

Grey in the Innovation Process

Keith G Jeffery STFC, Anne Asserson UiB

Abstract

The research lifecycle has multiple objectives materialised as outputs, outcomes and impacts. Typical outputs are research publications (including grey literature), patents and products such as research datasets and software, many kinds of art or prototype engineering artifacts. Outcomes include patent licence income, value of a company set up to exploit the output or trained research staff. Impacts include employment creation, a commercial product that saves lives or labour or development of a new field of knowledge and research such as genomics since the 1950s.

Commonly research in progress may be documented as grey literature – such as technical reports, laboratory notebooks or instructions for operating new equipment. There is a decision point when grey literature is produced.

One can innovate academically. The output is peer reviewed publications; the outcomes include developing trained researchers; the impact leading to a new field of research. This route provides academic recognition.

Alternatively one can innovate along the wealth-creation route. The output could be a patent; the outcome license income or a new company; the impact employment, dividends to shareholders or a new 'wonder product'. This route provides wealth and possibly improvement in the quality of life.

If research is published this usually precludes following the wealth-creation route since the novel idea is now in the public domain and not protected by patent(s).

Increasingly research funding and research performing organisations wish to demonstrate that the research they fund or do leads to impacts of relevance to society. Tracing of impacts back to the original research is not easy, partly because the eventual impact may not be known for many years. The key is an accurate recording of the research lifecycle including important dates so that the innovation cycle from idea to impact and back to further ideas can be demonstrated.

Recent work - especially in UK in the JISC-funded MICE project - has produced a taxonomy of outputs, outcomes and impacts. In parallel an extension to CERIF (Common European Research Information Format - an EU recommendation to member states) has been developed and approved by euroCRIS. This extension re-uses typical CERIF entities of persons, projects, organisations, publications, patents, products but relates them (with temporal validity and appropriate role) to the

production or utilisation of outputs, outcomes and impacts. Naturally grey literature is a key component within this model.

1 INNOVATION

Innovation may be defined as the development of new customers' value through solutions that meet new needs, inarticulate needs, or old customer and market needs in new ways (Wikipedia). There are essentially three kinds of innovation:

1. Academic innovation: leading to new research techniques and results and providing also trained researchers;
2. Commercial innovation: using the results of research to create wealth via patents or products taken up by industry to provide a commercial product or service that generates the wealth;
3. Societal innovation: using the results of research for improvement in the quality of life in environmental, health, cultural or social aspects.

Innovation is a process with a time dimension and along that time dimension are produced outputs, outcomes and impact.

The grey literature output of research sits within a much broader context of other outputs (such as peer-reviewed publications, patents, products). Nonetheless, as argued in (Jeffery and Asserson 2007) grey literature objects are a very important

output since they may lead to outcomes and impact. The outputs are not the innovation: this is achieved by utilising the outputs to produce outcomes and impact.

Outcomes are activities derived directly from the outputs: from research publications or grey literature one might have as outputs trained researchers and new research techniques. From patents (themselves considered grey literature) an outcome could be license income and possibly consultancy work (income) for the researchers to assist exploitation of the patent. From products an outcome could be the setting up of a spin-out company which employs people and produces products or services. Interestingly, many of the outcomes of research are documented by grey literature since patents, company technical reports, government internal (or external) reports are generated and these are not peer reviewed 'white literature'.

Impact is difficult to define but it is the effect the research has on society. Commonly impact is detected many years (commonly 10-15) after the research is completed.

Some examples may illustrate:

1. An impact may be millions of lives saved due to a drug made available after extensive trialling, produced by a pharmaceutical company after further in house development and based on an output of research at a university.
2. Impact may be due to a policy change - possibly enforced by law - based on research. An example is the reduced deaths from lung cancer in Western

society because of laws based on policies derived from research on the effect of tobacco on human health.

3. Similar examples exist in the environmental domain where research has led – in time – to policies concerning the provision of clean water supplies.

Of course in some of these cases there is associated wealth creation (e.g. for the pharmaceutical company in the first example) and associated provision of employment and hence further wealth creation.

