**SICSA Data Science: Scottish Linked Data Interest Group workshop**

November 9, 2014 in [Events](http://scot.okfn.org/category/events-2/), [Linked Data](http://scot.okfn.org/category/linked-data/), [Open Data](http://scot.okfn.org/category/open-data/)

The [School of Informatics](http://www.ed.ac.uk/schools-departments/informatics/) hosted the 3rd Scottish Linked Data Interest Group Workshop on the 10th September 2014, following on from two successful workshops earlier in the year organised by [Jeff Pan](http://homepages.abdn.ac.uk/jeff.z.pan/pages/) in the University of Aberdeen. The event brought together a diverse set of about 30 people with different interests and perspectives on Linked Data. Apart from Edinburgh, the geographical spread of participants covered Glasgow, Dumfries, Haddington and Aberdeen, and their backgrounds ranged across academia, the public sector (both local and central government), broadcasting and software development. Topics of discussion ranged from theoretical and research oriented through to applied and practical.

During the closing session, some participants confessed to feeling that the more technical talks were outside their comfort zone. Despite this, there seemed to a consensus that the heterogeneity of both the audience and the subject matter was a big plus and that it would be a mistake to have separate workshop tracks for, say, academia and public sector; each group can learn from the other.

**Summary of the talks:**

**András Salamon: Dynamic schema discovery**

By collecting together URIs that are similar into classes, the graph connecting the classes forms a high level description of the data. This is known as a largest simulation. When the links between data do not form cycles the largest simulation is essentially a traditional database schema. Although a largest simulation is too expensive to compute for web-scale graphs, it can be approximated by first deciding which are leaf vertices in the graph, and then making linear passes over the data to successively refine the simulation. The talk reviews recent progress in efficient approximation of largest simulations. This also leads to fast rematerialisation of tabular data that has been dematerialised into RDF.

Slides available [here](http://scot.okfn.org/files/2014/11/Andr%C3%A1sSalamon.pdf).

**Yuan Ren: Supporting Competency Question-driven Ontology Authoring**

Ontology authoring is a non-trivial task for novice authors who are not proﬁcient in logic. It is difﬁcult to specify the requirements to an ontology, and test whether the requirements have been satisfied. Our approach is based on a Competency Question-driven Ontology Authoring pipeline where authors use competency questions to specify their functional requirements. The answerability of competency questions, which can be verified by authoring tests, ensures the satisfiability of requirements. A dialogue-based authoring interface has been developed to extract the features and elements from competency questions and generate authoring tests. Reasoning support is invoked on the fly to perform testing, and informative feedback on the consequences of authoring actions is provided to users.

Slides available [here](http://scot.okfn.org/files/2014/11/YuanRen.pdf).

**Michael Roth: Semantic parsing to Linked Data**

I describe ongoing work in the [S-CASE project](http://www.scasefp7.eu/), which aims at assisting software engineering by providing a repository of existing and re-usable software components. Our work focusses on the task of parsing software requirements, which describe the functionality of a software system and are hence fundamental in the development of new software components. Software requirements are commonly written in natural language, making them prone to ambiguity, incompleteness and inconsistency. By mapping requirements to formal semantic representations, emerging problems can be detected at an early stage of the development process, thus reducing the number of ensuing errors and the development cost. Linked data provides an appropriate framework for representing the semantics of software requirements and enables the sharing and linking of functional information with other data relevant to the software domain.

Slides available [here](http://www.slideshare.net/microth/semantic-parsing-to-linked-data).

**Fiona McNeill: Dynamic data sharing for facilitating communication during emergency responses**

Emergency response situations are characterised by the need for fast, efficient data exchange. Post-disaster reports into responses usually highlight a failure to do this as one of the key impediments of a successful response. Current practice is heavily dependent on human interaction as the primary means of data sharing, and it is clear that some level of automation needs to be introduced to appropriately support these human users to effectively and appropriately share relevant parts of their large data sources. One of the key problems of automation is the incompatibility that exists, on several layers, between different data sources. I will describe a system, CHAIN, that is designed to automatically extract relevant parts of a data source when that source is queried, even if there are multi-layered incompatibilities between the query and the data source. This facilitates dynamic sharing of the data even when the data sources have not been pre-aligned.

Slides available [here](http://scot.okfn.org/files/2014/11/FionaMcNeill-.pdf).

**Amy Guy: Context-aware properties**

Many concepts in ontologies are ambiguous and have multiple valid interpretations depending on the circumstances in which they are to be interpreted. I suggests that rather than relying on the annotation given at the time of data publication, determining the meaning of data can take place at an application level at the time of data use, when additional information about the circumstances of its use is available. Creative media production on the web is used as a real-world scenario through which to explore how ambiguity of concepts might be utilised — rather than overcome — for the benefit of the end users of the data.

