

Identity Awareness and Re-use of Research Data in Veillance and Social Computing

Alexander Hayes
University of Wollongong, Australia
ah691@uow.edu.au

Professor Steve Mann
University of Toronto, Canada
mann@eecg.toronto.edu

Amir Aryani
RMIT University, Australia
amir.aryani@rmit.edu.au

Susannah Sabine
IEEE SSIT, Australia
susannah@susannahsabine.com

Leigh Blackall
La Trobe University, Australia
l.blackall@latrobe.edu.au

Pia Waugh
Department of Finance and Deregulation, Australia
pia.waugh@finance.gov.au

Stephan Ridgway
TAFE NSW, Australia
stephan.ridgway1@det.nsw.edu.au

Abstract—Identity awareness of research data has been introduced as a requirement for open research, transparency and reusability of research data in the context of eScience. This requirement includes the capability of linking research data to researchers, research projects and publications, and identifying the license for the data. This connectivity between research data and other elements in research ecosystems is required in order to make the data available and reusable beyond the initial research. In this paper, we examine these capabilities in the domains of veillance and social computing. The dataset cases presented in this paper articulate the challenges that researchers face as they seek to expose data created as a result of research activities.

Keywords: data re-use; architectures; mobile computing; communication; internet; research ecosystems; big data; research funding; veillance; social computing; technology

I. INTRODUCTION

Researchers working in the domains of Veillance and social computing are faced with the unenviable challenge of data management often beset with ambiguity of ownership and licensing conditions. The dataset cases presented in this paper exemplify the diversity of data types, locations, licensing conditions and ownership contention that often govern re-use of data.

The term Veillance [1] is used here to describe the domain within which surveillance [2], sousveillance [3], dataveillance [4] and the emergent concept of Uberveillance [5] are included. Social computing [6] is used in the context of this paper to describe the applied use of technologies that share a high degree of community formation and user level content creation. The new data-driven scientific era is highly focused on investigating research questions that requires datasets beyond a simple spreadsheet.

As researchers conduct research activities to position data as a valid, citable research output the eScience concept of identity awareness for research data [7] may inform researchers working in these domains to address key capabilities for data reuse that include:

- Resolving an identifier for data or contributors to the original entity
- Connecting identifiers for the same object or contributor across different services
- Identifying all the publications related to a particular dataset and vice versa
- Identifying all datasets from a contributor and vice versa
- Identifying the license for a dataset
- Parsing an entire identity chain, connecting data, people and publications as an interoperable and seamless service.

Jim Grey describes today's science as the fourth paradigm [8] whereas a thousand years ago science was empirical, in the last few hundred years theoretical, over the last few decades computational and today data exploration giving birth to the term 'eScience' where exploiting data is the key to scientific discovery. In recent years cross-disciplinary research projects have emerged which require collaboration of researchers from differing backgrounds, from humanities to physics or engineering, using datasets often scattered amongst multiple institutions and across continents.

In October 2010, the 'Final report of the High Level Expert Group on Scientific Data: A submission to the European Commission' entitled '*Riding the Wave*' aimed to answer one question: "How can Europe gain from the rising tide of scientific data?" [9]. The report describes the rising

significance of scientific data as an opportunity for governments and research institutions to provide researchers with a proper research e-infrastructure in order for researchers to collaborate on the same dataset from different domains, sharing a dataset across the globe.

Such collaborations can provide insights into some of the most difficult questions of our time, such as social and economic impact of climate change. The *Riding The Wave* report expresses the vision of a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data. In a sense, the physical and technical infrastructure becomes invisible and data themselves become the infrastructure – a valuable asset, with which science, technology, the economy and society can advance. Similar research infrastructure has been envisaged in Australia as part of the \$542m government investment in the National Collaborative Research Infrastructure Strategy¹ (NCRIS).

These investments have led to the development of a number of research data capture, storage and dissemination infrastructure projects in Australia. It is foreseeable that the systematic approach to curation and dissemination of research data will soon be a requirement for all publicly funded research projects. In this paper we extend the concept of identity awareness to include the access right (licence) for the research data and how this concept can be used to examine potential re-usability of research data in the domains of Veillance and social computing.

