audits was due to the auditor, or to the researchers actually making resources available and accessible after they received an email from OAM auditors.

A further test was therefore carried out:

> 29 projects from the same portfolio (EPSRC Systems Engineering) that ended in 2008 were audited using a set of pre-defined search criteria as presented above, but no email was sent to the PI, 15 of which were then re-audited independently using the same search criteria. The outcome of the re-audits confirms that the differences between the two tests is <3.

It can be concluded that where the search criteria are standardised, and where there is no bias from the shift in behaviour caused by the email, the difference between two auditors is <3 in 6 out of 15 re-audits, the rest presenting the same score, Fig.50:

EP/E057438/1	2	0	-2
EP/E057101/1	0	0	0
EP/C523865/1	2	0	-2
EP/F065647/1	0	0	0
EP/C541588/1	0	0	0
EP/E057047/1	6	3	-3
EP/E057152/1	0	0	0
EP/C512022/1	3	1	-2
EP/D03339X/1	0	3	3
GR/R97726/02	0	1	1
EP/F004370/1	0	0	0
EP/E01657X/2	0	0	0
EP/G049971/1	0	3	3
EP/E059872/1	0	2	2

Fig. 50: Variations in Scores between Auditors are <3 When Search Criteria are Constrained

External validation 1

Can the auditing instruments be used to audit *any* project, and not just systems engineering?

Yes, OAM has been used to successfully audit projects from NERC funding council:

Grant Number	Project Page on Funding Website	Project URL/webpage (other than funding body page) [ONE STAR]	Does the project publish proactive information [TWO STARS]	OA project resources [THREE STARS]	URI or unique identifier [FOUR STARS]	standard notation/formalism? [FIVE STARS]	raw data/code [SIX STARS]	Description/comment	document formats encountered when auditing	star rating	Total star rating and comments
NER/A/S/2003/00307	http://gotw.nerc.ac.uk/list_full.asp?pcode=NER%2FA%2FS%2F2003%2F00307	http://www.environmentalresearch.info/?ID=239949	no	no	no	no	no	A project page with only a little background information on a environmental project base website.	html, pdf	1	One - A project page with little or no information.
NER/T/S/2003/00729	http://gotw.nerc.ac.uk/list_full.asp?pcode=NER%2FT%2FS%2F2003%2F00729	http://www.environmentalresearch.info/?ID=1188725	Yes - Information on the project is provided along with contact information and grant information	no	no	no	no	A project page with proactive information	html, xls, xml	3	Three - A project page with proactive information
NE/B501947/1	http://gotw.nerc.ac.uk/list_full.asp?pcode=NE%2FB501947%2F1	http://www.environmentalresearch.info/?ID=1187025	Yes - Information on the project is provided along with contact information and yearly budgeting for the project.	no	no	no	no	A project page with proactive information	html, excel, xml	3	Three - A project page with proactive information
NER/T/S/2003/00487	http://gotw.nerc.ac.uk/list_full.asp?pcode=NER%2FT%2FS%2F2003%2F00487	http://creem2.st-andrews.ac.uk/documents/CreemAnn	Yes - Information about the running of the project and what has been carried out.	Yes	no	no	no	A project page that includes a good deal of proactive information and a list of references.	pdf	6	Six - A project page, proactive info and references listed
NER/T/S/2002/00166	http://gotw.nerc.ac.uk/list_full.asp?pcode=NER%2FT%2FS%2F2002%2F00166	Nothing found	no	no	no	no	no	Nothing found other than a link to a person's profile and a conference schedule.		0	Zero
NE/D008727/1	http://gotw.nerc.ac.uk/list_full.asp?pcode=NE%2FD008727%2F1	http://www.environmentalresearch.info/?ID=242526	no	no	no	no	no	A project page with only a little background information on a environmental project base website.	html, excel, xml	1	One - A project page with little or no information.
NER/A/S/2003/00348	http://gotw.nerc.ac.uk/list_full.asp?pcode=NER%2FA%2FS%2F2003%2F00348	http://www.environmentalresearch.info/?ID=1192106#01e2db98-a932-4351-a8e0-0d0fbc999300	Yes - Information on the project is provided along with contact information and yearly budgeting for the project.	no	no	no	no	A project page with proactive information	html, excel, xml	3	Three - A project page with proactive information
NE/B501004/1	http://gotw.nerc.ac.uk/list_full.asp?pcode=NE%2FB501004%2F1	http://www.environmentalresearch.info/?ID=1106045	Yes - Information on the project is provided along with contact information and yearly budgeting for the project.	no	no	no	no	A project page with proactive information	html, excel, xml	3	Three - A project page with proactive information

Fig. 51: External Validation, NERC Audits

External validation 2

What results would yield a series of audits carried out over projects selected via different criteria, say funded by the same council, but not in systems engineering?

Thirty EPSRC-funded Projects in Complex Systems were audited successfully, as per the data sheet in Fig. 52:

	D	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	AC	AD
1	Project Name an	Funding	Name an	Project P	Does t	If yes to the question above, enter URI here	Doe	If yes to t	Is the	If yes to t	Does	If yes to t	Does ea	If yes to t	Does the	If yes t	Does	If yes	Star r	What is t	Any comments?
2	EP/I016791/1	EPSRC	Professo	gow.epsr	no		no		no		no		no		no		no		0	0	
3	EP/H021779/1	EPSRC	Professo	gow.epsr	yes	http://erie.surrey.ac.uk/	yes	http://erie	yes		yes	http://erie	yes		no		no		0	0	
4	EP/E049370/1	EPSRC	Professo	gow.epsr	no		no		no		no		no		no		no		0	0	
5	EP/E049095/1	EPSRC	Lerner, P	gow.epsr	no		no		no		no		no		no		no		0	0	
6	EP/H022767/1	EPSRC	Kuna, Dr	gow.epsr	no		no		no		no		no		no		no		0	0	
7	EP/H00453X/1	EPSRC	Billings, I	gow.epsr	no		no		no		no		no		no		no		0	0	
8	EP/H05068X/1	EPSRC	Dr R San	gow.epsr	no		no		no		no		no		no		no		0	0	
9	EP/H010432/1	EPSRC	Professo	gow.epsr	no		no		no		no		no		no		no		0	0	
10	EP/G060169/1	EPSRC	Professo	gow.epsr	no		no		no		no		no		no		no		0	0	
11	EP/I005919/1	EPSRC	Mr JC Hil	gow.epsr	no		no		no		no		no		no		no		0	0	
12	EP/E062865/1	EPSRC	Professo	gow.epsr	no		no		yes	http://www	no		no		no		no		2	2	found a resource sep
13	EP/G05956X/1	EPSRC	Professo	gow.epsr	no		yes	http://eee	no		no		no		no		no		2	2	
14	EP/E049257/1	EPSRC	Beck, Pro	gow.epsr	no		no		no	http://www	no		no		no		no		2	2	resource was found s
15	EP/I019189/1	EPSRC	Professo	gow.epsr	yes	http://complexmatter.wordpress.com/	yes	http://com	no		no		no		no		no		3	3	
16	EP/E501621/1	EPSRC	Professo	gow.epsr	yes	www.ucl.ac.uk/secret/homepage	yes	http://www	no		no		no		no		no		3	3	
17	EP/F03685X/1	EPSRC	Dr G J Zi	gow.epsr	yes	http://www2.warwick.ac.uk/fac/sci/maths/research/events/	yes	http://www	no		no		no		no		no		3	3	
18	EP/H000380/1	EPSRC	Professo	gow.epsr	yes	http://www.uwe.ac.uk/research/heat/	yes	http://www	no		no		no		no		no		3	3	
19	EP/I019146/1	EPSRC	Dr r vitolo	gow.epsr	yes	http://emps.exeter.ac.uk/mathematics-computer-science/r	yes	http://emp	no		no		no		no		no		3	3	
20	EP/H031928/1	EPSRC	Dr RA Jol	gow.epsr	yes	http://www.maths.surrey.ac.uk/hosted-sites/research/topic	yes	http://www	no		no		no		no		no		3	3	
21	EP/G057737/1	EPSRC	Prof M Ba	gow.epsr	yes	http://simulacra.blogs.casa.ucl.ac.uk/2011/03/scale/	yes	http://sim	no		no		no		no		no		3	3	
22	EP/H02185X/1	EPSRC	Wilson, F	gow.epsr	yes	http://www.geog.ucl.ac.uk/about-the-department/news/nev	yes	http://www	no		no		no		no		no		3	3	
23	EP/I018492/1	EPSRC	Professo	gow.epsr	yes	http://focat.weebly.com/	yes	http://foca	yes		no		no		no		no		4	4	no results yet, but the
24	EP/G022682/1	EPSRC	31-Oct-1	gow.epsr	yes	http://www.cranfield.ac.uk/sas/risk/	yes	http://www	no	http://www	no		no		no		no		5	5	a resource was found
25	EP/E501311/1	EPSRC	Professo	gow.epsr	yes	http://www2.warwick.ac.uk/fac/cross_fac/comcom/dtcsite/	yes	http://www	yes	http://www	yes	http://www	no		no		no		6	6	the website had a pa
26	EP/E062083/1	EPSRC	Professo	gow.epsr	yes	http://socialsciences.exeter.ac.uk/politics/research/project	yes	http://soc	yes	www.ias.(	no		no		no		no		6	6	
27	EP/I01344X/1	EPSRC	Hall, Prof	gow.epsr	yes	http://www.itrc.org.uk/	yes	http://www	yes	http://www	no		no		no		no		6	6	
28	EP/H02171X/1	EPSRC	Fieldhou	http://gow	yes	http://sciddproject.wordpress.com/	yes	http://scid	yes	http://scid	yes	http://scid	yes	http://scid	no		no		15	15	
29	EP/E061850/1	EPSRC	Merali, M	gow.epsr	yes	http://www.gla.ac.uk/departments/emergenet/whatisemer	yes		yes	http://www	yes	http://www	yes	http://www	no		no		15	15	resource was found s
30	EP/G021163/1	EPSRC	MacKay, I	gow.epsr	yes	http://www2.warwick.ac.uk/fac/sci/maths/research/events/	yes	http://www	yes	http://www	yes	http://www	yes	http://www	no		no		15	15	
31	EP/E062075/1	EPSRC	Dr SJ Bel	gow.epsr	yes	http://emergingsustainability.org/content/about-emerging-	yes	http://emi	yes	http://emi	yes	http://emi	yes	http://emi	no		no		15	15	

Fig. 52: External Validation 2, NERC Audits – Different Field

Despite a different sample size, the number of projects audited in this test – 30 audits – is approximately a third of the main sample – 100 audits – the overall statistical distribution of the scores seems to be different in this segment, suggesting maybe there could be differences between systems engineering and other sectors (Fig. 53). This hypothesis would require further work to be explored, and is not within the immediate scope of this doctoral research.

	A	B
1	SCORE	NUMBER OF AUDITS
2	0	10
3	2	3
4	3	8
5	4	1
6	5	1
7	6	3
8	15	4

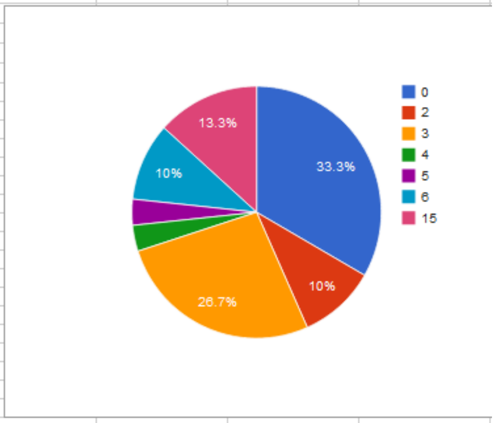


Fig. 53: 30 audits of EPSRC projects in complex systems show a different distribution than Systems Engineering

6.3.3 Evaluation of the Recommendations

The recommendations included in Chapter 7 at the time of writing are still being evaluated via feedback from the community of stakeholders – over 100 participants (mostly PIs and researchers) of the study were asked, via questionnaire, to evaluate the usefulness of the findings with a four point scale (very useful, useful, not so useful, useless) out of which approximately 15 replied, the majority of which suggesting 'very useful', Fig.54.

