

# Data, Data Everywhere – open, linked, interoperable?

Jon Mason

International Graduate Centre of Education  
Charles Darwin University  
Darwin, Australia  
jon.mason@cdu.edu.au

**Abstract**—The focus of standards development in the domain of information technology for learning, education, and training (ITLET) has shifted considerably since it first gained global attention in the 1990s. While the field has benefited from and endured disruption it has achieved sufficient buy-in from stakeholders for its core aim: interoperability of systems and learning content. This has been achieved in the midst of a range of diverse trends with regards to technological innovation, pedagogical practice, political imperatives, and socio-cultural sensibility. Among these trends has been the ‘open agenda’ – a movement with deep roots that has hitherto been focused upon access, intellectual property, benefit to the public domain, data sharing, and technical interoperability. Openness has generally been perceived as beneficial for the common good – although events in recent years provide potent evidence that it can be politically divisive when confidential information is leaked and governments legislate that their intelligence agencies have access to sufficient data to keep us safe. What implications does this have for ITLET? Apart from new capabilities the proliferation of data now associated with digital technology innovation brings with it a key development: data that has traditionally been distinct from a learning resource is fast becoming a learning resource in itself as ‘learning analytics’ is deployed. With new capabilities of e-learning systems there is a new imperative emerging: systems governance. Many of the issues associated with learning analytics – privacy, ownership, ethical, and business related – can all be understood as facets of governance. So, for the standards communities engaged in ITLET, interoperability must also be considered through the lens of governance. How can this be achieved? It is proposed in this short paper that in order to properly scope what is required that questioning be embraced in a rigorous and strategic manner.

Making sense of the emergent patterns of order and disorder in all this change is what we are compelled to do – but, are we asking the right questions about information technology standards when we do this? Is it fair to assume that systems interoperability is necessarily the primary goal for ITLET? Is systems governance just a policy issue? What do we need to consider in the development and deployment of data standards? What needs to be considered in the development of open protocols?

**Keywords**—analytics; standards; tracking; governance; interoperability; open; complexity

## I. INTRODUCTION

Not long after the Web was enthusiastically embraced by the education sector – and subsequently colonized by business and financial interests – global movements emerged that were

concerned with “defining the Internet architecture for learning” through producing interoperability specifications and “learning technology” standards that would enable a sustainable IT infrastructure and a vibrant marketplace for e-learning [1]. SC36 was a part of this early activity and, as a peak formal organization, it set about defining the scope of its business in terms of the types of IT systems involved and the forms in which learning content could be described and transported. While other organizations such as the IMS Global Learning Consortium and the IEEE Learning Technology Standards Committee were focused on working closely with vendors SC36 contributed its own focus and value proposition: internationalization. In many respects the work of SC36 was informed by technical specifications produced by these other standards bodies and initiatives, including formally constituted standards organizations as well as a broadening array of other consortia and movements. Prominent among these were those concerned with *openness* – open access, open source, open content, and open educational resources (OER). But despite this rich mix of innovation and standardization many stakeholders interpreted a competition taking place between standards bodies. Arguably, some still do.

Nearly two decades have passed since the first standardization efforts concerned with Internet-enabled learning first began. While the early focus of standards development was on such things as learning content, how to describe it, store it, share it, and reuse it, together with how to represent learner profiles in order to match these with appropriate content, things have now changed. These things are still important, but the implications of engaging with the Internet now have many more consequences. Consider the differences in two learning technology standards: (1) the IEEE Learning Technology Systems Architecture, published in 2003; and, (2) the Training and Learning Architecture: Experience API (xAPI) published by ADL in 2013. The former is concerned with systems interoperability; the latter, data flows instantiated as learner records.

We are now living in an increasingly globalized world in which data *is* as much as data *are* – and it is stubborn science that fails to recognise this. Likewise with *metadata*; language evolves and *data* commonly refers to *datasets* as well as *data streams*. Data flows in real-time, it has the power to cause havoc in financial markets, and it is rendered as evidence in courtrooms and research labs in ways never before possible. It

is proliferating at extraordinary volumes and potentially sourced from anything and anywhere. All kinds of digital devices can routinely capture and render any activity or utterance into data. And, this emerging era is also one in which data-intensive computing is now described as the “fourth paradigm” [2] of science. In recent parlance, it is now a world of Big Data. For researchers and agencies providing services to the public, the opportunity resides in *open data* and *linked data*.