Particularly, we focus on the nature of the data in social computing and the impact of domain specific limitations that prevent data re-use, access to data, the issues posed by the heterogeneity of data, privacy and the importance of infrastructure to support open access. The scope of limited studies in this paper draws upon developments in the open knowledge, open access agenda, open data strategy, digital open government, open social data, transnational research data alliances and data publishing [10].

II. IDENTITY AWARENESS OF RESEARCH DATA

In this section we outline the key elements that are considered critical to a research data ecosystem and then proceed to define the concept of identity awareness of research data.

The following key elements are considered critical to a research data ecosystem:

- *Dataset* - research data that has been used for academic publication²
- *Researcher* - The person or research group/s who has contributed to collecting or creating the data or used this data as part of an academic publication
- *Publication* - An academic document that reports on a research activity that has used the dataset as an input or an output

¹ <http://ncris.innovation.gov.au/>

² A research data repository may hold more than one dataset while a dataset is composed of multiple data elements about physical or digital objects

- *Grant* - The research project or research program that has funded or supported creating or collecting the research data
- *Licence* - The access rights to the dataset

In addition, differing contexts and contributing factors, which govern whether data can be accessed or considered for re-use are also considered.

There are four key components to identity awareness [11] in the context of research data:

1. Identifying the researcher (recipients) who have contributed to or who have used a dataset enabling researchers to receive credit for their effort upon curation, encouraging data re-use and fostering collaboration by connecting collaborators to the researcher
2. Identifying all publications (dissemination) related to a particular dataset derived by data citation, which in turn enables reproducing the scientific results and provides credibility to research data and scientific discoveries
3. Identifying the project (activity) that funded the dataset which facilitates the reuse of data by providing information about the research activity and demonstrates the reuse of data as the return of investment for publicly funded projects
4. Identifying the licence (condition) access rights for the dataset that is vital for enabling the reuse of research data.

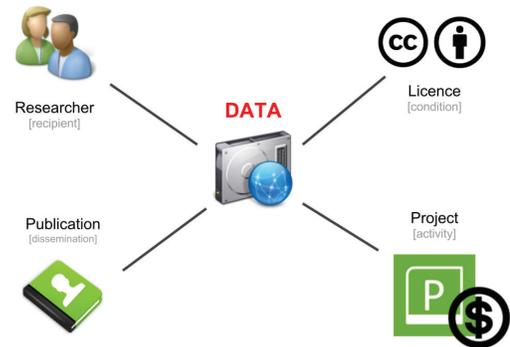


Figure 1. The Identity Awareness of Research Data - A diagrammatic representation of the vital components with which the identity awareness of research data can be determined.

III. DATASET CASES – RE-USE CHALLENGES

Using the afore mentioned four key components with which identity awareness of research can be determined, a methodological approach was devised to analyse a number of datasets from transdisciplinary research areas including astronomy, social sciences, engineering and education.

It is important to note that this methodological approach, where considering data as the central catalyst with which to consider re-use potential, identified possible configurations or conditions of access to data that contribute to re-use challenges including:

- Data accessibility - closed and accessible with permission or; closed and not accessible or; secured and personal or; open and accessible but proprietary or; open and accessible in open format
- Researcher/creator of the data - identifiable but not contactable or; identifiable and contactable or; not identifiable or; multiple creators
- Data publication status - repository with persistent identifier eg. DOI or; data paper linked to the data and resolvable or; data paper not linked to the data or; no data paper other publication type or; not published
- Data license type - total copyright or; total copyleft or; other personal provenance or; creative commons or; other proprietary
- Data funding source - (public funding or activity such as grants, support, inkind) or; no public funding (private/other) or; partial public funding (mixed source) or; 100% publicly funded or; unknown

1) The MACHO Project

The MACHO Project³ was a collaborative project between scientists from the Australian Mt. Stromlo and Siding Springs Observatory, which is a part of the Australian National University (ANU) in Canberra, Centre for Particle Astrophysics at the Santa Barbara, San Diego, the Berkley campuses of the University of California and the Lawrence Livermore National Laboratory.