Slides available [here](https://docs.google.com/presentation/d/11F6YGdVkGaz-VDV7A0M5oW1VBds-FEqaRrEyL9WPNCU/edit#slide=id.g3b14eaaad_016).

**Alex Tucker & Chiara del Vescovo: RES: Research and Education Space — what are we going to be tangled up with?**

The Research & Education Space (RES) is a project being jointly delivered by Jisc, the British Universities Film & Video Council (BUFVC), and the BBC. Its aim is to bring as much as possible of the UK’s publicly-held archives, and more besides, to learners and teachers across the UK.

At the heart of RES is [Acropolis](https://bbcarchdev.github.io/inside-acropolis/), a technical platform which will collect, index and organise rich structured data about those archive collections published as Linked Open Data on the Web. The collected data is organised around the people, places, events, concepts and things related to the items in the archive collections — and, if the archive assets themselves are available in digital form, that data includes the information on how to access them, and for which use they are copyright-cleared, all in a consistent machine-readable form. Building on the Acropolis platform, applications can make use of this index, along with the source data itself, in order to make those collections accessible and meaningful.

Slides available [here](http://www.slideshare.net/ChiaraDelVescovo/slides-38977237).

**Hai Nguyen: CURIOS Mobile — Linked Data Exploitation for Tourist Mobile Apps in Rural Scotland**

Many tourist mobile apps currently use narratives generated specifically for the app and often require a reliable Internet connection to download data from the cloud. These requirements are difficult to achieve in rural settings where many interesting cultural heritage sites are located. Although Linked Data has become a very popular format to preserve historical and cultural archives, it has not been applied to a great extent in the tourist sector. This talk describes an approach to using Linked Data technology for enhancing visitors’ experience in rural settings. In particular, we present CURIOS Mobile, the implementation of our approach and an initial evaluation from a case study conducted in the Western Isles of Scotland.

Slides available [here](http://scot.okfn.org/files/2014/11/HaiNguyen.pdf).

**Paolo Paolo: Integrating Know-How in the Linked Data Cloud**

Know-how available on the Web, such as step-by-step instructions, is largely unstructured and isolated from other sources of online knowledge. To overcome these limitations, we propose extending to procedural knowledge the benefits that Linked Data has already brought to representing and retrieving declarative knowledge. We describe a framework for representing generic know-how as Linked Data and for automatically acquiring this representation from existing resources on the Web. This system also allows the automatic generation of links between different know-how resources, and between those resources and other online knowledge bases, such as DBpedia. We discuss the results of applying this framework to a real-world scenario and we show how it outperforms existing community-based integration efforts.

Slides available [here](http://www.slideshare.net/PaoloPareti1/slideshare-copy-of-10-september-2014-3rd-scottish-linked-data-workshop-38975818).

**Peter Winstanley: The Share PSI 2.0 Thematic Network**

This talk presents the Share PSI 2.0 Thematic Network which was created to promote the sharing of information and experiences about the implementation of open data in the European public sector. The creation of this network is motivated by the growing awareness of the importance of data in different sectors of the future economy. An important goal to be achieved is making the process of opening data more attractive by maximising the returns on investment. This involves addressing a number of different problems, ranging from legal issues to data interoperability. Overall, the Share PSI 2.0 Thematic Network is bringing together a large number of government and academic organizations alongside others such as the World Wide Web Consortium (W3C), the Open Knowledge Foundation and the Open Data Institute.

Slides available [here](http://www.slideshare.net/PeterWinstanley1/2014-0910-sharepsi2talkscottishlinkeddatagroup).

**Ian Watt & Andrew Sage: Aberdeen Linked / Open Data Initiatives**

This talk reviews the Linked Data and Open Data strategies adopted by the City of Aberdeen Council and focusses on two recent initiatives: [Code for Europe](http://codeforeurope.net/) and [Code the City](http://codethecity.org/). These initiatives strived to achieve a high level of engagement with the community, allowing people to decide which problems to solve by using open data about the city. Many aspects were discussed, such as the technical challenges of creating a solid open data infrastructure, the actual datasets that were published and the applications developed on top of them.

Slides available [here](http://www.slideshare.net/ACC_IW/sldig-140910).

- See more at: http://scot.okfn.org/2014/11/09/scottish-linked-data-interest-group-workshop/#sthash.eLLYh1GK.dpuf