Information Visualisation Skills: an Initial Investigation of Industrial Demand and Educational Supply

Milena Dobreva[a] and Fernando Loizides[b]

[a]University of Malta, Malta
[b]University of Wolverhampton, UK
milena.dobreva@um.edu.mt, fernando.loizides@wlv.ac.uk

Abstract. Information Visualisation developed historically as a discipline enhancing the presentation and facilitating the understanding of information. Nowadays, it becomes one of the essential skills in presenting and using big data. This leads to an increase in the demand for professionals with sound information visualization skills, and subsequently is a catalyst for Universities to offer higher education courses or programmes in this area. In our paper, we offer some initial evidence on the current state of demand and supply. We present industry requirements for jobs in information visualisation, already running and established courses and future courses that are being planned in information visualisation using a mixed method approach combining interviews and content analysis. Being an initial scoping study, our research is providing a partial picture of the current state but it also helps to formulate questions for a further in-depth exploratory research.

Keywords: Information Visualisation, Education, Data Science, Big Data

1. Introduction and Motivation

With the phenomenon of information overload [1] and the increased use of Big Data applications [2], there is a growing need to present information in ways in which can be easily assimilated. Information Visualisation becomes one of the key factors in interfaces and information architecture of contemporary information systems. Their use does not require only the skills to prepare information visualisations, but also to make use of them. In general this means that two new types of skills need to be formed:

- Professional skills to create information visualisations; and
- Skills for the widest range of users, including policy makers and the citizens, to interpret and use further the visualisations.

© Springer-Verlag Berlin Heidelberg 2011

The latter had been in the focus of discussions initiated by the European Commission on the use of big data in the large policy making and the widely lacking skills for understanding visualisations is one of the primary concerns as expressed in [3].

Hence, the educational system has another trend to adapt to-and while the skills in the domain of visual literacy need to be formed in an early age, in this paper we concentrate on the higher education system which is preparing professionals. In the last few years, information visualization courses have begun to appear in higher education as part of the post-graduate and also undergraduate curriculum [4]. The extent of most are one off lessons, sometimes a single hour lesson, as part of a larger curriculum, such as computer science, data science, information science, and more recently, design courses. In addition, there have also began to exist, albeit very few, Master level as well as Bachelor level Degrees in information visualisation. Despite the large interest in including the subject of information visualisation within a higher education curriculum and the fact that there is ample material to guide teaching information visualisation, there is sparse research and reports on information visualisation educated specifically.

In this paper, we will present an initial content analysis of the descriptions of 11 existing courses, 60 job adverts enriched by 15 interviews with academics and professionals that have a vested interest in creating an information visualisation based course.

2. General Methodology

The research presented investigates three areas which provide evidence on the demand for information visualization skills in the industry, and the supply in terms of higher education courses aiming to prepare professionals with the required skills. The first is already existing courses on information visualisation, the second is potential and future courses and the third looks at jobs advertised specifically with information visualisation candidates in mind.

In order to analyse the existing courses and the job opportunities, we used the same methodology (content analysis) gathering data from published job adverts and-course descriptions, and adding a set of interviews. Firstly, we selected courses from universities which were within the 100 top universities in their respective countries according to the times higher education guide. The courses included 5 from Europe (UK, France, Netherlands, Greece and Germany) 3 from the United States, 1 from Australia, 1 from New Zealand and 1 from Brazil. Similarly, the job adverts we collected were selected from common on-line job search sites of each individual country (e.g. Monster.co.uk for the UK). The search was made on both the terms ‘information visualisation’ and ‘information visualization’ in order not to eliminate results from a specific dialect. The results gave us the textual descriptions of the different courses and jobs. These were then manually split into the different sections required; for example, the requirements needed to apply for a position or the description of what a course contained. These were then processed in order to extract the key words from the text. Stop words such as ‘and’ were eliminated. Words with different spelling (e.g. visualisation and visualization) were concatenated and similar words (e.g. ‘analysis’ and ‘analytical’, or singular and plural like ‘course’ and ‘courses’) were also joined.

Such concatenations are clearly identifiable because they appear as word fractions, e.g. ‘analys’ and ‘cours’. Lastly, the process of stemming was introduced [5] in order to reduce words to their basic form in order to identify the same meanings (e.g. ‘analysed’ and ‘analyse’ are reduced to ‘analys’). The frequency and the number of documents that each word appears in were then extracted in order to prioritize similar terms from different participants or adverts.