TABLE I. MACHO PROJECT

Dataset 1	Details
Description	Astronomy database
Researcher/s	Identifiable
URL	http://macho.anu.edu.au/
Access	Public; open; no login required
License	Public domain
DOIs	DOI minting nearing completion
Type	Digital: ~128,000 high resolution images
Publication	81 academic publications cite this data (Google Scholar as of April 2013)
Project	* US Department of Energy through the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48 * National Science Foundation, Center for Particle Astrophysics, University of California cooperative agreement AST-8809616 * Mount Stromlo, Siding Spring Observatory, Australian National University.

The project was formed to test the hypothesis that objects such as brown dwarfs or planets make up the

³ <http://macho.anu.edu.au>

majority of the dark matter that exists in the halo of our own galaxy. These MACHOs (Massive Compact Halo Objects) can be detected by observing the effect of their gravity on the light from extra-galactic stars.

In this search for Massive Compact Halo Objects (MACHOs), known dense star fields in neighbouring galaxies were repeatedly photographed, when visible, for a period of over 10 years. Each observation consisted of simultaneously taking images, eight through a red filter and eight through a blue filter. Both red and blue images for a given observation cover the exact same region of sky at the exact same point in time. The dataset consists of a catalogue of 76 million stars, several hundred observations of each on average. The total catalogue data (stars plus lightcurves) is about 12TB.

The dataset creation precedes digital object identifiers (DOI) however separate DOIs are currently being assigned to the overall images dataset, the star catalog and the photometry catalog. The project website contains an "Acknowledgement to use when referencing MACHO project data" used to determine via Google Scholar when the dataset has been cited.

In summary, The MACHO Project is an eScience big data dataset, an example that meets all four of the criteria considered critical for identity awareness of research data. Data is available via an open public domain license, the process of DOI minting has begun on this project to ensure greater tracking of exposure and re-use of the data and the project continues to be used by numerous researchers and cited in research papers, conference proceedings and eResearch projects.

2) Glogger.mobi Project

This data repository evolved from Mann's "Wearable Wireless Webcam" (1994) that was the first cyberlog or glogging service freely made available for public use.

This digital repository provides users of lifelogging and other technologies to submit, store and share their data online. The popularity of this open Internet accessible website grew until substantially and on February 22, 1995 Mann's Wearable Wireless Webcam became "Cool Site of the Day" [12] Joi Ito's chronology of "moblogging" (mobile logging) credits Mann as the inventor (originator) of the moblogging [13] Wearable Wireless Webcam, itself deriving from Mann's earlier CyborgLog⁴ which is a first-person recording of various physiological and environmental sensors in an activity, in which the person doing the recording is a participant in the activity. [14]

TABLE II. GLOGGER.MOBI

Dataset 2	Details
Description	Mobile glogging repository
Researcher/s	Identifiable
URL	http://glogger.mobi
Access	Login required – data otherwise open

⁴ Often abbreviated to 'glog'

Dataset 2	Details
License	User assigned – CC BY, GFDL, other
DOIs	None
Type	Digital: images, videos, audio, metadata
Publication	Numerous publications
Project	Privately funded

Glogger.mobi was created by Professor Steve Mann's team at University of Toronto and has 246,282 users⁵ and many terabytes of data in the form of images, audio, video, autosensory (self-sensing), physiological, sousveillance, as well as related metadata.

CyborGlogging pre-dates blogging (web logging), and differs from blogging/logging in that the cyborglog is created without conscious thought or effort by the person doing it [15]. EXIF and associated geospatial metadata is exposed with each data element, content syndication is available using RSS and downloads of data are offered in FLV, MPEG4 and the original high definition format.

In summary, Glogger.mobi Project meets 4 of the 4 criteria considered critical for identity awareness of research data, however, due to its iterative and dynamic state Glogger.mobi remains a contentious case as it no formal data repository status despite the fact that it is cited in numerous research journals, conference proceedings and e-publishing articles.

3) Digital Outback Project

The former Australian Flexible Learning Framework, [16] an Australian federal government initiative awarded funding to successful applicants from the private, government and industry based vocational training sector to embed the use of e-learning and emergent technologies in an education and training context.

Under the auspice of 'E-learning for Target Groups' this project engaged Australian Aboriginal learners throughout rural and remote communities in Western Australia.

