

Walled Gardens No More! The Expanding World of Metadata

with Lara Whitelaw & Dion Lindsay

lara.whitelaw@open.ac.uk

dion@dionlindsayconsulting.com

Lara Whitelaw was presented with the UKeIG Information Manager of the Year Award 2015 at the Intranet Now Conference in London in October 2015. She is University Metadata Development Manager at The Open University.

Dion Lindsay represented UKeIG at the conference and interviewed Lara for eLucidate a few weeks later.



Lara Whitelaw with the UKeIG Information Manager of the Year award flanked by keynote speaker James Robertson (Step Two Designs) and Dion Lindsay (UKeIG and Dion Lindsay Consulting Ltd) (Photo credit: Antonio Salgado, Capturise via the IntranetNow.co.uk conference)

Lara and the excitement of metadata work at the Open University

Lara: I work as part of the Information Management team. It's based in Library Services but we have University-wide roles. Within the tem there is a University Archivist, University Records Manager, an Information Manager, there's me as the Metadata Specialist, and the four of us work with each other to help support the University in all its activities. We also work independently, and I probably work most with individual projects

across the University and different groups to help them define their metadata requirements.

Dion: Which must give you a fair amount of both power and responsibility across the whole university. How would you sum it all up?

Lara: It's really very interesting. I'm involved in developing all kinds of metadata strategies and policies and making them more visible and trying to promote metadata as a thing we should be doing. It makes my job very diverse, working with all kinds of people and doing very different things in different months. And it means I have my toes in all kinds of areas.

Dion: How do the metadata policies work in practice? Is it that the producer of each package has to apply the metadata rules and come up with metadata and so on before the product can be launched?

Lara: No, we have various things going on. I chair the Web Standards Group and that applies a mandated policy across the whole university. And that gives basic metadata rules that need to be applied across all materials delivered on the Internet and intranet. They have to apply those and in the Web Standards Group we are working together to identify how well we are doing to ensure they are applied.

Dion: How strict is the mandated policy - does it mean materials will not go onto the Internet or intranet unless certain metadata fields are filled in?

Lara: Yes, to a certain extent. There is metadata that is applied at that level. It's not all the metadata that we want but it's what we think it's reasonable to ask for. And we use software to help us identify where we are doing well and where we are doing badly. And then we target site owners and offer them support and training to help them improve their metadata.

We have recently written a case study showing how applying metadata improves search results visibility within our intranet. And we offer workshops to help site owners applying metadata and use the publicity around the Metadata Case Study to help people in the workshops understand why metadata is important and why they should do it. We use a variety of means to get the message across.

Metadata Case Study

Dion: It's great to hear that there is a case study. The Open University create a massive amount of material and so it must be one of the best test sites for the effectiveness of metadata. Have you published that case study formally?

Lara: Without case studies you can't go to the material producers and say you must do this without giving them a good reason why. You can't just say, "because it's a good thing". Even if you can't always show that they individually gain, with the case study we can demonstrate that the University gains. And we can give feedback to the team that manages our search engine based on findings from case studies.

We haven't published formally. But we have published it within the University and we are continuing to improve the data so that we can give the best, most accurate picture of what is happening, and we do hope that we can publish it.

Automated extraction of metadata

I'm also doing a lot of work at the moment with [OpenLearn](#)¹ to understand how the metadata can automatically be extracted from video and PDF content and I really want to do more work around that because that's really exciting. I'm working with OpenLearn and an outside company. We provide the data and the outside company is providing the tools. We provide the expert understanding to help them tune the tools. So I analyse the metadata coming back and tell them where it's going wrong.

The power of metadata and the speed of change

Dion: Many information management professionals, when they have to be involved in a lot of day-to-day stuff, managing budgets and so on, sometimes wish they could be more involved in the analytical or theoretical side of information management. And then they think "But hold on. It can take forever to make these things happen. If I get my sleeves rolled up and really try to change the metadata practices here it could take forever." But you seem to have been able to achieve a lot of significant change quickly! The leveraging power of the right metadata work seems very strong.

Lara: Well, yes. It makes it very exciting. I'm tremendously lucky that my job was designed specifically to do this. It was recognised that our learning resources are tremendously valuable and unique to the University. And the reality is a very small amount of metadata work can make a huge difference to the visibility of content.

It's the analysis of relationships between concepts that I find really exciting and which gives metadata real power to break down the boundaries between silos. Working in professional information management in libraries, we sometimes work in arenas where we have neat systems that have nice structures to them, or at least some shape. So as information professionals we are sometimes in "walled gardens" of our own, and no one else gets to see our work. And it is actually metadata that helps us break down those barriers and make our repositories understandable to the world outside, and for example to the Google systems in the wild open spaces we have no control over!

So I'm interested in things like linked data and ways of identifying and defining relationships within our data and between our data and other kinds of data outside our systems so that it can work both ways: not only can other people understand our information better, but we can also hopefully pull in information from other systems and enrich our own material in ways that we would never be able to otherwise.