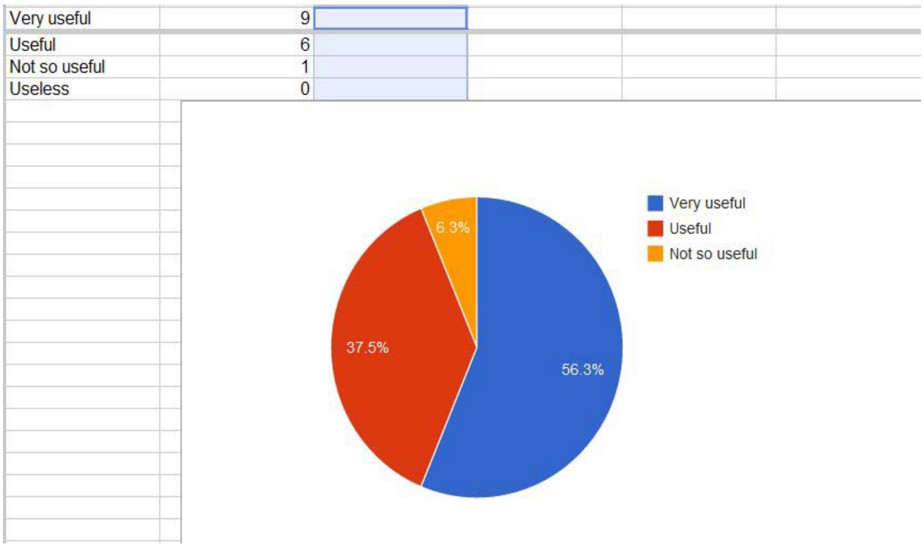


Fig. 54: Evaluation of Recommendations via Feedback

Further feedback is being elicited the evaluation of which remanded to future work.

6.4 OVERALL SELF-ASSESSMENT OF THE THESIS

In the first instance, the evaluation criteria for doctoral research were sourced from the appointed research advisors, and the following Examiners Report was supplied as the only reference for evaluation criteria for the University of Strathclyde[36]:

i) The thesis is a record of the candidate's original research or other advanced scholarship based on primary data sources or similar.
ii) The thesis makes an identifiable contribution to knowledge leading to original insight or understanding of a specific topic within the field of study.
iii) The oral examination was satisfactory.
iv) The literary style and presentation of the thesis are satisfactory.
v) The thesis is worthy of publication:
    (a) in part
    (b) in whole.
vi) The candidate possesses a good general knowledge and understanding of the field of study and has demonstrated an ability to conceptualise, design and implement an appropriate research strategy and programme.

No definition or agreed criteria are specified of what constitutes a 'contribution to knowledge' leaving the evaluation completely open to the subjective view point of the author or the examiner.

36 correspondence with supervisor, November 25th 2009

The overall lack of transparency and accountability and unspecified evaluation criteria are a known matter of concern for research scholars (Pearce, 2004):

> All institutions require and all students deserve academic procedures that are fair (on everyone), transparent (to everyone), and consistent (both in practice and outcome). This applies at the postgraduate research level as much as it does lower down the academic qualifications ladder. The current system for examining research degrees appears to work quite well most of the time in most universities across the UK, but sometimes it can go badly wrong and the process is not transparent. With little consistency in process within and between institutions, equity of treatment and assurance of quality cannot be guaranteed.

Fig. 55: Matters of Concern according to Pearce (2004)

Furthermore doctoral examination procedures are highly variable - they vary between institutions and between examiners, there are no clearly articulated common criteria or standards for the award of a PhD and the doctoral viva is usually a relatively private affair that takes place ‘behind closed doors (Pearce), a double edged sword. As a matter of due diligence, relevant terms of references for the evaluation of this doctoral research have been sourced and briefly reported below.

What is a Contribution to Knowledge

A contribution to knowledge *can* be formulated to answer the question 'what does research tell us that we did not know before?', but the range of answers can vary, from 'know what' to 'know how' to 'know how much'. Examples of diverse variety of doctoral contribution are provided in the list below (Rugg & Petre, 2004):

> *"- **Re-contextualization** of an existing technique, theory or model (applying a technique in a new context, testing a theory in a new setting, showing the applicability of a model to a new situation): showing it works – or that it doesn't – and why*

*- **Corroboration and elaboration** of an existing model (e.g. evaluating the effects of a change of condition; experimental assessment of one aspect of a model)*

*- **Falsification or contradiction** of an existing model, or part of one*

*- **Drawing together two or more existing ideas** and showing that the combination reveals something new and useful*

*- **Demonstration of a concept:** showing that something is feasible and has utility (or showing that something is infeasible and explaining why it fails)*

*- **Implementation of theoretical principle:** showing how it can be applied in practice; making concrete someone else's idea, and hence showing how it works in practice and what its limitations are*

*- **Codification of the 'obvious':** providing evidence about what 'everyone knows' (possibly providing evidence that received wisdom is incorrect)*

*- **Empirically-based characterization of a phenomenon of interest** (e.g. detailed, critical, analytic account of the evolution of an idea; detailed analytic characterization of a crucial case study or a novel chemical compound, or a new planet)*

*- **Well-founded critique of existing theory or evidence** (e.g. correlating the results of a number of existing studies to show patterns, omissions or etc.)."*

(Rugg & Petre, 2004)

The novel contributions to knowledge resulting from this research, and how they fit into the definition of 'what is a contribution' described above are presented in the next, and final part, of this thesis.

Finally, the quality model for this thesis has been designed taking into account explicit and structured terms of reference, as per the example provided by the institute of Education, University of London[37] and summarised below:

"PRESENTATION AND CLARITY"	"- The reader should be able to read the text without difficulty. - The text should be clear and 'tell a story'. -The submission should be 'user friendly'. The reader should be able to find his or her way around the submission, locating tables and Figs., and being able to cross reference with ease. A numbering system for chapters, sections, and, sometimes, paragraphs can be very helpful. - The style should be economic without unnecessary duplication or repetition. - The bibliography and/or reference list should be complete and accurate. - It should be possible to gain easy access to tables and figures relating to particular passages in the text, and to examine both data and commentary without effort. - The submission should be no longer than necessary."
"INTEGRATION AND COHERENCE"	"There should be logical and rational links between the component parts of the thesis. In some cases coherence will be achieved by a series of empirical studies or analyses which build one upon the other. In other words, there will be an intellectual wholeness to the submission. "
"CONTRIBUTION TO KNOWLEDGE "	"A submission for a PhD should be approximately equivalent in quantity and quality to at least two articles of a standard acceptable to a fully refereed journal. Where candidates have already had portions of their doctoral work accepted for publication, this is prima facie evidence of an adequate standard. Alternatively, the submission should be substantial enough to be able to form the basis of a book or research monograph which could meet the standards of an established academic publisher operating a system of critical peer review for book proposals and drafts."
"ORIGINALITY AND CREATIVITY"	"The research and the written submission should be the candidate's own work. However, the degree of independence shown may vary according to the research topic, since in some instances students will be working as part of a larger team, while in other instances they will be completely on their own. A candidate should show an appropriate level of independent working."
"REVIEW OF RELEVANT LITERATURE"	"Candidates should demonstrate that they have detailed knowledge of original sources, have a thorough knowledge of the field, and understand the main theoretical and methodological issues. There should not be undue dependence on secondary sources. The literature review should be more than a catalogue of the literature. It should contain a critical, analytic approach, with an understanding of sources of error and differences of opinion. The literature review should not be over-inclusive. It should not cover non-essential literature nor contain irrelevant digressions. Studies recognised as key or seminal [sic] in the field of enquiry should not be ignored. However, a student should not be penalized for omitting to review research published immediately before the thesis was submitted. A good literature review will be succinct, penetrating and challenging to read"

Table 22: Quality Criteria for Doctoral Dissertation in the UK

6.5 CONCLUSION

This section presents key considerations in relation to the quality planning and assessment of mixed method research in general, and detailed expositions of how various aspects of this thesis have been evaluated, comprehension of figures and tables reproducing the validation tests where appropriate.

[37] Education, University of London, http://www.ioe.ac.uk/doctoralschool/info-viva.htm

Each central aspect of the research and the thesis has been evaluated, and where necessary validated, taking into account standard, as well as novel approaches, as in the cases of PRISMA. Since the widest possible combination of research methods was adopted, similarly, the widest possible choice of evaluation methods has been applied, to ensure and demonstrate, above all 'credibility' and 'trustworthiness' of this research. In particular, priority has been placed in the documentation and transparency of key decisions made along every step of the research and dissertation writing process.

"Knowledge Speaks, But Wisdom Listens"

attributed to Jimi Hendrix

CHAPTER 7

CONCLUSION, RECOMMENDATIONS, WORK AHEAD

OVERVIEW OF THE SECTION

In this final section of the thesis, the research questions and hypotheses are briefly reconsidered, the outcomes and findings summarised, and a series of socio-technical recommendations issued. The main contributions to knowledge are identified, together with an overall discussion, followed by pointers to future work and concluding with final remarks.

7.1 SUMMARY OF OUTCOMES

This research addresses the challenges of accessing publicly-funded scholarly knowledge, in the context of a pragmatic, socio-technical approach.

Ongoing literature review referenced throughout this thesis provides background knowledge from a multi-disciplinary perspective in relation to the themes under study, together with a critical review of the same. The main outcomes of the research can be summarised as:

- A critical analysis of a recently published narrative review of 'future directions for knowledge sharing research', which identifies and addresses an important knowledge gap in the literature (Chapter 3).
- Initiatives, policies and legislation in support of knowledge sharing, such as the Open Access policies of public research funding councils in the UK, are explored and evaluated in relation to their usefulness to the knowledge seeker, and inconsistencies and some contradictions of terms are identified (Chapter 5, Research Component 1).
- Novel collection instruments have been developed, including a set of heuristics for the evaluation and qualitative benchmarking of knowledge sharing patterns of publicly-funded projects in research, the rationale and justification and technical outline of the collection instruments and supporting heuristics is provided in Chapter 4.
- A comprehensive portfolio of projects in the domain under study is audited, and re-audited, and the results are presented (Chapter 5, Research Component 2).
- An evaluation of the research methodology and findings is provided in Chapter 6, and the set of overall recommendations based on the findings is presented here, in Chapter 7.

The resulting outcomes and recommendations in combination constitute the foundation for the 'systemic' approach proposed with this research.

7.1.1 The Research Questions Revisited

The initial observations described in Chapter 1 provided the overall motivation and direction for this inquiry, however the formulation of the research question and corresponding hypotheses are largely the result of a working compromise, and have been designed deliberately to be general to allow for the widest possible modes of exploration of a seemingly intractable socio-technical problem that could not be precisely understood and defined a priori. In this section therefore questions and hypothesis are summarised and then loosely 'revisited' and reformulated to frame them in the context of the work done and of the outcomes achieved. In the paragraphs that follow, the original question and hypotheses are reproduced, with the addition of a brief commentary.

ORIGINAL Question and Hypothesis 1

Q1. How can the gap between T and P be identified?

H1. By gathering and comparing evidence of T and P.

An instrument to gather evidence of policies and practices has been developed and validated, as presented in Chapter 4. This hypothesis is TRUE (the gap between theory and practice can be identified via gathering evidence).

REVISITED:

The main purpose of the first research component (RC1) is to characterize the expectation that knowledge resources should be findable in the first place. From a different angle, the motivating research question could be formulated as:

What motivates the assumption that knowledge derived from research outputs should be accessible?

The 'access to scholarly research outputs' movement has been growing steadily since the Budapest Declaration, as introduced in Chapter 3. However there is no single policy, law or initiative that can legitimately guarantee that *'knowledge from publicly funded research should be publicly accessible on the internet*'. An exploratory evaluation of the knowledge sharing policy field was therefore undertaken to understand how the boundaries of the related initiatives intersect. Given the variety and heterogeneity, and varying degrees of effectiveness of the measures that directly or indirectly are devised to support knowledge sharing, a combination of literature review, analysis of disparate sources (including consultation with employees of research councils exchanges with peers and experts) were adopted. The primary focus of the exploration became 'open access policies' as these are proposed by the 'open access movement' as the core motivation for open access publishing of scholarly outputs.

<u>ORIGINAL Question and Hypothesis 2</u>

Q2. Is there a gap between T and P?

H2. Despite knowledge sharing policies, such as Open Access mandates being adopted by research funding councils, knowledge seekers cannot locate Open Access knowledge resources via online web searches.

Based on the portfolio of audited projects, reported in Chapter 5 and validated in Chapter 6, which relates to a specific sector in a specific country, the hypothesis is TRUE (yes, despite the existence of Open Access policies, based on the methodology and evidence gathering instruments developed, in a large number of cases, it was not possible for auditors emulating knowledge seekers' behaviours to locate papers or other knowledge resource via online searches). Other domains have been tentatively explored for hypothesis validation purposes, but the work carried out in this study at this stage relates to the UK only.

REVISITED:

Given the different types of knowledge sharing policies in existence, and in particular the open access policies adopted by the research funding councils, are these adequate to ensure access to research?

Based on the evidence gathered both in the policy evaluation RC1 and the outcome of project audits RC2, this research demonstrates that despite the unequivocal existence of open access policies, and despite literature stating the adherence to the policy is 'monitored' by the issuing research council (Section 5.1.1.4), it can be said that knowledge generated by research outputs is accessible only in a small percentage of the cases.

ORIGINAL Question and Hypothesis 3

Q3. How can the gap between T and P be reduced?

H3. By devising integrated socio-technical measures and interventions.

A series of recommendations is formulated and reported in this chapter. These have been tentatively evaluated, merely for completeness, by feedback provided by approximately 15% of the study participants as 'useful' and 'very useful', therefore the hypothesis that the integrated socio-technical measures are useful to reduce the gap between theory and practice is within the limitations afforded by the resources available for this study, as TRUE. Further research to test the validity of the proposed measures over a more extensive field is however advisable.