2 BACKGROUND

2.1 Previous Work

For more than two decades, the authors have worked on research information in the widest sense comprising information not only about grey literature (grey objects) but also all the outputs of research (products, patents, publications) and the context within which the research was done including projects, organizations, funding, persons, facilities, equipment, events. Within the GL community we have highlighted the issues as we see them:

1. the need for formal metadata to allow machine understanding and therefore scalable operations (Jeffery 1999);
2. the enhancement of repositories of grey (and other) e-publications by linking with CRIS (Current Research Information Systems) (Jeffery and Asserson 2004);

3. the use of the research process to collect metadata incrementally reducing the threshold barrier for end-users and improving quality in an ambient GRIDs environment (Jeffery and Asserson 2005);
4. an architectural model for scalable, highly distributed, workflowed repositories of grey literature based on hyperactive 'intelligent' documents (Jeffery and Asserson 2006).
5. A 'from 10,000 metres altitude' view of the grey information landscape 'Greyscape' based on the hypothesis that grey literature is the foundation for the knowledge economy (Jeffery and Asserson 2007).
6. An analysis of interoperation architectures among research information systems 'INTEREST' (Jeffery and Asserson 2008).
7. A proposal that Grey Literature should be seen within the context of e-Science supported by a CERIF-CRIS (Jeffery and Asserson 2009).
8. A proposed architecture 'GLASS' using CERIF metadata to demonstrate transparency in the Grey process (Jeffery and Asserson 2010).

Although this corpus of work demonstrates how CERIF provides the required context for processing grey objects, in particular the 'Greyscape' paper (Jeffery and Asserson 2007) related Grey Literature described by CERIF metadata to the knowledge economy and is thus a relevant piece of previous work related to the current topic: Innovation.

2.2 The Requirement

The requirement is to record the innovation derived from research and in the context of this community research recorded as grey literature (or grey objects) meaning outputs not formally peer-reviewed. The innovation is recorded as outcomes and impact. However, the outputs, outcomes and impact need to be related back to the research project, the persons involved, the organisations involved (e.g. university / faculty / department / research group or maybe a commercial company), the funding and funding organisation (including where appropriate research programme and topic), the facilities and equipment used in the research.

3 THE HYPOTHESIS

We assert that a solution – CERIF – exists already which covers these requirements. CERIF has already been in use widely in 42 countries for recording research activity and is an EU Recommendation to Member States. CERIF is maintained, developed and promoted by euroCRIS (www.eurocris.org) at the request of the European Commission. In particular, work done during the MICE (JISC-funded) project in UK extended the CERIF datamodel (which included already outputs) to include indicators and measurements which can record outcomes and impact (and aggregated outputs) (Gartner, Cox, Jeffery 2012). This proposed extension to CERIF was ratified through the euroCRIS process for inclusion in the CERIF datamodel, and provides a

The new entities for managing indicators and their measurements for the purposes of innovation (outputs, outcomes and impact) are indicated in the diagram above but dealt with in more detail below.

4.2 Indicators and Measurements for Outcomes and Impacts

The part of the datamodel concerning indicators and impacts is reproduced in detail here (Figure 2) and in particular it should be noted how the entities representing indicators and measurements relate to the base entities of CERIF (such as publication, patent, product and person, organisational unit, project etc) through the semantically rich temporally bound linking relations.