TABLE III. DIGITAL OUTBACK PROJECT

Dataset 3	Details
Description	Mobile blogging data project
Researcher/s	Identifiable
URL	http://moblog.net/mobdeadly
Access	Public; open; no login required
License	Mixed – stakeholder contention
Data Type	Digital - images, geolocation and other text based metadata
DOIs	None
Publications	Multiple Australian government reports & conference proceedings
Project	* Australian Flexible Learning Framework (AFLF) E-learning for Target Groups Grant Funding * Parliament and Civic Education

Dataset 3	Details
	Rebate (PACER) Program * Telstra, Countrywide Australia

Supported by the telco Telstra Australia, a number of mobile technologies including a O2 XDA PDA and camera enabled Nokia 6000 series mobile phones were taken into the West Australian remote communities of Parngurr, Punmu and Jigalong communities for use in educational and community activities.

This project sought to engage learners in online digital story telling activities that brought them into contact with international communities. This project successfully demonstrated the use, for educational purposes, of mobile blogging, or moblogging of digital images to remote servers using email GPRS. Due to cultural sensitivities around visual and audio digital data capture, the ownership of data in this cultural context was deemed as shared amongst the Aboriginal community. Strict permissions to publish were negotiated in the form of image watermarking for select digital artefacts.

The data remains accessible, however, data provenance and intellectual property rights are not clearly defined due to differing cultural understandings of the Parngurr community, the Australian Federal Government and various researchers involved in the research activity that generated the data.

In summary, the Digital Outback Project meets 2 of the 4 criteria considered critical for identity awareness of research data. A mixed license type state (due to multiple and differing investments from key stakeholders) affects data re-use. Despite the fact that the project has been cited in numerous research papers, conference proceedings and Australian Federal government reports the dataset has no formal citation or recognized acknowledgement statement.

4) The Biomass Heat Transfer Project

The Biomass Heat Transfer Project facilitated by researcher Leigh Blackall is openly accessible via the Internet to research data that is dynamically authored and made available for re-use.

Blackall cites the emergent concept of Networked Learning [17] as the premise by which a body of digital artefacts including images, video, audio, geolocation and related metadata is created, facilitated by the Internet. The 'Biodynamic Heat Transfer' project is an ongoing data collection that involves multiple researchers (contributors) dynamically updating the collection including text based conversational contributions that further the cause of open education resources (OER) [18].

TABLE IV. BIOMASS HEAT TRANSFER PROJECT

Dataset 4	Details
Description	Networked learning repository – mixed contribution relating to the development of biomass heat transfer.
Researcher/s	Identifiable
URL	http://goo.gl/10KFV

⁵ Statistics as of 14th February 2013

Dataset 4	Details
Access	Public; open; no login required
License	Mixed - CC BY SA & Public Domain and All Rights Reserved
DOIs	None
Data Type	Digital: text, video and images; document edit histories
Publications	Author considers text, data, and reviews as published (unrecognized)
Project	No funding attribution noted

The focus of this project according to Blackall is to demonstrate an emergent learning and accreditation process, one in which multiple contributors dynamically update a multiple format dataset distributed across multiple digital platforms, connected via hyperlinking and other syndication services such as RSS and XML.

According to the researcher this dataset challenges traditional scholarly data publication as to the near future:

- Trajectory of dynamic data publication;
- Convivial nature of collaborative peer review;
- Emergent case for hybrid data provenance;
- Informal and peer oriented academic accreditation;
- Shift from expert centred to participatory, multi-author data creation and data reuse

In summary, the Biomass Heat Transfer Project meets 2 of the 4 criteria considered critical for identity awareness of research data. The data is available via a number of open web services under a range of license conditions however, due to its iterative and dynamic state this dataset remains a contentious case with limited citations existing in formal research papers, journals, conference proceedings or associated literature. Blackall considers the data as published, although this premise is only partially acknowledged in established research communities.

5) *TalkingVTE Project*

The TalkingVTE Project dataset contains the archived podcast audio recordings made at annual Australian Flexible Learning Framework (AFLF) eLearning showcase events from 2006 through to 2011.