REVISITED:

The third and final research question and hypothesis was developed *a posteriori*. It can be rephrased as:

What can be done to make research outputs more accessible to knowledge seekers?

An initial set of socio-technical measures, intending to suggest a direction for future work, are developed from literature and working knowledge of good practices in relevant fields, in the hope of anticipating the 'so what' question that may arise from reading the analysis of the main findings.

7.1.2 Methodological Implications

A mixed method research design has proven useful to tackle an otherwise 'intractable' social mess. The particular *melange* of techniques adopted here integrates qualitative and fuzzy techniques of inquiry, such as heuristics to pragmatically capture at least in part some of the incommensurable aspects of the problem space (What to measure? How?), with a meta-analysis approach that aims to gather insights into outcomes from the different techniques, such as carrying out classical cross referencing and statistical analysis of the data from the audits. The positivist/quantitative turn was dictated by the desire to obtain at last some hard fact to include in the findings that could be validated, and to some extent replicated. Particularly interesting and useful is the inclusion of FOI requests as triangulation method to validate information obtained with other techniques (email exchanges).

The FOI Act is a piece of legislation enforced by the Information Commissioner Office which requires public authorities to reply to questions in writing, disclosing relevant evidence and documentation where available. As this research establishes, much information in relation to the implementation of these policies, for example to answer the question *how are the policies monitored*? is answered subjectively by different individuals, who do not provide any evidence to support their answers.

The FOI requests were the only method of inquiry that could provide conclusive reliable outcomes, and the legal right to access evidence to corroborate the replies provided by employees during exchanges. The immediate implication for this doctoral research is that the research outcomes corresponding to FOI Requests, where other outcomes were not reliable, have contributed to construct the validity of the study. In broader terms, FOI Requests, although a relatively novel instrument of inquiry, can be very valuable to research, and sometime the only way to obtain results, therefore their usage should be encouraged especially in social research where reliable and consistent information is otherwise difficult to obtain. Additionally, ethnography is referenced, merely for completeness, in Section 2.1.1.5 as an ancillary technique part of the wider mixed method approach adopted in some 'background 'portion of the research to attempt answering notionally exploratory questions such as:

- *how much do academic and practitioners, and in particular systems engineers know about open access?*
- *What is their level of awareness of the existence of such policies?*

The questions could not be answered from literature, considering that the only study (Swan, 2008)[38] that tackled a similar set of issues does not disclose the academic field/profile of its participants, and the dataset upon which the conclusions are drawn was not available for be further querying and manipulation, as discussed in Section 5.1.2. After a series of exploratory pilot questionnaires had produced disappointing outcomes (outcomes that could not be conclusively validated or that would require significantly more resources - time and funding - to be adopted reliably), it was decided that more qualitative approach could be

[38] A. Swan (2008)Key concerns within the scholarly communications process
Report to the JISC Scholarly Communications Group
http://www.jisc.ac.uk/whatwedo/topics/opentechnologies/openaccess/reports/keyconcerns

beneficial. The opportunity arose of leveraging ongoing participation in networking events, academic conference and professional association gatherings, listed in the corresponding appendix, as a participant and observer (note that observations also take place via ongoing participation in related online communities and projects).

7.1.3 Summary of the Findings

The thesis attempts to provide a reasonably structured account of what has been a intensively concurrent research design, with many activities taking place simultaneously. The table below attempts to provide a structured overview of the overall findings, some of which have come up in the corresponding narratives, for example, such as the findings acquired from literature review, and matching them to the corresponding recommendations, which are then discussed in section 7.2. The structured Table 23 presents mostly a linear view of a chaotic, non linear problem and solution space.

RESEARCH COMPONENT	FINDING	RECOMMENDATON
Literature review	1. Policies not included in future knowledge sharing research, see p94 2. Challenges for Doctoral Research in SE do not include knowledge access/reuse, see p28	R1. Include Policy in future of knowledge sharing research R2. Include knowledge access/reuse to doctoral challenges for system engineering/system sciences
RC1	Fragmentation Underspecification Contradictions	R3. Harmonize/Integrate R4. Specify Policies R5. Fact checking, Linking/Using URIs
RC2	Large numbers (over 50% in the case of this study) of publicly funded projects do not publish knowledge which is shared/accessible on the web, despite the existence of open access policies Research outputs do not adopted structured , explicit knowledge schemas	R6 Monitor evidence R7 Legislate (make the policies mandatory by law) R8 Adopt shared structured knowledge schemas for research outputs

RESEARCH COMPONENT	FINDING	RECOMMENDATON
Overall	KS policies exist, but a) their adherence is not monitored b) they do not support adequately the access requirements of k seekers knowledge seekers Lack of OA and KS awareness in the Engineering Research	R9 General Recommendations (Inform, Incentivate, Foster Culture of Knowledge Sharing)

Table 23: Findings to Recommendations

7.2 RECOMMENDATIONS

In this section some recommendations are suggested as an 'integrated set', devised to address the systemic issues that become evident through the analysis of the findings. The motivation and justification for these individual recommendations is found in literature and good practice, as discussed and referenced in detail throughout the thesis. The proposed socio-technical approach, emphasizes the co-dependency of these measures, as a good fit to address the entanglement of problem under investigation. For example, to monitor the adherence to a knowledge sharing/open access policy sustainably (with contained effort and limited burden to the tax payer), as prescribed in R6, it is necessary to adopt appropriate artefacts (use URIs to share knowledge), as described in R8 which in turn can be achieved by an appropriately detailed level of policy specification R4, which in turn needs pointing to suitable artefact adoption and that can only be mandated by legislation in R7, and so on. Some of the recommendations are broad and generalisable and applicable beyond the domain of publicly-funded research, listed at the bottom as General Recommendations.

7.2.1. Include Policy in Future Knowledge Sharing Research (R1)

Non empirical findings in this study demonstrate that knowledge, and especially technical knowledge, is best shared through the adoption of explicitly codified artefacts, yet the empirical findings of this research (RC1) show that the adoption of such artefacts is limited, at least in the population segment analyzed in this study. Literature analysis shows that policy is not included in future research directions for knowledge sharing. This recommendation to include policy as a key socio-technical measure is therefore aimed at filling that gap (Section 3.3).

7.2.2. Include Access to Research Outputs in Doctoral Research Challenges (R2)

The difficulty in accessing scholarly research outputs in systems engineering in the UK, as narrated in the initial chapters, also confirmed by the empirical findings in RC2, are a challenge for new doctoral students in the discipline. Addressing these challenges, as discussed throughout the thesis, is a complex matter for which no single 'silver bullet' solution exist. The recommendation is, at a minimum, to acknowledge the difficulties and barriers to accessing the state of the art in systems engineering research for current and future doctoral researchers and students, make them known to all interested parties and address them with suitable strategies accordingly.

7.2.3. Harmonise (R3)

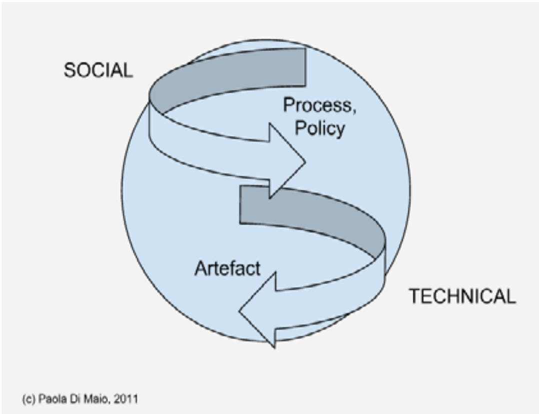
The fragmented and heterogeneous state of policies and legislation supporting knowledge sharing of scholarly outputs, can be confusing, and even lead to contradictory and conflicting initiatives. Although the existence of multiple policies is justified by the diversity of situations and cases, and may serve a purpose, their fragmentation and mutually conflicting aims can result in the uneven, erratic operationalisation of the policies.

The harmonisation and possibly alignment and integration between disparate policies and initiatives are desirable for at least two reasons: first, to avoid duplication, redundancy and possible conflicts arising from multiple policies, and secondly, to reduce the administrative cost of multiple policies development, implementation and management, and the costs of conflict resolution when they clash. Open Access policies should continue to encourage compliance from the bottom up – roughly corresponding to the self-archiving, also known as the 'green option' in Open Access jargon – but also encourage funding bodies and regulators to implement their policy via effective regulatory measures (mandates) that can be monitored.

7.2.3.1 Integration

At least three levels of integration are recommended, aimed at tackling three different aspects of the policy fragmentation problem.

- **Policy-artefact integration:** None of the policies and initiatives examined in the course of this research mandates specifies the adoption of suitable technical artefacts such as knowledge representation and formalisms (briefly discussed in R8) which emerge from research in the corresponding technical fields, such as knowledge engineering and web science, resulting in what this research identifies as a 'pragmatic gap'. It is important to establish a firm 'correspondence' between the policy and the mandates on the one hand, which can be called the social and organisational aspects of knowledge sharing, and the adoption of the knowledge sharing artefacts, conventions and standards, that can be defined as the technical aspects, because the two are facets of the 'same coin', as shown diagrammatically in the illustration below.



⅄ Fig. 56: The DAO of Socio-Technical Knowledge Sharing

- **Data, information and knowledge policies integration:** At the moment, different policies exist, some addressing knowledge, some data, and some information sharing. These should be aligned and where possible integrated, for the reasons mentioned above.
- **Geographic-jurisdiction integration:** Some of the terminology and concepts adopted in different knowledge sharing policies, such as, for example, the PSI directive in the EU and in the UK, present some discrepancies, as pointed out in Chapter 4 of this research. These should be, where possible, harmonised, to reduce the conflicts that contribute to the fragmentation of the policy landscape. Furthermore, there should be some integration and alignment between the 'institutional' policies, and the organisational policies.

Semantic integration is also recommended, as it emerges from the combination of other recommendations, such as R4 and R8.

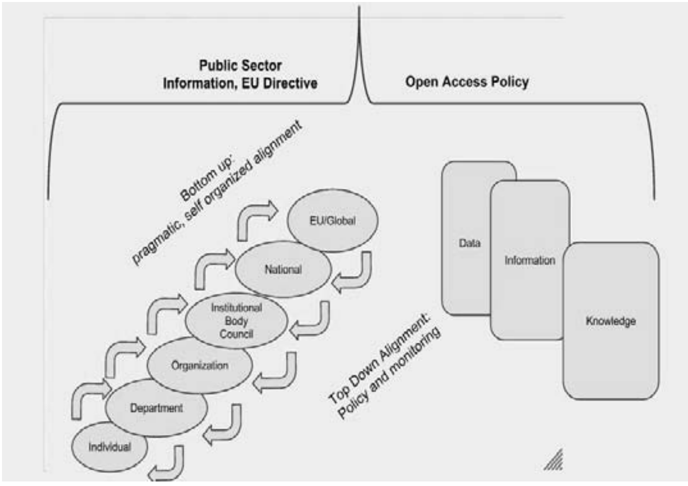


Fig. 57: Recommended Intervention: Integration and Alignment of the Fragmented Policies

7.2.4 Specify policies (R4)

It can be said that policies are rules governing the choices in behaviour (Sloman, 1994), derived from business goals, service level agreements or trust relationships within or between organisations (Damianou et al, 2002). In the context of this research, policies refer primarily to the 'mandates' of funding bodies that recommend the adoption of knowledge sharing behaviour, which are directly or indirectly related to a plethora of other policies, including international ones, as discussed in RC1. As such policies discussed here pertain to the domain of 'public policy' (Smith, 2003). The topic of publicly policy specification deserves a full in depth exploration of its own, and it is remanded to future work. Let it be recommended here that knowledge sharing policy specification should aim to define and where applicable align:

- the actors, or stakeholders and their role in the policy lifecycle (Bryson, 2003)

- the policy dimensions (Potucek, 2003)
- process, or workflow, that should be supported by the policy throughout its lifecycle. (Dais et al, 2008)
- requirements for adherence (a policy is adhered to if X requirements are met) (Abdullah et al, 2009)
- technical specifications, including artefacts and processes, required to ensure the policy is adopted and monitored as per the requirement (EU Public Consultation)[39]

If the monitoring of policies is to take place via the implementation of a web service[40] then the corresponding standards and languages that apply should also be specified.

7.2.5 Fact Check and Use Linked Data and URIs to Provide Evidence (R5)

Much is said in philosophy and epistemology of knowledge as 'just belief' (Gettier, 1963), and knowledge as truth (Russell, 1913) however 'scientific knowledge' rests, above all, on facts, whereby science itself is about verifiability and reproducibility (Rott, 2002). This research is developed in the context of an engineering discipline, in particular systems, web and knowledge engineering, whereby engineering is intended as 'the practical application of science to commerce or industry' (Fox, 1984). In related sections of this thesis the case for the use of URIs as appropriate formalisms for documentary evidence is made (Chapter 3). A recommendation is therefore derived that no claim or assertion of any fact, including policy adherence, should be considered reliable without evidence, necessary to establish at least some level of 'truth', as opposed to knowledge being based on beliefs or assumptions.