¹ [OpenLearn](#) aims to break the barriers to education by reaching millions of learners around the world, providing free educational resources and inviting all to sample for free courses that Open University registered students take.

The Digital Archive

Lara: The Digital Archive is an example where we've always had lots and lots of material from different sources for the various courses that we've taught since the University began but it's all been in library management systems, content management systems, sometimes obscured in ways that it's hard to see the bigger picture. It was only when we are able to define the relationships within the content and between courses and content that we are able to develop an interface which allows students and producers to see those materials as a whole and understand what each course was trying to teach and what the context of the course material was.

Dion: And that must be very satisfying not just for the users but also for the authors of the material who can now find new audiences for the material.

Lara: Definitely. We can actually say now that a course we are teaching today is actually conceptually the great-grandchild of the course that we were teaching 15 years ago, and that shows how the subject has moved on. And where we can see that 15 years ago we co-produced with the BBC materials to support the old course, that helps us understand enough to use clever tools to reach out to current BBC material because we've managed to define the relationships and that gives us great power that we didn't have before.

Professional networks

Dion: Where do you get your professional support from? Are there professional groups that you recommend in this field?

Lara: Well yes, obviously there is *UKeiG*. And there's also *ASIS&T*², which includes the Dublin Core Learning Resource Metadata Initiatives that are really key to helping us make our resources accessible on the web. And then there's obviously *ISKO*³. I attended their last conference and they always do interesting things and look at interesting aspects of how taxonomies can be used. *ASIS&T* looks more at the schema side of things and Dublin Core and suchlike which are more critical for me. All three are very good places to start and *ISKO* run very very good half-day events for people interested in niche areas. I would recommend anybody to go if they're interested in taxonomy management, where it currently sits and what the innovations are.

Intranets and Intranet Now

Dion: We met of course for the first time at Intranet Now⁴ in October 2015 - that was the second of the Intranet Now branded conferences. How did you enjoy the day?

Lara: Oh yeah, I loved it! It was really relaxed. Some of my conferences are really dense in what they are trying to get across and what your trying to learn and can be full of people being very serious, and Intranet Now was a lovely conference. Not that it wasn't about something serious, but people were very welcoming and I loved the whole relaxed atmosphere especially when people were able to come up on stage and suggest topics for

² Association for Information Science and Technology

³ International Society for Knowledge Organisation

⁴ Intranet Now – a conference for comms and intranet people. October 13 2015. Presentations available at <http://intranetnow.co.uk/speakers/>

the un-conference in the afternoon. And because intranets are, I think, frequently overlooked even though people working for an organisation interact with them all the time, Intranet Now gives a very diverse group of people with all kinds of different skills somewhere to discuss the innovations and opportunities and particularly the issues we have in these areas.

And as our work becomes more mobile and less constrained by physical space intranets become more important. It's not only a place where we are looking for information and using it as communications channels to do our day-to-day jobs, it is also somewhere we have a voice and a sense of place even if we are not located within the building with fifty people around us anymore. Intranet Now is recognising the effort people are putting in to improving this experience. I particularly liked the speaker⁵ who worked for an architectural firm who created an intranet for a fraction of what other organisations spend, and had gone in a completely different direction and created a truly beautiful intranet. I loved their promotional mugs, and how they spent hours deciding what exact shade of green should be used: because I'm really interested in the visual representation of information. That intranet worked really hard on the visual representation aspects -that was fabulous!

Semantic data

Dion: You mentioned "walled gardens" earlier and how metadata is helping break them open.

Lara: Yes. And this is fairly recent. Google didn't go near metadata for ages because it is way too easy to spoof metadata. There would have been too many people putting Britney Spears in the keywords to their metadata to just raise their content up the rankings. But then semantic data came along and Schema.org⁶ came along. Schema.org was developed by the search engines to help them give more structure to data and therefore bring back results with more information provided so that users can have more confidence that the results were things they were interested in and also so they can provide that faceted information. The little summaries that Google provides now at the top of a search results relies a lot on information provided by semantic data that is pulled from DBpedia⁷.

Dion: What final message would you like to leave our readers with?

Lara: It's just that although sometimes metadata work can seem really impenetrable, with lots of acronyms and lots of hard work for what seems like a little reward, in reality you can make a huge difference with it and it doesn't have to be lots of people slaving away in a back room handcrafting metadata and making it perfect within your perfect walled garden. Very few of us work in an environment now where that is even a remote possibility.

⁵ [Shaula Zanchi](#) at Robert Partington and Partners. For slide show see intranetnow.co.uk/speakers/

⁶ Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond. Schema.org is sponsored by Google, Microsoft, Yahoo and Yandex.

⁷ DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web

Instead we have to use a mixture of techniques but I still think that defining relationships using metadata to represent relationships between concepts, between the things that are important to you, can be really powerful and can give you huge amount of benefits. Persuading your colleagues to do a small amount will give them, and you, a massive benefit. That's what makes it such an exciting time to be involved in metadata right now!