These audio recordings provide a rich cross section of innovation in Australian e-learning and mobile learning conducted throughout the vocational educational and training sector. Most of the recordings were conducted in studio sessions, conference recordings or vox-pop interviews with educational leaders working in the areas of workforce development, workplace training, innovation and wearable computing.

TABLE V. TALKING VTE PROJECT

Dataset 5	Details
Description	Event audio podcasting
Researcher/s	Identifiable
URL	http://talkingvte.blogspot.com

Dataset 5	Details
Access	Login not required – RSS, Atom
License	CC BY Share Alike 2.5
DOIs	None
Data Type	Digital: audio recordings and accompanying metadata
Publications	No research publications
Project	Australian Flexible Learning Framework (AFLF)

The primary podcast utilizes a Google Blogger blog to syndicate to iTunes® via Google Feedburner. The audio data is made available under a Creative Commons Attribution-Non-Commercial-Share-Alike-3.0 license. The audio collection has survived the numerous funding iterations of the Australian Flexible Learning Framework by virtue of the efforts of the primary researcher Stephan Ridgway who continues to maintain the collection under a personal account profile.

In summary, the TalkingVTE Project meets 2 of the 4 criteria considered critical for identity awareness of research data. Whilst this dataset has limited research publication status the media contents, according to the researcher, are regularly accessed for download. Due to the archived state of the data, no further consideration has been made for DOI minting or further data publication, subsequently the dataset case will have no formal citation or recognized acknowledgement statement.

IV. ADVANTAGES OF DATA PUBLICATION

According to Pily and Fukasuka [19], peer-review and subsequent publishing of data has benefits for more than just the data scientists who create the datasets. As the previous datasets evidence, publishing data also benefits the funding bodies that pay for the data to be collected as well as the wider academic community.

It is also evident that twenty-first century research must make data dissemination a regular part of scholarly work flows. Better data dissemination can promote analytic rigor and transparency; reduce inefficiencies and duplication of effort, and open new research opportunities for larger scale and multidisciplinary inquiry. This need pertains to many areas of the humanities, social sciences and natural sciences. [20]

In the last two years data journals have evolved in a number of forms that differ markedly from database portals, dataset services or publications that only describe where a dataset may be accessed. [21] The primary purpose of data journals is to expose datasets and make them accessible for other researchers to reuse. Data journals also facilitate researcher activity that ties safe and trustworthy repositories for the preservation of underlying research data to the publication. For findability, citability and re-use, it is of ultimate importance that researcher validated data are well linked with publications to ensure their inclusion in the record of science. [22]

In the research domain of Veillance and that of social computing the challenge is not only how to bring about identity awareness of research data, but also how to support initiatives that guarantee true personal information asset management which in turn ensures that widespread permissions and access to the data can be managed while protecting individual privacy and security. [23]

V. THE IMPORTANCE OF DATA IN RESEARCH

Traditionally researchers have been at the mercy of the 'publish or perish' yardstick. Catch cries of 'show me the data' or 'data as gold' suggest a growing awareness of the priority and importance of data in the research community and the need for publishers to support data publication.

Datasets have several important qualities that differ from manuscripts as datasets can often be quite large (scale) and full of complex interrelationships between various tables and multimedia files (image, videos, GIS etc.) [20] There is multitude of proprietary and open format data softwares required to access, transfer, resolve or visualise such data.

Paradoxically, governments who institute policy to ensure quality assurance of data management and subsequent publication are now challenged also to take and apply the lead from galleries, libraries, archives museums and research organisations to manage rich media data content across whole of government. [24] The rapid shift to smart infrastructures that depend on Internet connectivity has profoundly altered the data landscape for researchers and access to these repositories has subsequently become easier to handle big data. As this paper attests, data repositories, data publishing, advanced citation and curation techniques facilitate re-use of data in turn lowering access barriers and avoiding hours of wasted research activity.

With a push for greater financial scrutiny and accountability, researchers who would have once operated in obscurity are now subject to a model of open contribution. This is particularly the case in the Veillance domain and that of social computing as in some cases the data itself is a result of activities that are based upon recordings of activities using technologies bound by ethical constraints and privacy considerations. Identity awareness of research data is undeniably first and foremost concept used extensively to guide research activities in the eSciences, however it may transfer neatly to the research activities in the domains of Veillance and social computing.