---

39 http://ec.europa.eu/enterprise/sectors/ict/files/overview_report_results_consultation_en.pdf
40 Web Service Architecture http://www.w3.org/TR/ws-arch/

When sharing knowledge on the web, the mechanism to provide verifiable evidence is to use hyperlinks to corresponding documents. In related work, the linked data model is explored as a possible formalisation for fact checking.[41] The architecture of the collection instruments developed for the study, KAF and OAM, uses URIs of knowledge resources as documentary evidence of the existence of knowledge resources corresponding to Open Access.

7.2.6 Monitor (R6)

The existence of one or more knowledge sharing policies, such as the Open Access declarations published on the website of the funders, does not mean, as the outcomes of the audits reported in Chapter 4 of this thesis show, that these policies are adhered to. Monitoring is an essential part of policy management. Ideally, a monitoring process should be an integral part of the funding process (no accessible schemas and URLs = no funding?). In related work a web service to automate the monitoring of policy compliance is explored as part of the Open Access Monitor future development plan.

7.2.7 Legislate (R7)

At the moment, the provisions for commercial knowledge transfer are entered into contractually, whereby contracts are legally binding instruments enforced by a firmly established piece of legislation, such as Contract Law. By contrast, Open Access policies, despite being named 'mandates' at the time of writing were not mandatory at all, and are at best still mostly 'guidelines', carrying no legally binding obligations. The relation between Open Access policies and knowledge transfer agreements is strongly asymmetrical in the law, and favours the latter.

---

[41] Provenance and Linked Data Workshop, e-Science Institute, University of Edinburgh 2011

Additionally, clearer guidelines should be specified in Open Access mandates as to what level of data, information and knowledge should be made freely accessible, and which levels can be protected by patents and copyright to allow research outputs to benefit from commercialisation opportunities and economic gain via Knowledge Transfer agreements.

7.2.8 Adopt Shared Formalisms (R8)

To achieve optimal knowledge sharing potential of codified knowledge resources, such as technical knowledge, it is necessary to adopt appropriate conventions, formalisms and artefacts. This is sometimes referred to as 'semantic and pragmatic' integration'. Knowledge artefacts must take into account the knowledge structures used in the corresponding knowledge domain, an example of an explicit representation of a knowledge structure in systems engineering is derived from analysing corresponding domain knowledge sources, and provided below. Some of these technical conventions are well established, such as:

- use a publicly-accessible online page or website
- share resources using URIs
- adopt shared/standardised notations and formalisms

These have been encoded in the set of heuristics proposed. However as of to date, no single shared knowledge schema exists that researchers and practitioners in systems science can adopt when trying to make their research outputs more useful and more easily accessible and reusable to others. In a separate paper (Di Maio, 2011) the development of a shared vocabulary for systems engineering using automated text extraction is reported. The analysis of relevant portions of the systems engineering body of knowledge results in a sample 'reference model', provided as an example of 'shared knowledge model'.

LIFECYCLE PHASE	KNOWLEDGE ASSET *document, specification	FORMAT	NOTATION/FORMALISM	SHARING MECHANISM
analysis	requirements specification	narrative structured text	natural language, pseudo-code	
design	system diagram	diagram	ER, DF, UML	
development	system specification	narrative structured text	Natural language pseudocode	image word document spreadsheet pdf html xml rdf owl other
installation	operating manual user guide	narrative diagrams	Natural language graphics	
testing	test plan	structured text	natural language charts	
acceptance	contract	narrative	natural language	
support	user feedback tickets	narrative	natural language	
	feedback			

Table 24: Reference Model of Knowledge Sharing in SE, Di Maio

Table 24 is an example of how a set of 'typical' system development lifecycle phases can correspond to a set of knowledge artefacts, logically articulated, represented and shared using appropriate formalisms, notation and file formats. Similar domain dependent knowledge reference schemas can be developed and adopted in other domains. The overall general recommendation (acquired from web science) is to model or map knowledge resources leveraging shared vocabularies and schemas, and make the resources accessible via URIs on the web. This particular recommendation has been an underlying mantra in web science in the last twenty years, however as shown by the empirical findings of this study, in the majority of cases it is still largely disregarded by researchers, practitioners as well as by policy makers. Furthermore, domain-specific knowledge reference models, as per the example provided above, can mitigate at least in part the lack of more sophisticated shared codification standards, such as formal ontologies, which are arguably more ambitious to achieve.

7.2.9 General Recommendations (R9)

Institutions, as bureaucratic organisations, tend to be 'passive', and to follow directions issued from 'the top' by their governing bodies. For a policy carrying strategic implications for the advancement of science at global and national level, such as a policy for Open Access to scholarly publications, it is necessary for everyone in the research supply chain, starting from the researchers as the main producers of new knowledge, to wholeheartedly embrace it. What good is a policy emitted by a funding body if institutions do not adhere to it? Institutions have primary responsibilities toward the public at large, as well as toward the public funding councils and the research community. General recommendations for research institutions and funding bodies are:

a) **Inform:** provide regular training about knowledge sharing (policies, technologies) and where necessary technical support for researchers. issue guidelines for researchers, including recommending the adoption of existing knowledge schemas and knowledge sharing artefacts and good practices, where these are available, and stimulate the innovative development of new ones.

b) **Embrace a culture of knowledge sharing:** although this may imply a disruptive overhaul of preconstituted knowledge and 'power' hierarchies.

c) **Incentivate:** additional benefits, including funding, should be provided for researchers who adhere to good knowledge sharing practices, as they truly serve advancement of science goals. Budgetary considerations play a role in the effective implementation of knowledge sharing policies, such as how Open Access is monitored. Currently, based on the outcome of various FOI requests in this study, it emerges that monitoring is carried out manually and notionally, mostly via one or two individuals at each funding council, using statistical sampling. Without the appropriate and systematic use of intelligent web-based technologies, the costs of monitoring policy implementation could exceed the benefits. Simple policy monitoring processes should be mandated and integrated in the funding procedures and automated via the adoption of simple artefacts and techniques, such as for example, 'linked data'.

Additionally, community involvement and participation in the monitoring process should be encouraged. Crowd-sourcing should be considered, as demonstrated in principle via the public version of Open Access Monitor.

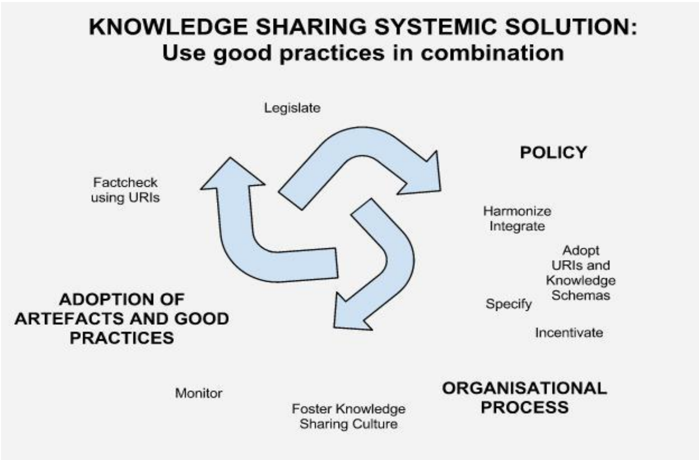


Fig. 58: Systemic Solution, Recommendations in Combination

7.3 CONTRIBUTION TO KNOWLEDGE

This research started with questions that could not be answered from existing scholarly literature, and resulted in evidence and empirical findings that directly answer the research questions, from which it is possible to confirm that the main hypotheses are true. To gather evidence, a novel methodology, a set of custom collection instruments and heuristic indicators, has been developed. The theoretical contribution includes: the identification of a socio-technical research space where policies and procedures with the adoption of web-based artefacts and techniques, illustrated in Fig.3, are more tightly coupled, and the synthesis of conceptual and systemic knowledge resources, illustrated in Fig.16.

The overall outcomes of the research can be summarised as:

1) **Empirical data as Evidence.** Large scale studies have been published, that report the growth of Open Access publications (Björk et al. 2010; Laakso et al. 2011) consisting of statistical analyses of large datasets. No systematic analysis has been carried out before that surveys a particular sector, given the lack of suitable evaluation methodology and instruments. The results obtained in this study show that from a knowledge seeker point of view, for approximately half of publicly-funded research projects, no papers or other knowledge resources can be located and accessed via searches on the open web, despite the existence of Open Access policies of the respective funding councils. This empirical data, despite its limitations, should contribute to inform policy makers of the effectiveness of the existing policies, and motivate and justify further evidence-based research.

2) A **critical analysis** of existing scholarly literature identifies and addresses two key knowledge gaps.

3) A **novel Knowledge Auditing methodology** and prototype implementation, to support auditing of open web-based resources using URIs and FOI to gather and validate evidence, which allows to heuristically evaluate and benchmark knowledge sharing behaviours against the funders mandates, as well as enabling parametric comparisons and ranking. It is noted that although many knowledge auditing methodologies exist, none evaluated in this study support web based knowledge audits with a tool that uses URIs.

4) **A novel set of heuristics** to evaluate the degree of knowledge sharing (or the degree of openness) that serves as benchmark in the auditing process.

5) **A set of recommendations**.

This research, in various stages, has been presented at international doctoral consortium (EUSEC, 2010) as well as other national and international workshops, symposia and conferences. Portions of this research have been published in fully referenced scholarly publications.

7.4 LIMITATIONS

Making a (modest) contribution to such a wide and open field of study such as 'knowledge sharing' has been possible only by constraining the scope of work to match a limited time-frame for its completion, which resulted inevitably in limitations, the most obvious of which are:

- further in depth analysis of topic areas related to recommendations (such as for example 'Public Policy Specification') is remanded to future work
- the scope of the study of cases, as well as its systematic review is limited mainly to one sector, in one country (systems research, UK)
- the quantitative analysis of the data can be susceptible to marginal error, especially where some of the data provided was a bit ambiguous (some projects have the same name but different project numbers, and vice versa.)

The roughness of the data has been addressed when selecting studies for inclusion, yet the percentages obtained from analysing the dataset should be considered indicative, rather than exact figures.

It is *not* in the scope of this study to name *all* the existing initiatives and groups that promote and lobby for the OA agenda, to index the volumes of research regarding economic and social benefits of Open Access, or to evaluate the role of individual initiatives or the activities of selected funders against another. A selection of disciplinary perspectives, issues and key references has been taken into account and included in this dissertation, mostly where these were critical and necessary to expound and address the research questions, at the expense of more complete analysis of all possible references.

7.5 DISCUSSION

In our contemporary networked society the active participation of individuals in all governance practices of institutions, is encouraged. It is the responsibility of individual researchers sometimes to monitor and even police the institutions they work for and operate in compliance with key strategic policies and best practices, but this cannot happen without researchers understanding the political and practical implications of information policies, and how they are used to control and manipulate knowledge flows, private commercial interests, and organisational hierarchies. Seeking knowledge is a key human behaviour, which influences social evolution (Douglas & Wykowski, 2010). Open web based technologies remove many of the physical barriers that prevent individuals to access knowledge freely, however many other barriers exist. Knowledge sharing policies, such as Open Access policies, an overview of which is provided in this thesis, are designed to reduce some of the non-physical barriers to knowledge sharing, such as restricted access to paying subscribers only. However these policies are being developed without taking into account the standpoint and requirements of knowledge seekers.

It is important that the effectiveness of these policies is evaluated against baseline criteria, such as '***does this policy solve the knowledge seeker problem?'*** Knowledge seekers have a place in contributing to policy making. This inquiry consolidates the notion that knowledge, to qualify as such, needs to be supported by verifiable evidence. In the context of increasingly complex, dynamic and rapidly evolving socio-technical realities, the ability to challenge pre-constituted assumptions with documented hard facts is a necessity. Truth can be unglamorous and controversial, especially when it challenges establishments operating to disguise it, and that a researcher's job can be a lonely, unsung affair, even when the work being done is fundamental and the results exciting. A wide combination of research methods can sometimes be the only way to tackle a complex, messy problem; however it can be difficult to argue, justify and validate. More specifically, this research has provided an opportunity to carry out a scientific investigation in an uncombed, emergent and trendy field of inquiry where much more work needs to be done. Combining diverse modes of inquiry and thought paradigms and complementing techniques to answer the research question and reach a conclusion teaches that even rigorous scientific process can be enhanced with creativity, self-determination and endurance.

7.6 FUTURE WORK

Much work lies ahead in resolving conflicts and optimising the distribution of key resources, such as 'new knowledge', which are essential for decision makers at any level. As this research shows, for normative interventions to be functional and effective they must be tackled systemically, and from many angles: policy makers should take a good look at knowledge sharing practices as a whole, and seriously consider individual instruments and initiatives based on the appraisal of factual evidence, and learn how to recognise the vacuity of measures which are not transparently documented.