Future course information was extracted in one of two ways. Firstly, through a workshop that was run as part of the KnowEscape COST Action TD1210 (*Analyzing the dynamics of information and knowledge landscapes*) at the Electronic Publishing Conference 2015 (ElPub 2015) in Malta. 8 participants were given questions to answer in written form and an open discussion took place followed by follow-up interviews from participants for clarification. Further participants were recruited from specific disciplines and interviewed online with a semi-structured interview, covering the same subjects as the workshop participants. Responses were gathered from academics on a global scale, with opinions of academics working in Australia, Canada, Germany, Greece, The Netherlands, New Zealand, Malta, and the UK.

3. Findings and Conclusions

The paper will be presenting the initial evidence which makes it possible to compare the content of sources which illustrate industrial demand on information visualization with the current response of the higher education system. Since this initial exercise already shows some substantial differences between the industrial and academic approach, as well as a huge variation in the focus of the various academic sources consulted, we intend to expand the study in the future. Within this article we have not added a systematic review of education-related papers from the traditional InfoVis conferences because our initial analysis showed that not every conference includes an education-related sessions, and in most cases such sessions include papers exploring the use of a particular tool in education, which is an additional area for exploration which was not in the primary focus of this first scoping study,

The analysis which will be presented in detail in the paper showed some serious differences in the content and main areas of focus of existing university programmes and industrial demand. While the courses created for the future put a greater emphasis on the visual literature, the connection to the industrial needs is still one of the weaker components factored in the curriculum development. Thus the need to understand industrial needs better is one of our main findings. In general our work touches mostly the professional skills while the more generic skills in interpreting and using visualisations remain an area where the possible approaches need to be discussed further.

However, an additional analysis of the research trends in information visualization, expanding the work presented by Ke, Börner, and Viswanath in 2004 [6] and bringing it to the most recent research, would be a helpful addition to add the dimension of state-of-the-art research to the identified demands and supply trends.

We also envisage further refinement when considering the **specific needs of various professional communities**. For example it would be interesting to explore how the current visualization advances in the domain of digital libraries [7] are reflected in the programmes of iSchools and other information science departments.

There is a discussion on the best approaches [8], [9] but it definitely would develop further – and with the type of evidence we started to gather the needs of specific domains can be compared to the 'general' information visualization demands.

And finally, one substantial topic for the future would be how to introduce to information visualization different types of specialists who start to use large data sets on a new scale and need information visualization skills. This aspect had been addressed in the first theme of emerging themes within the domain of use of data for policy making in the recent report on Data for Policy: "Concerns about the availability of relevant skills in public sector organisations, e.g. skills related to data collection, data analytics and interpretation of (visual) data" [3]. Again, in such areas of widening the use of visualization by professionals who have not been trained within their education, there will be a need to offer professional courses filling the gap.

Understanding better the demands and supply, and fine-tuning the specific requirements of professional communities and wider users, would require a substantial additional gathering of data. However, we do hope that the current article puts a foundation for adding further refinements and extensions to the exciting domain of how to prepare professionals in the domain of information visualization.

Acknowledgements

The research work presented in this article is partially supported by the FP7 grant 316087 AComIn "Advanced Computing for Innovation", funded by the European Commission in the FP7 Capacity Programme in 2012–2016, and by the COST action TD1210, KNOwESCAPE.

References

1. Bawden,D. and Lyn Robinson. The dark side of information: overload, anxiety and other paradoxes and pathologies. *J. Inf. Sci.* 35, 2, 2009.
2. Chen, M. et al. Big Data: Related Technologies, Challenges and Future Prospects. Springer, 2014.
3. Data for policy. A study of big data and other innovative data-driven approaches for evidence-informed policy making.2015. Available on http://media.wix.com/ugd/c04ef4_8dc2051fbb51459fa9eeac70be46ddbe.pdf
4. Kerren, A., Stasko, J. T., & Dykes, J. Teaching Information Visualization. *Information Visualization*, 65.
5. Porter, M. F. An algorithm for suffix stripping. *Program* 14, no. 3 (1980): 130-137.
6. Ke, W., Börner, K. and Viswanath, L. (2004) Major Information Visualization Authors, Papers and Topics in the ACM Library. IEEE Information Visualization Conference, Houston, Texas, Oct 10-12, 2004.
7. Scharnhorst. A. Walking through a library remotely – Why we need maps for collections and how KnoweScape can help us to make them? CoRR abs/1503.06776 (2015)
8. Hemsley, J. et al. Visualization Pedagogy in iSchools. Workshop description. In iConference 2015 Proceedings. 2015.
9. Ostergren M., Hemsley J., Belarde-Lewis M., Walker S. (2011) A Vision for Information Visualization in Information Science. In: iConference 2011, February 8-11, 2011, Seattle, WA, USA. pp. 531-537.