VI. DILEMMA IN VEILLANCE AND SOCIAL COMPUTING

The explosion of emergent technologies in a research context provides new challenges for researchers and organisations alike in the collection, collation and curation of data.

As we usher in a new era of interdisciplinary, transdisciplinary and collaborative data oriented research activities, researchers need support from their research organizations in their quest to manage data storage, access and re-use. By re-envisioning eResearch infrastructures there is a need for more than just discipline oriented or purpose specific secure research infrastructures to address the requirements of data intensive research activities.[25]

Of great concern are organisations that purposefully lock down access to data generated as part of research activities in the domains of Veillance and social computing, challenged and misinformed of the intent underpinning the open data agenda. The dilemma for researchers in these domains is the misconception that the agenda is driven solely by funding mandates for scrutiny and exposes that which had previously remained under lock and key. This mistrust in turn affects the agile methodological approach that researchers rely on to expedite research in a timely and competitive manner.

By virtue of providing open access to research data, reuse is now shaping whole communities, where researchers are now at the forefront of actionable change, not simply librarianship without a mantle. The dilemma in the domains of Veillance and social computing can be resolved by organisations carefully engendering trust, by engaging key stakeholders through research communities in activities that dispel myths and provide greater insight into the identity awareness of research data. This awareness is particularly important where person or human subject related information is an integral part of those research activities undertaken.

Managing person related information has become a new hot currency and an asset in fighting for the governance of the next phase of the Internet. Many leading data gatherers do massive business on personal data. Rights on the vast amounts of person related data are practically in the hands of data vendors and governments: they own this data, may make profit on it or provide personal data for a basis of decision-making. [26] It is evident that researchers supported by their organisations and who act out of an informed state of identity awareness of research data, make best use of data without compromising the subject in the process.

CONCLUSION

This paper provides evidence of the challenges that are faced by researchers in the domains of Veillance and social computing as they consider the identity awareness of research data in an endeavor for the re-use of research data.

This paper also provides an account of the key components and conditions that guide data intensive research activities in the domains of Veillance and social computing in an effort to meet to varying degrees the requirements of identity awareness of research data.

Clearly articulated connections between data, researchers, publications and research grants in this paper provide the impetus for researchers to strategise research activities to position themselves competitively in contemporary research ecosystems. A significant finding in this paper is that researchers have an opportunity to conduct research activities to position data as a valid and citable research output.

This paper also identifies barriers that are faced by researchers seeking to re-use datasets in the domains of Veillance and social computing, a need for a unified and co-ordinated international approach to identity awareness of research data and the ongoing challenges for researchers as they deal with privacy, security and well being of personally identifiable data [27] that emerges from research activities.