Work ahead includes the development of open and collaborative information and knowledge infrastructures capable of supporting the pragmatic integration of the disparate aspects of complex socio-technical issues, of which knowledge sharing is an example, and the synthesis of existing technically proven techniques and established good practices to optimised knowledge flows is required. Policies can play a big role in the success of well informed decision making, provided they are not notional or cosmetic, or developed without looking at reality in the face. The convergence between 'academia', 'industry' and the 'public sector', begins with unrestricted, shared and 'user driven' knowledge flows and models. Adequate research and practice integration methods need to be developed, tested and applied in 'the real world'. More specific issues that emerged in the course of this research that were not exhausted within the scope of this study, come as natural pointers to future work, such as:

- Further research into knowledge sharing policies harmonization and specification
- Refinement of Open Access Monitor into an automated web service
- Further investigation of why, despite the existence of mandates and explicit monitoring of the same reported by literature, FOI requests do not yield evidence of the same.
- Ethnographic research, case studies and qualitative analyses of selected projects/audits of particular interest could be carried out.

The outcomes of this study invite more evidence-based research. The methodological approach and research instruments developed here could be applied to different socio-technical problem sets.

One of broad, long term tasks ahead that this research provides input to, is the development of a culture and sustainable economic models to ensure that knowledge is considered and handled as ‘public good’.

7.7 FINAL REMARKS

A knowledge seeker’s journey never ends.

This research crosses many disciplinary boundaries, and leverages a wide combination of methods to bridge different dimensions in the problem space. Most importantly, it produces a novel approach to tackle ’intractable’ socio-systemic entanglement and produces observable and replicable outcomes; surely an opportunity to experiment and learn, however a mere particle in an infinite and evolving universe of discourse. Scientific inquiry is one way of knowing the world, and there is no single path to it. Submitting a doctoral thesis (before the funding finishes) is an important milestone which this researcher finally reaches with these concluding remarks, fully aware that is also represents the beginning of just another phase of an endless cycle.

REFERENCES

Abdullah, N.S., Sadiq, S. and Indulska, M., (2009). “Emerging challenges in information systems research for regulatory compliance”, in: *Proceedings of the 22nd International Conference on Advanced Information Systems Engineering*, Berlin, Heidelberg: Springer-Verlag.

Ackoff R. L., (1971). “Towards a System of Systems Concepts”, *Management Science,* Vol. 17 no. 11, pp.661-671.

Ackoff, R. L., (1989). “From Data to Wisdom”, *Journal of Applied Systems Analysis,* Vol. 16, pp.3-9.

Aamodt, A. and Nygård, M., (1995). “Different roles and mutual dependencies of data, information, and knowledge - an AI perspective on their integration”, *Data and Knowledge Engineering* , Vol. 16, pp.191-222, Elsevier.

Alavi, M. and Leidner, D.E., (2001). “Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues”, *MIS Quarterly.*

Alter, S. and Dennis A.R., (2002). “Selecting Research Topics: Personal Experiences and Speculations for the Future”, *Communications of the Association for Information Systems* , Vol. 8, pp.314-329 .

Argote, L. and Ingram, P., (2000). “Knowledge transfer: A Basis for Competitive Advantage in Firms”, *Organizational Behavior and Human Decision Processes*, 82 (1), pp.150-169.

Bailey, C.W., (2005). "The role of reference librarians in institutional repositories", *Reference Services Review*, Vol.33, No.3, pp.259-267
in: Kyriaki-Manessi, D., "The Impact of Open Access Policies on Libraries: The New Era in Publishing Industry", Technological Educational Institution of Athens, Greece.

Balkin, (2006). Personal notes [online]
Available at:
<http://balkin.blogspot.com/2006/04/what-is-access-to-knowledge.html>
[Accessed: September 2010].

Batet, M., Gibert K. and Valls, A., (2007). "The Data Abstraction Layer as Knowledge Provider for a Medical Multi-agent System", *Proceedings K4CARE*, Vol.4942, pp.87-100, Springer-Verlag.

Baxter, G.D. and Sommerville, I., (2011). "Socio-technical systems: From design methods to systems engineering", *Interacting with Computers*.

Beck, T. and Buchanan-Smith, M., (2010). "Joint evaluations coming of age? The quality and future scope of joint evaluations". Alnap.org. [online]
Available at: <http://www.alnap.org/pool/files/7rha-Ch3.pdf >, [Accessed: February 2011].

Becker, J., Dreiling, A. and Ribbert, M., (2003). "Meta-Model-based Approaches to Information Systems Engineering", in *Proceedings of the 2003 Information Resources Management Association Conference*. Philadelphia, PA, USA.

Beer, M., (1998). “Organizational behavior and development”. Boston: Division of Research, Harvard Business School.

Bellinger, G., (2004). “Knowledge Management-Emerging Perspective”. John Wiley & Sons. [online]
Available at: <http://www.systems-thinking.org/kmgmt/kmgmt.htm>, [Accessed July 2011].

Berners-Lee, T., (1998). Design Issues. [online]
Available at: <http://www.w3.org/DesignIssues/Webize.html>.

Berners- Lee, T., Lassila, O. and Hendler, J., (2001). “The Semantic Web”, *Scientific American*, 284 (5) pp.34-43.

Berlin Declaration, (2003). Open Access. [online]
Available at: <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung>.

Berry, D., (2010). “Adaptive Clinical Trials: The Promise and the Caution”, *Journal of Clinical Oncology*, American Society of Clinical Oncology.

Bertalanffy von, L., (1968). “General System Theory: Foundations, Development, Applications”. New York: George Braziller.

Besemer, H., (2006). “Gathering evidence about the effectiveness of ‘open access’ publishing policies in agriculture”, *Agris Blogpost* . [online]
Available at: http://agriscontent.wordpress.com/2006/08/23/gathering-evidence-about-the-effectiveness-of-%E2%80%9Copen-access%E2%80%9D-publishing-policies-in-agriculture/, [Accessed: September 2011].

Bizer, C., Heath, T. and Berners-Lee, T., (2009). “Linked Data - The Story So Far”, *International Journal on Semantic Web and Information Systems* (IJSWIS), 5(3): 1-22. DOI: 10.4018/jswis.2009081901.

Björk, B.C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., et al., (2010). “Open Access to the Scientific Journal Literature: Situation 2009”. *PLoS ONE* 5(6): e11273. doi:10.1371/journal.pone.0011273.

Bjorner, D., (2007). “Domain Engineering, A Software Engineering Discipline in Need of Research”, in: SOFSEM (2000): *Theory and Practice of Informatics*, Vol. 1963 of *Lecture Notes in Computer Science*, pp.1-17. Springer-Verlag.

Boaz, A., and Ashby, D., (2003). “Fit for purpose? Assessing research quality for evidence based policy and practice”. London: ESRC UK Centre for Evidence Based Policy and Practice.

Bock, G.W., Zmud, R.W. and Kim, Y.G., (2005). “Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate”, *MIS Quarterly*, pp.87-11.

Boisot, M., (1995). “Information Space: A Framework for Learning in Organisations”. London: Routledge.

Boisot, M., (1998). “Knowledge Assets, Securing competitive advantage in the information economy”. Oxford: Oxford University Press.

Bouchard, T.J. Jr., (1976). "Field research methods: Interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures", in: Dunnette, M.D. (ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally College Publishing Company.

Browne, J., Sackett, P.J. and Wortmann, J.C., (1995). "Future manufacturing systems- Towards the Extended Enterprise", *Computers in Industry*.

Bryson, J.M., (2003). "What To Do When Stakeholders Matter: A Guide to Stakeholder Identification and Analysis Techniques", *National Public Management Research Conference*, Georgetown University Public Policy Institute, October 9-11.

Buneman P., Khanna S., and Tan, W. C., (2001). "Why and Where: A Characterization of Data Provenance", in *ICDT*.

Calise, M., de Rosa, R. and Fernandez, X. (2010). "Electronic Publishing, Knowledge Sharing and Open Access: A New Environment for Political Science", *European Political Science*, Vol.9, S50-S60, doi:10.1057/eps.2010.35.

Campbell, J. P., Daft, R. L. and Hulin, C. L., (1982). "What to study: Generating and developing research questions". Beverly Hills, CA: Sage.

Caracelli, V. J., and Greene, J. C. (1997). "Crafting mixed-method evaluation designs", in: Greene, J.C. and Caracelli, V.J. (Eds), "Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms", pp.19-32. San Francisco: Jossey-Bass.

Carlile, P.R., (2002). "A pragmatic view of Knowledge and Boundaries: Boundary Objects in New Product Development", *Organization Science,* Vol.13 Issue 4.

Carr, L., DeRoure, D., Harnad, S., Hey, J. et al., (2003). "Recommendations for UK Open Access Provision Policy", *UK Government Science and Technology Committee*. [online]
Available at: http://users.ecs.soton.ac.uk/harnad/Temp/UKSTC.htm
[Accessed: July 2011].

Cherns, A.B., (1987). "The principles of socio-technical design revisited", *Human Relations,* Vol. 40.3, pp.153-161.

Chou, S. W. and He, M. Y., (2004). "Creation Knowledge Management: The Distinctive Roles of Knowledge Assets in Facilitating Knowledge", *Journal of Information Science,* 30: 146.

Chourabi, O., Pollet, Y. and BenAhmed, M., (2010). "Ontology based experience Management for System Engineering Projects", *Proceedings of I-KNOW.*

Ciliska, D., Cullum, N. and Marks, S., (2001). "Evaluation of systematic reviews of treatment or prevention interventions", *EBN user's guide,* Vol. 4, October 2001.

Clark, T., Sammut, P. and Willans, J., (2008). "Applied Meta-modelling: A Foundation for Language Driven Development". Sheffield: Ceteva.

Clegg, C.W., (2000). "Socio-technical principles for system design", *Applied Ergonomics.*

Cohendet, P., Meyer Krahmer, F., (2001). “The theoretical and policy implications of knowledge codification”, *Journal: Research Policy* .
Collis, B., (2008). “A Pedagogy For Learners In The Co-Creation Of Knowledge And The Problems That Confront It In Practice”, *Proceedings of the LICK Symposium*, Edinburgh.

Colosi, L. and Dunifon, R., (2006). “ What’s the Difference? ‘Post then Pre’ & ‘Pre then Post’. Cornell University Publishing . [online]
Available at: <http://www.citra.org/Assets/documents/evaluation%20design.pdf>.

Conklin, J., (2005). “Dialogue Mapping: Building Shared Understanding of Wicked Problems”. Wiley: 1st edition, 18 November 2005.

Cook, T. D. and Campbell, D. T., (1979). “Quasi Experimentation: Design and Analytical Issues for Field Settings”. Chicago: Rand McNally.

Cooper, C.A., Alderliesten, R., La Rocca, G. and Benedictus, R., (2009). “In Search of a Knowledge Based Preliminary Design Method of Complex Aircraft Wings”, *7th Annual Conference on Systems Engineering Research*, Loughborough University.

Cooper, H.M., (1984). “ The problem formulation stage”, in: Cooper, H.M. (ed.), *Integrating Research: a Guide for Literature Reviews*. Newbury Park (CA): Sage Publications, 1984.

Cowan, R. and Foray, D., (1997). “The economics of codification and the diffusion of knowledge”, *Industrial and Corporate Change*, 6, pp.595-622.

Creswell, J. W., (2006). “Mixed Methods Research: Chapter 14 Mixed Research: Mixed Method and Mixed Model Research”, [online] <http://khup.com/download/37_keyword-mixed-methods-research/chapter-14-mixed-research-mixed-method-and-mixed-model-research.pdf> From: Creswell, J. W., (2006) “Designing and Conducting Mixed Methods Research”, 01-Creswell (Designing)-45025.qxd 5/16/2006 8:31 PM.

Creswell, J. W., (2008). “Research design: Qualitative, quantitative, and mixed methods approaches”, (3rd ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J, W., & Plano Clark, V., (2007). “Designing and Conducting Mixed Methods Research”. Thousand Oaks, CA: Sage Publications.

Cummings, J., (2003). Knowledge Sharing: A Review of the Literature. [e-journal] Available at: <http://lnweb90.worldbank.org/oed/oeddoclib.nsf/docunidviewforjavasearch/d9e389e7414be9de85256dc600572cao/$file/knowledge_eval_literature_review.pdf>. [Accessed: June 2011].

Czarnecki, K. and Eisenecker, U., (1999). “Generative Programming: Methods, Techniques and Applications”. Addison-Wesley.

Dais, A., Nikolaidou, M., Alexopoulou, N., Anagnostopoulos, D., Wimmer, M., Scholl, H. and Ferro, E., (2008). “Introducing A Public Agency Networking Platform towards Supporting Connected Governance”, *Springer 5184*, pp.375-387.

Damianou, N., Bandara, A.K., Sloman, M. and Lupu, E.C., (2002). "A Survey of Policy Specification Approaches", Imperial College of Science Technology and Medicine, London [online]
Available at: <http.www. doc.ic.ac.uk/~mss/Papers/PolicySurvey.pdf>.