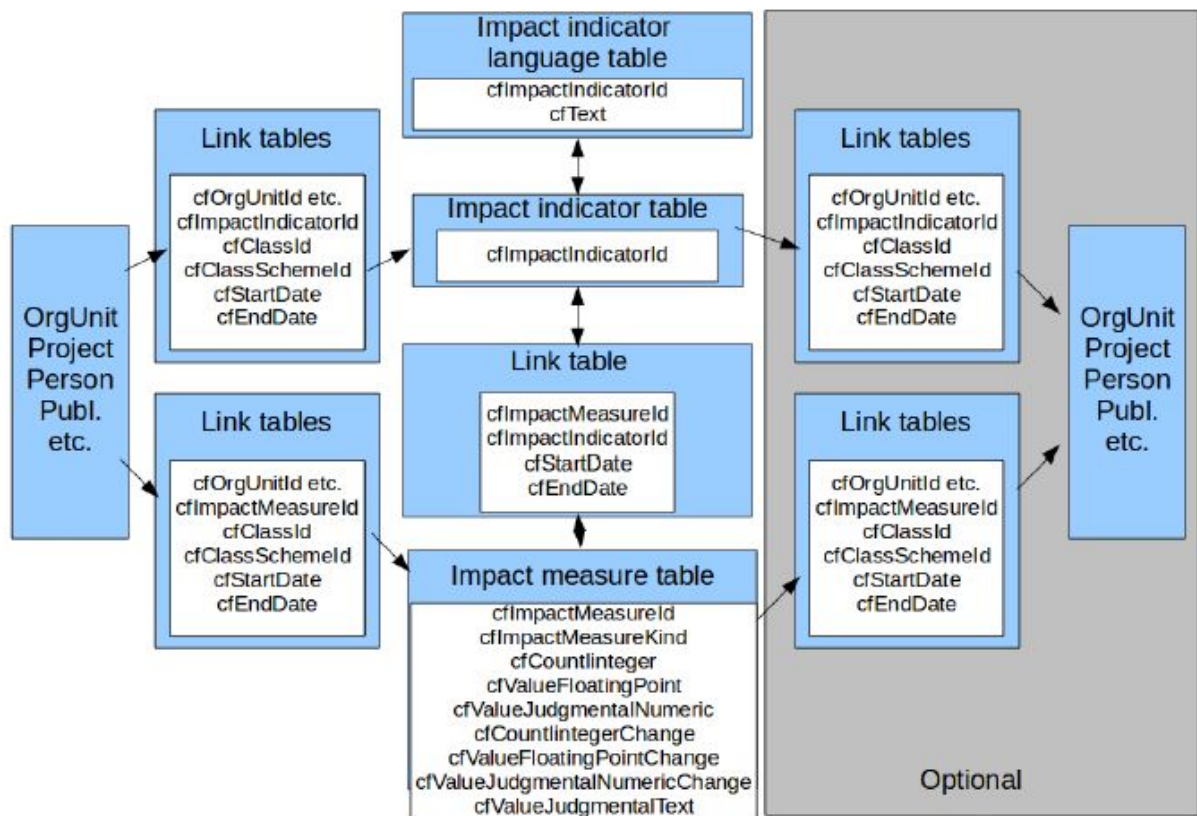


Figure 2: The MICE Datamodel within CERIF

It should be noted that the right side of the diagram (i.e. the instances of entities affected by or benefitting from the outcome or impact) is optional. Furthermore the instances of entities on the right hand side as e.g. beneficiaries of outcomes may appear on the left side as initiators of the transition from outcomes to impacts.

For each indicator there are one or more measurements. Not all measurements are relevant for every indicator; in fact usually only one measurement is appropriate for an indicator. The measurements include an integer count (e.g. how many publications were produced by a person or group in a given time period or how many lives were saved by a new drug); a floating point measurement (e.g. amount of licence income for a patent) and a judgement expressed numerically (e.g. quality on a scale of 1-10). In addition there are the 'delta' measures which record change and compare the value for one period of time with another. Examples would include the increased number of publications, the increased licence income or the improvement in quality. Finally an attribute is made available for a textual statement on judgement to justify the measure for an indicator or to express less precisely the estimated quality.

The MICE project produced a detailed taxonomy of indicators (Gartner, Cox, Jeffery 2012) that could be used, but there are others in the scientometrics and bibliometrics fields. CERIF can, of course, allow the use of any scheme of indicators due to its flexible semantic layer feature.

