

Information Provenance

Prof. Simon Rogerson

**Originally published as ETHicol in the IMIS Journal Volume 15
No 1 (February 2005)**

In January 2005 the UK public acquired five important new rights to information held by public authorities.

- *The Freedom of Information Act 2000* comes into force, after a 4 year delay to give authorities time to prepare. The Act applies to central government bodies and to English, Welsh and Northern Ireland public authorities. It also applies to the House of Commons, the House of Lords and to the Welsh and Northern Ireland assemblies.
- *The Freedom of Information (Scotland) Act 2002* applies to the Scottish Executive, the Scottish Parliament and Scottish public authorities
- *The Environmental Information Regulations 2004* provide a separate right of access to environmental information held by UK public authorities. Some private bodies, including utilities and contractors providing environmental services on behalf of authorities, are also covered. The regulations implement an EU directive.
- *The Environmental Information (Scotland) Regulations 2004* provide a similar right of access to environmental information held by Scottish public authorities and certain private bodies.
- *Amendments to the Data Protection Act 1998* strengthen people's rights to see personal information about themselves held by public authorities throughout the UK, including Scotland. The Act already allows people to see computerised personal data about themselves and medical, social work, housing and school records. The amendments significantly improve the right to see other paper records.

(source: The Campaign for Freedom of Information, Press release: 31 December 2004, www.cfoi.org.uk/foi311204pr.html (<http://www.cfoi.org.uk/foi311204pr.html>))

These rights give a new impetus to raising awareness of the need for information integrity and the implications of the lack of it. Information integrity is about accuracy, consistency and reliability of information content and information systems. Madhavan Nayar (26 July 2004) explains that, "Digital information is becoming as pervasive and essential as air, water, electricity and canned food. Increasingly, we rely on such information for our livelihood, lifestyle and even life itself. Ironically, however, information has not been the focus of interest thus far in this information age." If such information is questionable

then decisions and actions which are based upon it could be flawed and unsafe. How many situations like this will come to light when people gain access to information using their new rights? How many situations like this remain hidden in the organisations not covered by the new legislation? The expectation that digital information is dependable and trustworthy is reasonable. But how can dependability and trustworthiness be demonstrated? Trustworthiness is an intrinsic reality. Its perception, particularly in the beginning, depends critically on the perception of certain extrinsic forms (signs, labels, messages, etc) that are understood to represent the presence of underlying trustworthiness. In the case of digital information these extrinsic forms should represent the Information Provenance. Information provenance fixes the origin and network of ownership thus providing a measure of integrity, authenticity and trustworthiness. It provides an audit trail showing where information originated, where it has been and how it has been altered. In this way people would be able to consider how much credence they would give to a piece of information before acting upon it.

Consider this example. In the course of its enquiries a police authority collects information about an individual. This information is held within the police authority's information systems. Such information is allowed to be shared with a number of other authorised agencies across a secure network. Access is instigated by the agencies so no track is kept of where the information has been shared. Once this happens the copies of this information become legally owned by the recipient agencies. Agencies update this information for their own purposes and based upon their own intelligence. These new versions of the information are passed onto other authorised agencies. The police authority then updates the information about the individual based on new evidence. Agencies are not aware of this and continue to use their own version of the information.

In this situation there exist multiple copies of the information across a complex network of agencies. Copies are not the same and there is no mechanism in place to ensure that they are the same. Clearly the integrity of the information is questionable but those receiving it are likely to be unaware of this. Decisions may be made based on this untrustworthy information, that have detrimental effects on the individual. If the information had been accompanied by the information provenance then decision makers would be able to see how the information had changed and therefore consider how safe it was. Also provenance would provide a method to track back to the provenance of original information held in the information systems of the police authority to check whether the original information had altered since it was first accessed.

From this example which is based upon a real situation it can be seen the information provenance is a powerful instrument in improving information integrity and trustworthiness. According to Fox and Huang (2003) there are four levels associated with information provenance

- *Level 1 Static* This focuses on provenance of static and certain information
- *Level 2 Dynamic* This considers how the validity of information may change over time
- *Level 3 Uncertain* This considers information whose validity is inherently uncertain
- *Level 4 Judgment-based* This focuses on social processes necessary to support provenance.

If such concepts of provenance were incorporated into the practices and processes around information then the management and governance of information would be much improved. Just think - if every piece of information was accompanied by its provenance what a different information society we would live in.

Please send your views on ethical and social responsibility issues and cases of ethical dilemmas to:

Professor Simon Rogerson
Director
Centre for Computing and Social Responsibility
Faculty of Computing Sciences and Engineering
De Montfort University
The Gateway
Leicester
LE1 9BH
Tel:(+44) 116 257 7475
Fax:(+44) 116 207 8159
Email:<srog@dmu.ac.uk>
Home Page: (<http://www.ccsr.cse.dmu.ac.uk>)