Davenport, T.H. and Prusak, L., (1998). "Working Knowledge: How Organizations Manage What They Know". Cambridge, MA: Harvard Business School Press.

Davies, H., Nutley S. and Walter, I., (2005). "Assessing the impact of social science research: conceptual, methodological and practical issues", *A background discussion paper for ESRC Symposium on Assessing Non-Academic Impact of Research.*

De Geytere, T., (2005). "SECI model (Nonaka Takeuchi)" [online]
Available at: <http://www.12manage.com/methods_nonaka_seci.html>
[Accessed 29 October 2005].

Dentler, K, Cornet, R., et al. (2010). "Comparison of Reasoners for large Ontologies in the OWL 2 EL Profile", *Semantic Web Journal*. [e-journal]
Available at:
<http://www.semantic-web-journal.net/content/new-submission-comparison-reasonerslarge-ontologies-owl-2-el-profile>.

Denzin, N. K., (1978). "The research act: A theoretical introduction to sociological methods". New York: McGraw-Hill.

Di Maio, P., (2007). “Emergency and Poverty Relief ExpertFinder”, *First International ExpertFinder Workshop*, Berlin, Germany. [e-journal] Available at: <http://expertfinder.info/efw2007/items/maio.php>.

Di Maio, P., (2008). “Ontologies for Crisis Management”, in: Fiedrich, F. and van de Walle, B. (Eds), *Proceedings of the 5th International ISCRAM Conference*. Washington, DC, USA.

Di Maio, P., (2009) “Toward Pragmatic Dimensions of Knowledge Reuse and Learning on the Web”, *The 4th AIS SigPrag International Pragmatic Web Conference Track ICPW 2009)*, at the International Conference on Semantic Systems, Sep. 2009. Graz, Austria.

Dosi, G. and Marengo, L., (1994). “Some elements of an evolutionary theory of organisational competences”, in: England, R.W. (Ed.), *Evolutionary Concepts in Contemporary Economics*. The University of Michigan Press, Ann Arbor, MI.

Douglas, N. and Wykowski, T. S., (2010). “From Belief to Knowledge: Achieving and Sustaining an Adaptive Culture in Organizations”. CRC Press 1st Edition pp. 216.

Dusink, L. and Katwijk, J.V., (1995). “Reuse Dimensions”, *The ACM SIGSOFT Symposium On Software Reusability* (SSR’05), pp.137-149.

Eemeren, F., Grootendorst, R. et al. (1996). “Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments”, Lawrence Erlbaum Associates. Mahwah, NJ.

Emery, F. E., and Trist, E. L., (1960). “Socio-technical systems”, in: Churchman, C. W., and Verhulst, M. (Ed.), *Management Science Models and Techniques,* Vol. 2, pp. 83-97. Oxford, UK: Pergamon.

EU Commission Recommendation on the management of intellectual property in knowledge transfer activities and Code of Practice for universities and other public research organisations, (2008). Brussels, 10.4.2008 C(2008)1329. [e-journal] Available at: <http://ec.europa.eu/invest-in-research/pdf/ip_recommendation_en.pdf >.
[Accessed: April 2010].

EU PSI Directive (2003). [online] Available at: <http://ec.europa.eu/information_society/policy/psi/actions_eu/policy_actions/index_en.html>.
Fahey, L., and Prusak, L., (1998) “The Eleven Deadliest Sins of Knowledge Management”, *California Management Review.*

Federal Research Public Access Act. [online] Available at: <http://www.ala.org/ala/issuesadvocacy/access/accesstoinformation/publiclyfundedresearch/s1373.cfm>.
[Accessed: June 2011].

Feigenbaum, E.A. and McCorduck, P., (1983). “The fifth generation” (1st ed.). Addison-Wesley.

Feyerabend, P., (1975). “Against Method. Outline of an Anarchistic Theory of Knowledge”. London: New Left Books.

Fikes, R., Cutkosky, M., Gruber, T.R. and van Baalen, J., (1991). “Knowledge Sharing Technology Project Overview”, *Technical Report,* Stanford University, Knowledge Systems Laboratory.

Fishman, D., (1999). “The case for pragmatic psychology”. New York: New York University Press.

Fitzallen, N.E. and Brown, N.R., (2007). “Evidence-based research in practice”, in: *Engaging Pedagogies*, 26-30 November, Adelaide, SA.

Fluehr-Lobban, C., (1994). “Informed Consent in Anthropological Research: We Are Not Exempt”, *Human Organization*, 53, No. 1/Spring 1994, in: Murphy Dingwall 2007.

Focus Technical Brief, (2005). [online]
Available at: <http://www.ncddr.org/kt/products/focus/focus9/>.
[Accessed: September 2010].

Fox, J., (1984). “A short account of knowledge engineering”, *The Knowledge Engineering Review*, 1(1), pp.4-14.

Fox, J., (2011). “Formalizing knowledge and expertise: where have we been and where are we going?”, *The Knowledge Engineering Review,* Vol. 26, pp.5-10. Cambridge University Press.

Fransman, M., (1994). “Information, knowledge, vision and theories of the firm”, *Industrial and Corporate Change*, 3 (3), pp.713-75.

Gardner, P.L., Fong,Y. and Huang, R. L., (2010). "Measuring the Impact of Knowledge Transfer from Public Research Organizations: A Comparison of Metrics Used Around the World", *International Journal of Learning and Intellectual Capital*, Vol. 7, Nos. 3-4.

Garvin, D.A., (1988). "Managing Quality", Free Press.

Genova, G., (2009). "Modelling and metamodelling in Model Driven Development", *What is a metamodel: the OMG's metamodelling infrastructure* , Warsaw. [e-journal]
Available at: <http://www.ie.inf.uc3m.es/ggenova/Warsaw/Part3.pdf>.

Genzuck, M.A., (2000). "Synthesis of Ethnographic Research Center for Multilingual, Multicultural Research", University of Southern California.

Gettier, E., (1963). "Is Justified True Belief Knowledge?", *Analysis* 23, pp.121-123.

Gómez, N., Bustos Gonzalez, A., Santillan Aldana, J. and Arias, O., (2009). "Open access indicators and information society: the Latin American case", *OCLC Systems Services*, Vol. 25, Emerald.

Gorod, A., Sauser, B. and Boardman, J., (2008). "Paradox: Holarchical View of System of Systems Engineering Management", in: *IEEE International Conference on System of Systems Engineering*, Sch. of Systems & Enterprises, Stevens Institute of Technology.

Grant, R.,(1996). "Toward a Knowledge-Based Theory of the Firm", *Strategic Management Journal,* Vol. 17, Special Issue: Knowledge and the Firm, pp.109-122.

Gray, G.E., (2004). "Concise Guide to Evidence-based Psychiatry", *American Psychiatric Publications*, Washington, DC.
Gruber, T., (1995). "Toward principles for the Design of Ontologies Used for Knowledge Sharing", *International Journal Human-Computer Studies* Vol. 43, Issues 5-6, pp.907-928.

Guba, E.G., (1990). "The Alternative Paradigm Dialog", in: Guba, E.G., (Ed.) *The Paradigm Dialog*. London: Sage Publications, pp.17-27.

Guba, E.G. and Lincoln, Y.S., (1988). "Do inquiry paradigms imply inquiry methodologies?", in: D.M. Fetterman (Ed.), *Qualitative approaches to evaluation in education. The silent revolution,* pp.89-115. New York: Praeger.

Guha, R., McCool, R. and Fikes, R., (2004). "Contexts for the semantic web", *International Semantic Web Conference*, Vol. 3298 of Lecture Notes in Computer Science.

Harsh, O.K., (2008). "Reusable Data, Information, Knowledge and Management Techniques", *Journal of Knowledge Management Practice*, Vol. 9, No. 3.

Hatcher, S., Butler, R. and Oakley-Browne, M., (2005). "Evidence-based Mental Health Care". Livingstone, Edinburgh: Elsevier Churchill.

Hauser, H., (2010). “The Current and Future Role of Technology and Innovation Centres in the UK”, *A report for Lord Mendellson*. [e-journal]
Available at: <http://www.innovateuk.org/_assets/pdf/other-publications/hauser-review.pdf>.
[Accessed: June 2011].

Heidegger, M., (1962). “Being and Time”. New York: Harper & Row [Originally in German, 1927].

Hendler, J. and van Harmelein, F., (2008). “Handbook of Knowledge Representation”. van Harmelein, F., Lifschitz, V. and Porter B. (Eds.). Elsevier B.V.

Hendriks, P.H.J., (1999). “Why Share Knowledge? The Influence of ICT on Motivation for Knowledge Sharing”, *Knowledge and Process Management*, 6(2), pp.91-100.

Hollnagel, E., (2008). “The Changing Nature of Risk”, *HFESA Hournal, Ergonommics Australia*, Vol.22, No.1

Hollnagel, E., (ed.) (2003). “Handbook of cognitive task design”. Mahwah, NJ: Erlbaum.

Holsapple, C.W., (2004). ”Handbook on Knowledge Management: Knowledge matters”. Birkhäuser.

Holten, R., (2000). “Framework and Method for Information Warehouse Development Processes”, *Proceedings of the Data Warehousing 2000 Conference*, pp.135-163.

Holz, B., (2003). “Human Knowledge: Foundations and Limits”. [online] Available at: <http://humanknowledge.net/>.
[Accessed: August 2011].

Honour, E. and Valerdi, R., (2006). “Advancing an ontology for systems engineering to allow consistent measurement”, *Conference on Systems Engineering Research*, Los Angeles, CA.

Horn, R.E., (2001). “Knowledge Mapping for Complex Social Messes”, a Stanford University presentation to the *Foundations in the Knowledge Economy conference* at the David and Lucile Packard Foundation, July 16, 2001.

Horsborough, D., (2003). “Evaluation of qualitative research”, *Journal of Clinical Nursing* 12: pp.307-31.
Itami, H., (1987). “Mobilizing Invisible Assets”. Boston, MA: Harvard University Press, USA.

Jamali, R.H., Nicholas, D. and Huntington, P., (2005). “The use and users of scholarly e-journals: a review of log analysis studies”, *Aslib Proceedings.*

Jankowski, N.W., (ed.) (2009). “The Contours and Challenges of e-Research”, *E-Research: Transformations in Scholarly Practice*, Introduction. New York: Routledge, pp. 3-34.

Jarvenpaa, S.L. and Staples, D.S., (2000). "The Use of Collaborative Electronic Media for Information Sharing: An Exploratory Study of Determinants", *Journal of Strategic Information Systems*, 9(2-3), pp.129-154.

Jick, T.D., (1983). "Mixing qualitative and quantitative methods: Triangulation in action", in: van Maanen, J. (Ed.), *Qualitative methods.* Beverly Hills: Sage.

Johnson, R.B. and Ongwubuzie, A.J., (2004). "Mixed Methods Research: A Research Paradigm Whose Time Has Come", *Educational Researcher*, Vol.33, No.7, pp.14-26.

Jones, S., (2009). "A report on the range of policies required for and related to digital curation", Digital Curation Centre, University of Glasgow, UK.

Joubish, M.F., Khurram, M.A., Ahmed, A., Fatima, S.T. and Aider, K.H., (2011). "Paradigms and Characteristics of Good Qualitative Research", *World Applied Sciences Journal,* Department of Education, Federal Urdu University. Pakistan: IDOSI Publications.

Kalawski, R.S., (2009). "Grand Challenges for Systems Engineering Research, *Proceedings of the 7th Annual Conference on Systems Engineering Research*, Loughborough University.

Kim, M.P., (2005). "Knowledge Retention Enhances Performance-Based Management", *DCMA Communicator*, FALL 2005.

Kingston, J., (1994). "Linking Knowledge Acquisition with CommonKADS Knowledge Representation", AIAI-TR-156. [e-journal]
Available at:
<http://www.cos.ufrj.br/~jano/LinkedDocuments/_papers/aula07/Linking%20Knowledge%20Acquisition%20with%20CommonKADS%20Knowledge%20Representation.pdf> [Accessed: October 2011].

Kitamura, Y., Kashiwase, M., Fuse, M. and Mizoguchi, R., (2004). "Deployment of an Ontological Framework of Functional Design Knowledge", *Advanced Engineering Informatics*, Vol.18, Issue 2, DOI:10.1016/j.aei.2004.09.002.

Kogut, B. and Zander, U., (1992). "Knowledge of the firm, combinative capabilities, and the replication of technology", *Organisation Science* 3 (3), pp.383-397.

Kuhn, T.S., (1970). "The Structure of Scientific Revolutions", 2nd edition. Chicago: University of Chicago Press, pp.55.

Küsters, R., Truderung, T. and Vogt, A., (2010). "Accountability: definition and relationship to verifiability", *Proceedings of ACM Conference on Computer and Communications Security* 2010 pp.526-535.

Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B.C. and Hedlund, T., (2011). "The Development of Open Access Journal Publishing from 1993 to 2009" Source: PLOS ONE, Vol. 6 Issue: 6, Article Number: e20961.

Laszlo, K., (2011). "From Systems Thinking to System Being", Keynote speech, *International Symposium of the System Sciences*, Hull July 2011.

