index and enrich numerous datasets, through cross-dataset inference.

.

The Information Systems Laboratory of the Institute of Computer Science of FORTH, designs and develops innovative indexes, algorithms and tools to assist the process of semantic integration of data at a large scale. This endeavour started two years ago, and the current suite of services and tools that have been developed are known as “LODsyndesis” [L1]. As shown in Figure 1, the distinctive characteristic of LODsyndesis is that it indexes the contents of all datasets in the Linked Open Data cloud [1]. It then exploits the contents of datasets to measure the commonalities among the datasets. The results of measurements are published on the web and can be used to perform several tasks. Indeed, “global scale” indexing can be used to achieve various outcomes, and an overview of the available LODsyndesis-based services for these tasks can be seen in Figure 2.

First, it is useful for dataset discovery, since it enables content-based dataset discovery services, e.g., it allows answering queries of the form: “find the K datasets that are more connected to a particular dataset”. Another task is object co-reference, i.e., how to obtain complete information about one particular entity (identified by a URI) or set of entities, including provenance information. LODSyndesis can also help with assessing the quality and veracity of data, since the collection of all information about an entity, and the cross-dataset inference that can be achieved, allows contradictions to be identified and provides information for data cleaning or for estimating and suggesting which data are probably correct or most accurate. Last but not least, these services can help to enrich datasets with more features that can obtain better predictions in machine learning tasks [2] and the related measurements can be exploited to provide more informative visualisation and monitoring services.

The realisation of the services of LODSyndesis is challenging because they presuppose knowledge of all datasets. Moreover, computing the transitive closure of all “owl:sameAs” relationships is challenging since most of the algorithms for doing this require a lot of memory. To tackle these problems LODsyndesis is based on innovative indexes and algorithms [1] appropriate for the needs of the desired services. The current version of LODsyndesis indexes 1.8 billion triples from 302 datasets, it contains measurements of the number of common entities among any combination of datasets (e.g., how many common entities exist in a triad of datasets), while the performed measurements are also available in DataHub [L2] for direct use. It is worth noting that currently only 38.2% of pairs of datasets are connected (i.e., they share at least one common entity) and only 2% of entities occur in three or more datasets. The aforementioned measurements reveal the sparsity of the current datasets of the LOD Cloud and justify the need for services for assisting integration at large scale. This method is efficient, the time for constructing the indexes and performing all these measurements being only 22 minutes using a cluster of 64 computers.

Over the next two years we plan to improve and extend this suite of services. Specifically, we plan to advance the data discovery services and to design services for estimating the veracity and trust of data.

This project is funded by FORTH and it has been supported by the European Union Horizon 2020 research and innovation programme under the BlueBRIDGE project (Grant agreement No 675680).

Links:
[L1] http://www.ics.forth.gr/isl/LODsyndesis/
[L2] https://datahub.io/dataset/connectivity-of-lod-datasets

References:

[1] M. Mountantonakis and Y. Tzitzikas, “On measuring the lattice of commonalities among several linked datasets,” Proceedings of the VLDB Endowment, vol. 9, no. 12, pp. 1101-1112, 2016.
[2] M. Mountantonakis and Y. Tzitzikas, “How Linked Data can aid Machine Learning-based Tasks,” in International Conference on Theory and Practice of Digital Libraries, 2017.

Please contact:

Yannis Tzitzikas
FORTH-ICS and University of Crete
+30 2810 391621
tzitzik@ics.forth.gr

On the Interplay between Physical and Digital World Accessibility

by Christophe Ponsard, Jean Vanderdonckt and Lyse Saintjean (CETIC)

Many means of navigating our physical world are now available electronically: maps, streets, schedules, points of interest, etc. While this “digital clone” is continually expanding both in completeness and accessibility, an interesting interplay scenario arises between physical and digital worlds that can benefit all of us, especially the mobility-impaired.

Physical world accessibility is about ensuring that features of physical places (e.g., shops, tourist attractions, offices) are designed to accommodate the (dis)abilities of people visiting them in order to ensure optimal access for all, including the estimated 15% of the population living with some kind of impairment. Our world is undergoing digitisation at an ever-increasing rate. Online maps are becoming so precise that the inner structure of buildings is often captured, tours are available, while locations are ranked by users, etc. As a consequence, many new opportunities have arisen for breaking accessibility barriers and better combining the physical and digital aspects of accessibility. For example, the use of the available online information can be complemented with physical measures to help assess the accessibility of buildings and modes of transport. Combining digital maps, crowd-sourced information, open data and expert assessments can help produce up-to-date accessibility data at a larger scale and reduced cost.