For learning, education, and training, this new context feeds the ongoing digital revolution: disrupting, empowering, and bringing challenges, opportunities, and choices to learners, educators, and institutions alike. A notable example shows a new complexity in which *data is becoming a digital learning resource* as much as the content it may be associated with. Once upon a time the boundaries between learning content and data were quite distinct but with the growth of *learning analytics* there is now a meshing. And, this is happening in an increasingly cloud-enabled world of modular services which challenge monolithic systems and architectures.

We are also living in an intricately interconnected world that often manifests in a new tribalism while accommodating democratic and fundamentalist expression alike; in which the surveillance society has arrived by stealth; where IT systems have blurred privacy with security; and, in which the *open agenda* in the deployment of IT systems does not necessarily translate into wise action or social benefit.

Making sense of the emergent patterns of order and disorder in all this change is what we are compelled to do – it is, as Dervin suggests, a “mandate of the human condition” [3]. As a community of practice concerned with IT standards development, however, are we asking the right questions about what is required when we do this? Is it reasonable to assume that systems interoperability is necessarily the primary goal for ITLET? What about systems governance – is this just a policy issue? What do we need to consider in the development and deployment of data standards? What needs to be considered in the development of open protocols? How do we know when our environmental scan is sufficient? If human ethics committees are necessary in granting clearance to researchers who collect data from human participants then what ethical systems need to be in place in the digital domain for learning analytics?

Such questions concerning learning analytics and many others readily discoverable in the public domain suggest that the way forward for standardization in this space is not yet clear. Based upon findings associated with implementing the Question Formulation Technique (QFT) this paper argues that one way forward may be to use a structured approach to question formulation which involves identifying even more questions so that a deeper analysis can proceed [4]. Research has shown that the QFT can assist students in formulating and refining their own questions and that deeper inquiry benefits from explicit and prolonged questioning [4, 5]. For the purposes of this paper, and based on workshops already completed [10] it is proposed that QFT sessions may assist in scoping activities that often foreground all standards work [6].

## II. LEARNING ANALYTICS

### A. The Standards Context

Understanding the place that learning analytics might occupy in the ITLET standardization space is informed by historical perspective. Organizations such as the IEEE Learning Technology Standards Committee and the IMS Global Learning Consortium (IMSGLC) first identified learner profiling – understanding and specifying the learner’s abilities, requirements, and preferences – to be an important piece of an overall *architecture* or *abstract model* for IT standards that could support e-learning. Over time, this work has been complemented by activities of the W3C on the Web Accessibility Initiative and been developed further within the IMSGLC and SC36 as Access for All (AFA). This agenda has certainly been about systems interoperability but it is also concerned with optimizing the alignment of technological innovation specific to individual needs, preferences, and contexts.

Within the schooling sector – where there are particular kinds of issues associated with access to Internet-based content – the focus in standards development has therefore had a particular managerial flavour to it. This is best exemplified by the Schools Interoperability Framework Association (SIFA), which has produced a series of specifications based upon data models and infrastructure that support the sharing of administrative data – about schools and students – at the jurisdictional level. In Australia, for example, recent efforts by the National Schools Interoperability Program (NSIP) have leveraged the SIFA work and been focused on projects such as a national identity framework, a learning services framework, and inter-jurisdictional data mobility. Significantly, in the context of this paper, the focus of this work is very much to do with systems interoperability even though there are many governance-related issues to deal with.