Lane, J.A. and Valerdi, R., (2010). “Accelerating System of Systems Engineering Understanding and Optimization through Lean Enterprise Principles”. [e-journal] Available at: <http://csse.usc.edu/csse/TECHRPTS/2010/usc-csse-2010-504/usc-csse-2010-504.pdf>.

Lanter, D.P., (1991). “Design of a Lineage-Based Meta-Data Base for GIS”, *Cartography and Geographic Information Systems*, Vol. 18.

Lee, H., Sajjad, F., Kamal, M. and Irani, Z., (2011). “Workflow technology as an e-participation tool to support policy-making”, *Journal of Enterprise Information Management*, Vol.24, Issue 2, pp.197-212.

Lee, T.W., Mitchell, T.R. and Sablynski, C.J. , (1999). “Qualitative research in organizational and vocational psychology, 1979 - 1999”. *Journal of Vocational Behavior*, 55, pp.161-187.

Leonard-Barton, D., (1995). “Well Springs of Knowledge: Building and Sustaining the Sources of Innovation”. Harvard Business School Press.

Levitt R., Martin S., Nutley S. and Solesbur, W., (2010). “Evidence for accountability”, Nuffield Foundation.

Liebowitz, J., (1999). *The Knowledge Management Handbook*. Boca Raton, FL: CRC Press.

Liebowitz, J., Rubenstein-Montano, B., McCaw, D. and Buchwalter, J., (2000). “The knowledge audit”, *Knowledge and Process Management*; Jan/Mar 2000; 7, 1; ABI/INFORM Global.

Lynne, M., (2001). "Toward a Theory of Knowledge Reuse: Types of Knowledge Reuse Situations and Factors in Reuse Success", *Journal of Management Information Systems*, Vol.18, Issue 1.

Madill, A. and Jordan, A., (2000). "Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies", *British Journal of Psychology*, Wiley Online Library.

Majchrzak, A., Neece, O.E. and Cooper, L.P., (2000). "Knowledge Reuse For Innovation - The Missing Focus In Knowledge Management", *Academy of Management Annual Conference*, Washington D.C., USA.

Mayr, P., (2006). "Constructing experimental indicators for Open Access documents", *Institute of Library and Information Science*, Humboldt-University, Berlin, Germany.

Mearns, M.A. and du Toit, A.S.A., (2008). "Knowledge audit: Tools of the trade transmitted to tools for tradition", *International Journal of Information Management*, 28(1), pp.161-167.

Medeni, T., Iwatsuki, S. and Cook, S., (2008). "Reflective Ba and Refractive Ma in Cross-Cultural Learning", in: Putnik, G.D. and Cunha, M.M. (Eds.) *Encyclopedia of Networked and Virtual Organizations Idea Group.*

Mellor, S.J., Scott, K., Uhl, A. and Weise,D., (2004) "MDA distilled: principles of model-driven architecture", *Computers*, Addison-Wesley Professional.

Menon, T. and Pfeffer, J., (2003). “Valuing Internal vs. External Knowledge: Explaining the Preference for Outsiders”, *Management Science* Vol. 49 No. 4: Special Issue in Managing Knowledge in Organizations: Creating, Retaining, and Transferring Knowledge (Apr 2003) pp.497-513.

Mertens, D., (1998). “Research methods in education and psychology: integrating diversity with quantitative & qualitative approaches”. Sage Publications.

Mertens, D., (2010). “Transformative Mixed Methods Research Qualitative Inquiry Qualitative Inquiry” 2010, 16, pp.469.

Miller, F., (1986) “Use, appraisal, and research: A case study of social history”, *The American Archivist:* 49(4), pp.371-392.

MODAF Metamodel Description Document, [e-journal]
Available at: <http://www.mod.uk/NR/rdonlyres/0360559B-3D6B-4198-BD92-D2B783A7E831/0/20090310_MODAF_Meta_Model_V1_0_U.pdf>.
[Accessed: 01 August 2011]

Morris, C.W., (1946). “Signs, Language, and Behavior”. New York: Prentice-Hall.

Mosteller, F., and Boruch, R. (Eds.), (2002). “Evidence matters: Randomized trials in education research”. Washington, DC: The Brookings Institute.

Murphy, E. and Dingwall, R., (2007). “Informed Consent, Anticipatory Regulation and Ethnographic Practice”, *Social Science & Medicine,* 65, 2223-223.

Neighbors, (1980). “Software Construction Using Components”, *Doctoral Dissertation*, University of California, Irvine. [e-journal]
Available at: <http://www.bayfronttechnologies.com/thesis.pdf>.
[Accessed: October 2010].

Nelson, K.M. and Cooprider, J.G., (1996). “The Contribution Shared Knowledge to IS Group Performance”, *MIS Quarterly*, 20(4), pp.409-432.

Newell, A., (1980). “The Knowledge Level”, Presidential Address, *American Association for Artificial Intelligence*, Stanford University.

Newell, W.H., (2007). “Distinctive Challenges of Library-based Inter-disciplinary Research and Writing: A Guide”, *Issues in Integrative Studies.* [e-journal]
Available at:
<http://www.units.muohio.edu/aisorg/PUBS/ISSUES/25_newell.pdf>.

Nicholas, D., Huntington, P., Dobrowolski, T., Dobrowolski, T., Rowlands, I., Jamali, H.R. and Polydoratou, P., (2005). “Revisiting ‘obsolescence’ and journal article ‘decay’ through usage data: an analysis of digital journal use by year of publication”, *Information Processing & Management: an International Journal - Special Issue,* Vol. 41, Issue 6. Tarrytown, New York: Pergamon Press, Inc.

Nielsen, J., (1994). “Heuristic evaluation”, in: Nielsen, J., and Mack, R.L., “Usability Inspection Methods”. John Wiley & Sons: New York, NY.

Nishida, K., (1990). “An inquiry into the good”, Yale University Press.

Nissen, M. and Levitt, R., (2002). “Dynamic Models of Knowledge-Flow”, *Dynamics Working Paper 76*, Center for Integrated Facility Engineering, Stanford University.

Nonaka, I. and Konno, N., (1998). “The concept of ’Ba’: Building foundation for Knowledge Creation”, *California Management Review*, Vol. 40, No.3.

Nonaka, I. and Takeuchi, (1995). “The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation”. New York, NY: Oxford University Press.

Norman, D.A. (1980). “Twelve Issues for Cognitive Science”, *Cognitive Science 4 1-32*. University of California, San Diego. [e-journal]
Available at: <http://psych.stanford.edu/~jlm/pdfs/Norman80TwelveIssues.PDF>. [Accessed: December 2010].

Norris, D.M., Mason, J., Robson, R., Lefrere, P. and Collier, G., (2003). “A Revolution in Knowledge Sharing”, Educause Review 15, [e-journal]
Available at: <net.educause.edu/ir/library/pdf/erm0350.pdf>.

OAIS, (2002). “Reference Model for an Open Archival Information System”, *Recommendation for Space Data System Standards*. [e-journal]
Available at: <http://public.ccsds.org/publications/archive/650x0b1.PDF>. [Accessed: September 2010].

OECD Declaration, (2007). “OECD Principles and Guidelines for Access to Research Data from Public Funding”, *Organisation for Economic Co-operation and Development* [e-journal] Available at: http://www.oecd.org/dataoecd/9/61/38500813.pdf [Accessed: July 2010].

O'Leary, Z., (2004). “The essential guide to doing research”. London: Sage.

Packer, M.J., (2010). *The science of qualitative research.* New York: Cambridge University Press.

Pan, S.L. and Scarborough, H., (1999). “Knowledge Management in Practice: An Exploratory Case Study of Buckman Labs”, *Technology Analysis and Strategic Management* 11(3) pp.359-74.

Patton, M.Q., (1988). “Paradigms and pragmatism”, in: Fetterman, D.M. (Ed.), *Qualitative approaches to evaluation in education. The silent revolution* pp.89-115. New York: Praeger.

Paynter, R., (2009). “Evidence-based research in the applied social sciences”, *Reference Services Review* ,Vol. 37 No. 4, 2009 pp. 435-450, Emerald.

Pearce, L., (2004). “How to Examine A Thesis”. Open University Press.

Plano Clark, V., (2010). “The Adoption and Practice of Mixed Methods: U.S. Trends in Federally Funded Health-Related Research”, *Qualitative Enquiry*. Sage.

Poddar, A. and Ruthven, I., (2010). “The emotional impact of search tasks”, *Proceedings of the 3rd Information Interaction in Context Symposium* (IIiX 2010).

Popper, K.R. (1963). “Conjectures and Refutations: The Growth of Scientific Knowledge”. New York: Routledge and Kegan Paul, pp.88.

Potucek, M., (2003). “Dimensions of Public Policy: Values, Processes, Implementation and Results”, in: Potucek, M., Leloup, L., Jenei, G. and Váradi, L., (Eds) (2003). “Public Policy in Central and Eastern Europe: Theories, Methods and Practices”, Bratislava: NISPAcee.

Prencipe, A., and Tell, F., (2001). “Inter-project learning: processes and outcomes of knowledge codification in project-based firms”, *Research Policy 30* 1373-1394. [e-journal]
Available at: <http://fandango.cs.unitn.it/~rcuel/filippo/sciserv.pl.pdf>.
[Accessed: January 2011].

Polkinghorne, D., (2007). “Validity Issues in Narrative Research”, *Qualitative Enquiry*. Sage Publications.

Raudenbush, S., (2002). “Identifying scientifically-based research in education”, Invited speaker at the *Scientifically Based Research Seminar*, U.S. Department of Education, Washington DC.

Geisberger, E., Broy, M., Berenbach, B., Kazmeier, J., Paulish, D. and Rudorfer, A., (2006). “Requirements Engineering Reference Model (REM)”, Technische Universität München, [e-journal]
Available at: <http://www4.informatik.tu-muenchen.de/publ/papers/TUM-I0618.pdf>
[Accessed: February 2011].

Regulation 1515, (2005). “The Re-use of public Sector Information Regulations” [online]
Available at: <http://www.legislation.gov.uk/uksi/2005/1515/contents/made> [Accessed: November 2010].

Reza-Tazari, M., (2003). “A Context-Oriented RDF Database”, in Proceedings *Semantic Web Databases Conference*.

Rhodes, D. and Valerdi, R., (2007). “Enabling Research Synergies through a Doctoral Research Network” for *Systems Engineering Systems Engineering*, Vol. 10, Issue 4.

Richter, F. and Vettel, K., (1995). “Successful Joint Ventures in Japan: Transferring Knowledge Through Organizational Learning”, *Long Range Planning*, Vol. 28, No. 3.

Ritchey, T., (1996). “Analysis and Synthesis On Scientific Method - Based on a Study by Bernhard Riemann”, (Revised Version) *Systems Research*, 1991, Vol. 8, No. 4, pp.21-41, Thesis Publishers.

Ritchey, T. (2005). “Wicked Problems: Structuring Social Messes with Morphological Analysis”, adapted from a lecture given at the Royal Institute of Technology in Stockholm.

Rizzo, A., Pozzi, S., Save, L. and Sujan, M., (2005). “Designing complex socio-technical systems: A heuristic schema based on Cultural-Historical Psychology”, *ACM International Conference Proceeding Series*; Vol. 132, Proceedings of the 2005 annual conference on European association of cognitive ergonomic.

Rittel, H. and Webber, M., (1973). “Dilemmas in a General Theory of Planning,” *Policy Sciences 4*. Amsterdam: Elsevier Scientific Publishing, pp.155-159.

Rott, H., (2003). “Lehrer’s dynamic theory of knowledge”, in: Olsson, E. (Ed.), *The Epistemology of Keith Lehrer, Kluwer, Dordrecht*, pp.219-242.

Rugg, G., and Petre, M., (2004). “The Unwritten Rules of Ph.D. Research”. Open University Press: McGraw – Hill.

Russell, B., (1913). “Theory of Knowledge: The 1913 Manuscript”, in: Eames, E.R. & Blackwell, K. (Eds.) (1984). London: George Allen & Unwin.

Russell, B. (1957). in: *The Scientific Outlook* (2009), Taylor & Francis, US.

Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B. and Richardson, W.S. (1996). “Evidence-based medicine: what it is and what it isn’t”, *British Medical Journal*, Vol. 312 No. 7023, pp.71-2.

Salomon, G., (1991). “Transcending the qualitative-quantitative debate: The analytic and systemic approaches to educational research”, *Educational Researcher,* 20(6), pp.10-18.

Sanders, L., (2008). “An Evolving Map of Design Practice and Design Research”, *Interactions Magazine*, Nov-Dec.

Sang, M.L. and Soongoo, H. (2002). “An enterprise-wide knowledge management system infrastructure”, *Industrial Management & Data Systems*, Vol. 102 Issue 1, pp.17 - 25.

Sauser, B. and Boardman, J., (2006). "From Prescience to Emergence: Taking Hold of System of Systems Management" in *27th Annual National Conference of American Society for Engineering Management*. Huntsville, AL.