Physical accessibility requirements are now well defined and documented, such as in the ISO21542 standard. Assessing physical accessibility is not straightforward because it requires the identification of physical obstacles to accessibility relevant to specific kinds of impairment, e.g. a doorstep for a wheelchair, the presence or not of braille on lift buttons for blind people, the printing of light labels for the vision-impaired. Experts can reliably evaluate this accessibility by applying a well-defined measurement procedure and web-oriented reporting like Access-I [L1]. However, the limited number of trained experts greatly restricts the ability to carry out such assessments at a large scale and keep the information up-to-date. On the other hand, dedicated social media apps like jaccede.com or wheelchair.org enable a crowdsourcing approach through the collection of partial information reported by different users, each with her own point-of-view, a method which may, however, lead to inconsistency. Both approaches are naturally complementary and can efficiently be combined through the use of open linked data techniques [2].

Although information and communication technologies are a clear enabler for physical accessibility, they also raise new obstacles in that the end-users then need to consult a computer or mobile interface to make sure that a place is physically accessible. The Web Accessibility Initiative (WAI) [L2] provides specific guidelines, such as the Web Content Accessibility Guidelines (WCAG) to help web developers make websites accessible while dynamic content is addressed through Accessible Rich Internet Applications (ARIA). Many useful evaluation and repair tools are also proposed, including for tailored and optimised usability and accessibility evaluation [4]. Many organisations propose practical labels that are based on WCAG, like AnySurfer in Belgium [L3].

Figure 1 graphically depicts the interplay between physical and digital accessibility, thus raising interesting questions about the current level of awareness about each kind of accessibility, the possible correlation between them, and what kind of synergies can further enhance user experience. In order to answer these questions, a preliminary study was carried out in Belgium by combining data from Access-I (for physical accessibility) and AnySurfer (for digital accessibility). Some methodological alignment was performed to define comparable notions of "unacceptable", "satisfactory" and "good" levels of accessibility for physical, sensory and mental impairments. The survey was conducted on places already assessed (so with some bias). Figure 1 shows a satisfactory level of physical accessibility in 60% of cases, of digital accessibility in 70% of cases and both in 33% of cases. The generally better level of digital accessibility, compared with physical accessibility, might be explained by the increased importance of digital presence and the lower effort required to achieve digital accessibility. Indeed, a website is easier and less expensive to repurpose than a building. The fact that only one third of the places studied currently achieve both dimensions means that these aspects are still largely being considered separately.

The current scope of both accessibility and digital assessments should also be reviewed: Physical accessibility is not limited to building boundaries but starts when the intention to travel to a place arises. And this scenario, eventually leading to the use of some service or equipment inside a building, relying on using a computer and/or a mobile interface like a smartphone or tablet to query the web about transportation, opening hours, and user specific accessibility. Physical and digital accessibilities then closely intertwine with retroaction loops illustrated in Figure 1.

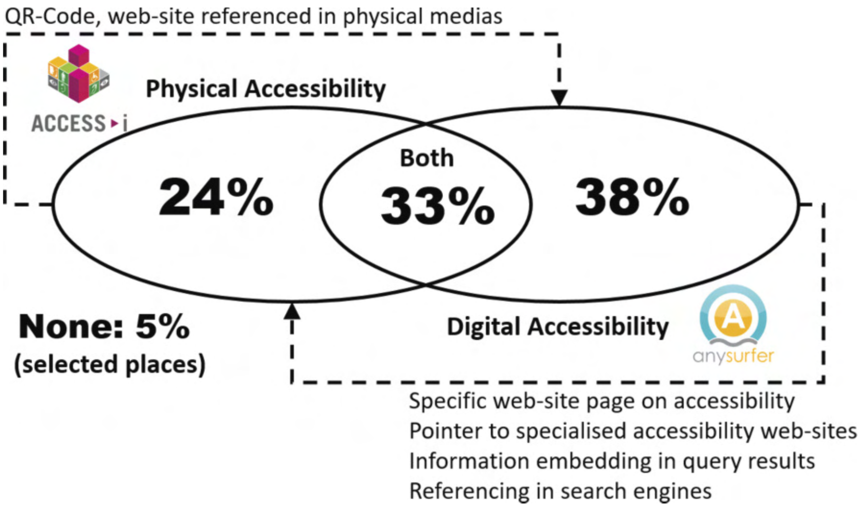

Figure 1: The interplay between physical and digital accessibilities.

In addition to deepening our survey on the relations between physical and digital accessibility, we are also currently investigating how to effectively support accessibility using a digital mobile companion exploiting aggregated open accessibility data [3].

Links:
[L1] http://www.access-i.be
[L2] https://www.w3.org/WAI
[L3] http://www.anysurfer.be/en

References:
[1] C. Ponsard, V. Snoeck: "Unlocking Physical World Accessibility through ICT: A SWOT Analysis", Proc. of ICCHP'2014.
[2] C. Ponsard et al: "A Mobile Travel Companion Based on Open Accessibility Data", Proc. of ICCHP'2016.
[3] J. Vanderdonckt, A. Beirekdar: "Automated Web Evaluation by Guideline Review", Journal of Web Engineering, Vol. 4, No. 2, 2005, pp. 102-117.

Please contact:
Christophe Ponsard, CETIC, Belgium
+32 472 56 90 99, christophe.ponsard@cetic.be