### B. Scope

So, how is *learning analytics* currently understood? Wikipedia summarizes it as:

“the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs”

Such definitions are contested; however there is also sufficient evidence of commonality of understanding for there to be numerous examples of systems being deployed, typically as dashboards within learning management systems – but also, as an added service layer within many social media applications. Research also shows that there are numerous reports and case-studies available that describe working systems. But in formal settings where these systems report on student data an important question arises: *to what extent is there informed consent for this data collection?*

This last question can be seen to fit with many like it as recent research indicates [7, 8]. Moreover, because there appear to be so many of these kinds of questions that some reports, such as [7] group these into ethical and legal issues and then cluster them into sub-groupings. The fine detail of these

groupings includes issues such as consent, transparency, ownership, accuracy, validity, privacy, stewardship, and control – with a conclusion, that a *Code of Practice* is what is first required in this space. Such a conclusion accords with established practice in the standardization space where codes of practice and technical reports often precede standards that detail technical specifications. In closely related work [9] provides a meta-analysis of these questions after collating them with questions sourced from the European Learning Analytics Community Exchange (LACE) project. In this meta-analysis, which aims to bridge the “problem space” to the “solution space”, the issues emerging from all questions are summarised under five headings:

- Validity
- Risk-based
- Ownership
- Implementation
- Business case

Aligned with this work, [10] reports on a workshop specifically focused on learning analytics interoperability and structured to include a session using the QFT in order to elicit as many questions. In a similar manner to [9] these questions are grouped – although, because of the interoperability focus the groupings have a systems perspective and span:

- Requirements analysis
- Scope
- Interoperability
- Quality
- Evaluation
- Design
- Ethics
- Best practice

Thus, it can be seen that the inclusion of one more keyword, interoperability, yields related but substantially different semantics. This suggests, then, that further research into this area is needed. While questioning can be seen as the primary tool for inquiry it is also important to consider that it can seek or invite many different responses – answers, facts, data,

information, advice, explanations, understanding, reasons and dialogue.

### III. CONCLUSION

Two conclusions emerge from this discussion:

1. Significant issues concerning IT systems that capture, use, and share data suggest that *data and systems governance* is an area that ITLET standardization must address.
2. The requirements gathering phase of standards development could benefit from structured approaches to the formulation, collation, and analysis of questions, such as the QFT.

### REFERENCES

- [1] J. Mason, and H. Pillay, “Opening Digital Learning to Deeper Inquiry”, in *The International Handbook of E-learning*, Vol II, M. Ally and B. Khan, Eds. New York, NY: Routledge, 2015, pp. 1-10.
- [2] T. Hey, S. Tansley, and K. Tolle, Eds., *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft research: Washington, 2009.
- [3] B. Dervin, “What Methodology Does to Theory: Sense-Making Methodology as Exemplar”, in *Theories of Information Behavior*, K.E. Fisher, S. Erdelez and L. McKechnie, Eds., *ASIS&T Monograph Series*. Medford, NJ: Information Today, 2005, pp.25-30.
- [4] D. Rothstein and L. Santana, *Make just one change: Teach students to ask their own questions*. Cambridge, MA: Harvard Education Press, 2011.
- [5] R.J. Murnane, N.S. Sharkey and K.P. Boudett, “Using student-assessment results to improve instruction: Lessons from a workshop”. *Journal of Education for students placed at risk*, vol. 10-3, 2005, pp. 269-280.
- [6] T. Hoel and J. Mason, J. “Problematizing the Way we do Standards – Focusing more on Scope and Rationale”, *Proc. Open Forum during ISO/IEC JTC1 SC36 Mtg*, Eastern China Normal University, Shanghai, 2011.
- [7] N.Sclater, “Code of practice for learning analytics - A literature review of the ethical and legal issues”. London, UK: Jisc, 2014. Online [http://repository.jisc.ac.uk/5661/1/Learning\\_Analytics\\_A-Literature\\_Review.pdf](http://repository.jisc.ac.uk/5661/1/Learning_Analytics_A-Literature_Review.pdf)
- [8] S. Slade and P. Prinsloo, “Learning analytics ethical issues and dilemmas.” *Amer. Behav. Scientist* vol. 57, 2013, pp. 1510-1529.
- [9] T. Hoel, J. Mason and W. Chen, “Data Sharing for Learning Analytics – Questioning the Risks and Benefits”, unpublished.
- [10] J. Mason, T. Hoel, W. Chen and C. Stracke, “Questions as Data: Outcomes and Perspectives from the 1st ICCE Learning Analytics Workshop”, Nara, Japan, unpublished.