REFERENCES

- [1] E. Ackermann, "Why Smart Glasses Might Not Make You Smarter," *IEEE Spectrum*, pp. 1–6, 2013.
- [2] D. Lyon, "Surveillance Society: Monitoring Everyday Life," Buckingham & Philadelphia: Open University Press, 2001, pp. 1–7. DOI: 10.1080/01972240309486
- [3] S. Mann, J. Nolan, and B. Wellman, "Sousveillance: Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments," *Surveillance & Society*, vol. 1, no. 3, pp. 331–355, 2002.
- [4] R. A. Clarke, "Information Technology and Dataveillance," *Communications of the ACM*, vol. 37, no. 5, pp. 498–512, 1988.
- [5] K. Michael and R. Clarke, "Location and Tracking of Mobile Devices: Überveillance Stalks the Streets," *Computer Law and Security Review*, no. October, p. 30, 2013.
- [6] M. Parameswaran and A. Whinston, "Research issues in social computing," *Journal of the Association for Information Systems (JAIS)*, vol. 8, no. 6, pp. 336–350, 2007.
- [7] A. Aryani and A. Burton "Identity Awareness: Toward an Invisible e-Infrastructure for Identifying," eResearch Australasia, 2012, Sydney, Australia. DOI: [10.6084/m9.figshare.155650](https://doi.org/10.6084/m9.figshare.155650)
- [8] Tony Hey, Stewart Tansley, and Kristin Tolle, "The fourth paradigm: data-intensive scientific discovery," 1st ed.: Microsoft Research, 2009.
- [9] J. Wood, T. Andersson, A. Bachem, and C. Best, "Riding the wave: How Europe can gain from the rising tide of scientific data. Final report of the High Level Expert Group on Scientific Data - European Commission.," European Union, 2010. [Online] Available: <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf> [Accessed: 10-May-2013].
- [10] M. Napolitano, "Open Data Challenges and Opportunities," Open Knowledge Foundation, pp. 1–54, 2012.
- [11] A. Hayes and A. Aryani, "The Identity Awareness of Research Data," 2013. [Online]. Available: http://figshare.com/articles/Identity_of_Research_Data/_157165. [Accessed: 26-Mar-2013].
- [12] S. Mann, "Wearcam.org as roving reporter"; (c) Steve Mann, Feb. 1995," 1995. [Online]. Available: <http://wearcam.org/eastcampusfire.htm>. [Accessed: 17-Feb-2013].
- [13] J. Ito, "Joi Ito's Web: A chronology of articles events and resources," 2003. [Online]. Available: <http://radio-weblogs.com/0114939/outlines/moblog.html>. [Accessed: 17-Feb-2013].
- [14] S. Mann, "Mann: About Glogging," 2013. [Online]. Available: <http://www.glogger.mobi/about.php#glogger>. [Accessed: 17-Feb-2013].
- [15] S. Mann, *Intelligent Image Processing*. John Wiley and Sons, 2001, p. 384.
- [16] "Australian National Vocational Elearning Strategy (NVELS)," 2013. [Online]. Available: <http://www.flexiblelearning.net.au>. [Accessed: 17-Feb-2013].
- [17] L. Blackall, "What is a definition of networked learning?," *Blogger*, 2011. [Online]. Available: <http://www.leighblackall.com/2011/06/what-is-definition-of-networked.html>. [Accessed: 1-May-2013].
- [18] L. Blackall, "Networked Learning: The Biomass Heat Transfer Project," *Alexander Hayes: Archive.org*, 2013. [Online]. Available: <http://archive.org/details/NetworkedLearningBiomassHeatTransferProject>. [Accessed 8-May-2013].
- [19] D. Pilat and Y. Fukasaku, "OECD principles and guidelines for access to research data from public funding," *Data Science Journal*, pp. 1–24, 2007.
- [20] "Open Context Data Publishing: Editorial Policies & Author Guidelines," *Open Context - Web based research data publication*, vol. 1. The Alexandria Archive Institute, pp. 0–10, 2012.
- [21] "F1000 Research Provides Researchers with Fast and Transparent Science Publishing Model," F1000, pp. 0–3, April 2013. [Online] Available: <http://f1000.com/resources/Open-Science-Announcement.pdf>. [12-May-2013].
- [22] A. Farquhar, "Joint Statement from DataCite and the International Association of STM Publishers on the Linkability and Citability of Research Data," no. June. DataCite, Copenhagen, pp. 4–5, 2012.
- [23] B. Mulcronney, "For the People, By the People: A New Deal for Personal Information Assets," 2013. [Online]. Available:

<http://wethedata.org/2013/02/04/a-new-deal-for-personal-information-assets/>. [Accessed: 17-Feb-2013].

- [24] P. Waugh and A. Stringer, "My NZ Open Data & Digital Government Adventure," 2013. [Online]. Available: <http://pipka.org/2012/12/18/my-nz-open-data-and-digital-government-adventure/>. [Accessed: 17-March-2013].
- [25] W. Michener, "Five New Paradigms for Science and an Introduction to DataONE," *Educause*, vol. 47, no. 2, pp. 50–51, 2012.
- [26] K. Ekholm and P. Karhula, "Sleepwalking toward a control society: Ten Must-Know Trends," *www.ifla.org*, pp. 1-15, 2012.
- [27] "Data Protection Directive: European Union," 2012. [Online]. Available: http://en.wikipedia.org/wiki/Directive_95/46/EC_on_the_protection_of_personal_data. [Accessed: 1-May-2013].