Another UK project, Snowball, (SnowballProject) has produced a set of indicators for university benchmarking. However, the indicators are dominantly to record performance and less to record outcomes and impact. The 'Snowball Recipe Book' (produced by Elsevier which was a project partner) (SnowballRecipes) was launched at a recent euroCRIS Members' Meeting.

The recently initiated Indicators Task Group of euroCRIS is exploring and researching the available techniques and intends to produce a canonical set of indicators and associated measurements that can be used for benchmarking and comparison across outputs, outcomes and impact. This will provide the basis for measuring innovation and especially innovation in the Grey Process.

5 CONCLUSION

From the above we may conclude:

1. CERIF provides an appropriate data structure for recording innovation, including within the GREY process;
2. It is being used in significant systems tracking outputs, outcomes, impact related to contextual, temporal, geospatial metadata;
3. euroCRIS has an Indicators Task Group dealing with new scientometrics (including bibliometrics) and new methods for detecting impact (backward chaining)

6 REFERENCES

(CERIF) www.eurocris.org/cerif

(Gartner, Cox, Jeffery 2012) Richard Gartner, Mark Cox, Keith Jeffery 'A CERIF-based schema for encoding research impact' The Electronic Library xxx.yyy 2012 Emerald Publishing (in press)

(Jeffery 1999) Jeffery, K G: 'An Architecture for Grey Literature in a R and D Context' Proceedings GL'99 (Grey Literature) Conference Washington DC October 1999 <http://www.konbib.nl/greynet/frame4.htm>

(Jeffery 2004b) Jeffery, K.G.; 'The New Technologies: can CRISs Benefit' in A Nase, G van Grootel (Eds) Proceedings CRIS2004 Conference, Leuven University Press ISBN 90 5867 3839 May 2004 pp 77-88 (available under www.eurocris.org)

(Jeffery and Asserson 2004) K G Jeffery, A G S Asserson; Relating Intellectual Property Products to the Corporate Context; Proceedings Grey Literature 6 Conference, New York, December 2004; TextRelease; ISBN 90-77484-03-5

(Jeffery and Asserson 2005) K G Jeffery, A G S Asserson 'Grey in the R&D Process'; Proceedings Grey Literature 7 Conference, Nancy, December 2005; TextRelease; ISBN 90-77484-06-X ISSN 1386-2316

(Jeffery and Asserson 2006c) Keith G Jeffery, Anne Asserson: 'Hyperactive Grey Objects' Proceedings Grey Literature 8 Conference (GL8), New Orleans, December 2006; TextRelease; ISBN 90-77484-08-6. ISSN 1386-2316 ; No. 8-06-X

(Jeffery and Asserson 2007) Keith G Jeffery, Anne Asserson: 'Greyscale' Opening Paper in Proceedings Grey Literature 9 Conference Antwerp (GL9) 10-11 December 2007 pp9-14; Textrelease, Amsterdam; ISSN 1386-2316

(Jeffery and Asserson 2008) Keith G Jeffery, Anne Asserson: 'INTEREST' Proceedings Grey Literature Conference Amsterdam 8-9 December 2008 in Tenth International Conference on Grey Literature : Designing the Grey Grid for Information Society, 8-9 December 2008, Science Park Amsterdam, The Netherlands ed. by Dominic J. Farace and Jerry Frantzen ; GreyNet, Grey Literature Network Service. - Amsterdam : TextRelease, February 2009. GL-Conference series, ISSN 1386-2316; No. 10. - ISBN 978-90-77484-11-1.

(Jeffery and Asserson 2010) Keith G Jeffery, Anne Asserson 'GL Transparency: Through a Glass, Clearly' Proceedings Twelfth International Conference on Grey Literature,

Prague December 2010 Ed Dominic Farace and Jerry Frantzen 95-100, TextRelease
Amsterdam Series ISSN 1386-2316; no 12; ISBN 978-90-77484-16-6

(SnowballProject) <http://www.snowballmetrics.com/>

(SnowballRecipes) <http://www.snowballmetrics.com/wp-content/uploads/Snowball-Metrics-Recipe-Book.pdf>

(Wikipedia) www.wikipedia.org