Savin-Baden, M. and Fisher, A., (2002). "Negotiating 'Honesties' in the Research Process", *British Journal of Occupational Therapy,* 65(4), pp.191-193.

G20 DWG, (2011). "Scaling Up Knowledge Sharing for Development- A Working Paper" for the *G-20 Development Working Group*. [e-journal]
Available at:
<http://einstitute.worldbank.org/ei/sites/default/files/Upload_Files/G-20_DWG_report_on_pillar_9_10_June.pdf>

Schoop, M., de Moor, A. and Dietz, J., (2008). "Pragmatic Web Manifesto", *Communications of the ACM.*

Schmitt, J. F., (n.d.) "A Systemic Concept for Operational Design". [unpublished article] Available at:
http://www.au.af.mil/au/awc/awcgate/usmc/mcwl_schmitt_op_design.pdf
[Accessed: April 2010].

Schreiber, G., Wielinga, B. and Breuker, J., (1993). "KADS: A principled approach to knowledge-based system development". London: Academic Press.

Schreiber, G. and Wielinga, B., (1993). "Knowledge Model Construction", in: *KADS: A principled approach to knowledge-based system development*. London: Academic Press.

Schubert, P., Lincke, D., and Schmid, B., (1998). “A Global Knowledge Medium as a Virtual Community: The NetAcademy Concept”, in *Proceedings of the Fourth Americas Conference on Information Systems.*

Sharma, R. and Chowdhury, N., (2007). “On The Use Of A Diagnostic Tool For Knowledge Audits”, *Journal of Knowledge Management Practice,* Vol. 8, No. 4.

Shavelson, R.J. and Towne, L., (2002). “Scientific research in education”. Washington, DC: National Research Council, National Academy Press.

Sheffield, J., (2005). “Systemic knowledge and the V-model”, *Department International Journal of Business Information Systems.*

Shimizu, H., (1995). “Ba-Principle: New Logic for the Real-time Emergence of Information”, *Holonics*, 5/1 (1995), pp.67-69.

Shulman, L.S., (1988). “Disciplines of inquiry in education: An overview”, in: Jaeger, R.M. (Ed.),*Complementary methods for research in education,* pp.3-17. Washington: AERA.

Sillitto, H., (2006). “Systems Engineering and Social Systems Engineering for Network-Enabled Capability”, EUSEC.

Simmhan, Y.L., Plale, B. and Gannon, D., (2005). “A Survey of Data Provenance in e-Science”, *SIGMOD Record*, Vol. 34, No. 3.

Sloman, M., (1994). “Policy Driven Management for Distributed Systems”, *Journal of Network and Systems Management*, Chap 2, pp.333.

Smith, B., (2003). "Public Policy and Public Participation – Engaging Citizens and Community in the Development of Public Policy", prepared for *Population and Public Health Branch*, Canada

Smith, E., (2001). "The role of tacit and elicit knowledge in the workplace", *Journal of Knowledge Management*, Vol. 5 No. 4.

Sober, E., (1981). "The Principle of Parsimony", *British Journal for the Philosophy of Science*, Vol.32(2), pp.145-156.

Sommerville, I., (2004). *Software Engineering*, 7th edition, Chapter 2.

Song, H. and Kondou, S., (2006). "Regional Knowledge Creation in Cluster Promotion Organizations Through 'MA' Management", Proceedings of the 7th International Symposium on Knowledge and Systems Sciences, *Towards Knowledge Synthesis and Creation*, September 22-25.

South Africa, Dept. of Science & Technology, Organisation for Economic Co-operation and Development, (2005). "Integrating Science & Technology into Development Policies: An International Perspective", OECD Workshop. [online] Available at:
<http://www.oecd.org/document/35/0,3746,en_2649_34499_35048867_1_1_1_1,00.html>
[Accessed: February 2011].

Spender, J.C., (1996). "Making knowledge the basis of a dynamic theory of the firm", *Strategic Management Journal 17.*

Stanage, S.M., (1987). *Adult Education and Phenomenological Research.* Malabar, Florida: Robert E. Krieger Publishing Company.

Stewart, J. and Williams, R., (2005). "The Wrong Trousers? Beyond the Design Fallacy: Social Learning and the User", in Rohracher, H., (2005). *User involvement in innovation processes. Strategies and limitations from a socio-technical perspective*. Munich: Profil-Verlag.

Stojanovic, N., Stojanovic, L. and Handschuh, S., (2002). "A Framework for Knowledge Management on the Semantic Web", in: *The Eleventh International World Wide Web Conference*, Poster Track, Hawaii.

St.Pierre, E.A., (2002). "Science rejects postmodernism", *Educational Researcher, 31(S),* pp.25-27.

Studer, R., Benjamins, V.R. and Fensel, D., (1998). "Knowledge Engineering: Principles and Methods, Data and Knowledge Engineering (DKE)", 25(1-2), pp.161-197.

Suber, P., (2009). "Timeline of the Open Access Movement: Formerly called the Timeline of the Free Online Scholarship Movement". [online]
Available at: <http://www.earlham.edu/~peters/fos/timeline.htm>
[Accessed: March 2011].

Sveiby, K.E., (1997). "The New Organizational Wealth: Managing and Measuring Knowledge-Based Assets". San Francisco: Berrett-Koehler Publishers, Inc.

Sveiby, K.E., (2001). “Methods for Measuring Intangible Assets”, updated 27 April 2010. [online]
Available at: <http://www.sveiby.com/articles/IntangibleMethods.htm >
[Accessed: May 2011].

Swan, A., (2008). “Key concerns within the scholarly communications process”, *Report to the JISC Scholarly Communications Group*. [online]
Available at:
<http://www.jisc.ac.uk/whatwedo/topics/opentechnologies/openaccess/reports/keyconcerns>
[Accessed: December 2011].

Szulanski, G., Cappetta, R. and Jensen, R.J., (2004). “When and how trustworthiness matters: Knowledge transfer and the moderating effect of causal ambiguity”, *Organization Science*, 15, pp.600—613.

Tao, F., Millard, D., Woukeu,, A. and Davis, H., (2005). “Semantic Grid based e-Learning using the Knowledge Life Cycle”, in: *International Workshop on Applications of Semantic Web Technologies for E-Learning* (SW-EL).

Tham, K., Donald, H. and Kim, M., (2002). “Towards Strategic Intelligence with Ontology Based Enterprise Modelling and ABC”, in *Proceedings of the IBER Conference*, Las Vegas, Nevada, USA, October.

Tolk, A., (2003). “Beyond Technical Interoperability - Introducing a Reference Model for Measures of Merit for Coalition Interoperability”, in Proceedings of *The Command and Control Research and Technology Symposium* (CCRTS), CCRP Press.

Thorley, M., (2007). “The Data Sharing Policies of the UK Research Councils: Principles and Practices”, in: *Berlin 5 Open Access: From practice to impact: Consequences of Knowledge dissemination*, Padova.

Tuomi, I., (1999). “Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory”, *System Sciences*, HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference.

Umpleby , S.A. , (2001). “Knowledge Management From The Perspective Of Systems Theory And Cybernetics” , *Society for the Advancement of Socio-Economics,* Amsterdam, The Netherlands, June.

UNDP, (2002). *Handbook on Monitoring and Evaluating for Results*, Evaluation Office. [e-journal]
Available at: <http://www.undp.org/evaluation/documents/HandBook/ME-HandBook.pdf>
[Accessed: October 2011].

Unico Report, (2008). “Metrics for the Evaluation of Knowledge Transfer Activities at Universities”. [e-journal]
Available at: <http://ec.europa.eu/invest-in-research/pdf/download_en/library_house_2008_unico.pdf>
[Accessed: June 2011].

Uschold, M. and Gruninger, M., (1996). “Ontologies: Principles, methods and applications”, *Knowledge Engineering Review*, Vol.11(2), pp.93-136.

Valentine, J.C., Pigott, T.D., and Rothstein, H.R., (2010). "How many studies do you need?: A primer on statistical power in meta-analysis", *Journal of Educational and Behavioral Statistics,* 35, pp.215-247.

Valerdi, R., (2010). "Heuristics for Systems Engineering Cost Estimation", *IEEE* Systems Journal 5(1), pp.91-98.

Vasconcelos, J.B., Kimble, C. and Gouveia, F., (2000). "A Design for a Group Memory System using Ontologies", *Proceedings of 5th UKAIS Conference.* Cardiff, Wales: McGraw Hill.
Vicente, K.J., (1999). "Cognitive Work Analysis: Toward safe, productive, and healthy computer based work". Mahwah, NJ: Erlbaum.

Von Krogh, G., Ichijo, K. and Nonaka, I., (2001). "Emergence of 'Ba': A Conceptual Framework for the Continuous and Self-transcending Process of Knowledge Creation", in: Nonaka, I. and Nishiguchi, T. (Eds.) *Knowledge Emergence: Social, Technical and Evolutionary Dimensions of Knowledge Creation*, pp.13-30. Oxford: Oxford University Press.

Yin, R., (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage Publishing.

Wang, S. and Noe, R.A., (2010). "Knowledge sharing: A review and directions for future research", *Human Resource Management Review* 20, pp.115-131.

Weber, R., (2007). "Knowledge Management in Call Centres: Addressing Failure Factors in Knowledge Management", *The Electronic Journal of Knowledge Management,* Vol. 5 Issue 3, pp.333 - 346.

Webb, C., (1992). “The use of the first person in academic writing: objectivity, language and gatekeeping”, *Journal of Advanced Nursing*, 17, pp.747-752.

White, B., (2003). “Web Accessibility, Mobility and Findability”, *Proceedings of Web Congress*, Latin American, First Latin American Web Congress.

Williams, R., Stewart, J. and Slack, R., (2005). “Social learning in technological innovation: Experimenting with information”. Elgar Publishing.

Willig, C., (2001). “Introducing Qualitative Research in Psychology: Adventures in Theory and Method”. Buckingham: Open University Press.

Zhang, Z. and Siekmann, J., (2007). “Knowledge Science”, *Engineering and Management Second International Conference*, KSEM 2007, Melbourne, Australia, November 28-30, Proceedings- Lecture Notes in Computer Science, Vol. 4798.

von Zweck, C., Paterson, M. and Pentland, W. (2008). “The Use of Hermeneutics in a Mixed Methods Design”, *The Qualitative Report*, Vol.13 No.1, pp.116-134.

Other References Not Cited in Text:

Ackoff, R. L., (1993). “From Mechanistic to Socio-Systemic Thinking”, *Proceedings of Systems Thinking in Action Conference,* Pegasus Publishing [online] Available at: <http://acasa.upenn.edu/socsysthnkg.pdf>

Bazeley, P., (n.d.). "Mixed methods workshop notes". [online] Available at: <http://www.era.edu.au/myfiles/uploads/Mixed%20methods%20workshop%20notes.pdf> [Accessed: June 2010].

Challeplioglou Artemis Foundation of Biomedical Research of the Academy of Athens Vasilakaki Eugenia DBS SA Informatic Services Company. [e-journal] Available at: <http://www.inforum.cz/pdf/2006/Kyriaki_Manessi_Daphne.pdf>.

Committee on Facilitating Interdisciplinary Research, (2005). *Facilitating Interdisciplinary Research*, National Academies Press.

Currie, G., Finn, R. and Martin, G., (2007). "Spanning boundaries in pursuit of effective knowledge sharing within networks in the NHS", *Journal of Health Organization and Management,* Vol. 21 No. 4/5, pp.406-417. Emerald Group Publishing Limited.

Hanrahan, M.U., (2001). "The body in question in research methodology: 'Death of the Author' and 'the Pearly Gates of Cyberspace', in: Singh, P. and McWilliam, E.L. (Eds.), *Designing Educational Research: Theories, Methods and Practices*. Post Pressed, Tenerife, Qld, pp.233-248.

http://www.w3.org/TR/ws-arch/

http://ec.europa.eu/enterprise/sectors/ict/files/overview_report_results_consultation_en.pdf

Improving knowledge transfer between research institutions and industry across Europe [e-journal]
Available at: <http://ec.europa.eu/invest-in-research/pdf/download_en/knowledge_transfe_07.pdf>.

Interim Report of the Lisbon Expert Group, (2007). "Towards open and systemic research and innovation policies for Europe".

Laws, S., (2011). "Giving Effect to Policy in Legislation: How to Avoid Missing the Point", *Statute Law Review* 32(1), pp.1-16. Crown Copyright 2011.

Limpiens, et al., (n.d.). "Linking Folksonomies and Ontologies for Supporting Knowledge Sharing: a State of the Art". [e-journal]
Available at: <http://isicil.inria.fr/v2/res/docs/livrables/ISICIL-ANR-EA01-FolksonomiesOntologies-0906.pdf>.

Pan, S.L. and Leidner, D.E., (n.d.). "Bridging communities of practice with information technology in pursuit of global knowledge sharing".

Polanyi, M., (1966). "The Tacit Dimension". New York: Anchor Day Books.

INDEX OF APPENDICES