Information Science education in Darmstadt

Melanie Siegel and **Margot Mieskes**

Darmstadt University of Applied Sciences, Germany

Abstract

Information Science can be complicated to explain. This obvious weakness is in fact an asset of the science: Students get a solid basic education with a technological focus and are versatile. Information Science can and must review its content and adapt to the requirements of society and the working environment on a regular basis. Besides the solid basics, the curriculum contains innovative knowledge-driven and project-based content.

Keywords: information science topics, research and education, society and education, teaching methodology

Introduction

Information and Library Science is – as the name suggests – a very broad area with many scientific applications (as in Lee 2018). As such, it is hard to explain briefly what this topic is about. This is also reflected in study courses which have to cover a broad variety of topics before students are able to focus on their specific interests and strengths. Additionally, Information Science deals with information, which becomes more and more important in our information society.

A survey of graduates at our University shows that the degree programme effectively prepares graduates for positions not only in libraries, but increasingly also in the data processing industry. At a symposium on information science, the students showed a variety of their projects, which was met with lively interest by the attending companies from the IT sector. The diversity and interdisciplinary nature of the subject should therefore be seen as an opportunity rather than a shortcoming. The necessary constant redefinition of the subject with reference to developments in industry and society is a major strength of information science.

Teachers and researchers in this area therefore have to balance the need for a solid basic education in a vast study field, while also teaching state-of-the-art methods and bring recent research results into the lecture hall as well.

Information Science Topics

Three major areas define Information Science: structuring and presentation of information, search for information and generation of information – independent of whether we are in a digital or analogue domain.

In order to structure and present information, knowledge in databases, thesauri, ontologies and visualization are required. In recent years, this moved from using expert knowledge and hand-crafted data to employing automatic means to extract and aggregate information from more or less structured data. Therefore, a curriculum in Information Science has to cover topics ranging from classical computer science topics, such as basic programming skills, database knowledge and visualization to topics from the area of language technology and linguistics.

Search for information has moved from information broking to information retrieval, which again shares elements with computer science. Additionally, information has to be made accessible from the vast amount of structured and unstructured data, such as e-Mails, blogs, social media, but also audio-visual data. Methods allowing to access these types of data have to go beyond keywords, but rather integrate semantic knowledge and relations.

The third area is the generation of information. The first priority here is to generate high-quality information, i. e. to ensure the quality of information (e. g. Erdmann et al. 2017). In addition, there is the structuring of information with reference to its function (e. g. Muthig and Schäflein-Armbruster 2008), and also the ideas of semantic annotation. Adapting the presentation of information to target groups and thus facilitating inclusion is another important task. Siegel and Lieske (2015) show how this concern can be supported by language technology methods. Finally, this area also includes the translation of information whose support by automatic procedures has nowadays become standard (see Porsiel 2017).

Linguistic Information and Language Technology

A large part of the information is encoded in linguistic form, as text. The information of the characters in the text can only be obtained by determining the meaning of the language, i. e. its semantics. While humans understand language intuitively, understanding is a significant computing power for a machine. How, for example, should a machine automatically recognize what is taught in a "girls' trade school"? This includes cultural knowledge and knowledge about the context in which the word appears. However, information is only really valuable when it is linked. A basic course in the first and second semesters of information science therefore deals with the central questions: What is meaning? What can be done with semantic analysis?

Humans can understand and process language intuitively. We are able to draw considerable semantic inferences. However, in the age of large amounts of data that need to be searched, structured, understood, and produced, we depend on machines that support us in this. For information science, this means focusing on technology that can automatically process the meaning (semantics) of the language, in order to generate information from language, namely semantic technology.

Semantic technology is first of all language technology, such as technology for processing morphology, automatic recognition of named entities, recognition of collocations, recognition of word variants, automatic resolution of ambiguities, semantic analysis of sentences, learning of semantic relationships from annotated texts, and much more. Further, technology to structure texts and information and to present the meaning, such as XML, meta tags and ontology description languages is needed (see also Agogo and Hess 2018)

Information Science and Society

Information Science highly interacts with society. On the one hand, we have to teach our students to make meaningful use of the tools available, on the other hand, we have to teach them to act responsibly. Therefore, they need to acquire knowledge and skills relevant for today’s labour market, but which also enable them to adapt to changes that will no doubt occur 10 or 20 years during their active life. But in light of current events, such as the Cambridge Analytica affair and the data protection regulation changes in the EU, we also have to make our students aware of juridical and ethical issues involved in their work and the technology and tools they use.

An important part of information science studies in Darmstadt is an internship. Students get to know the job market for information scientists and the work processes in the companies. Companies are also discovering the study of information science as a qualification that can support important work processes in the company. The compulsory reports written by students once they finished their internship also covers questions concerning the elements of their study course they found valuable, which were unimportant and which should be extended. This information, which comes from about 60 students per year give us useful feedback what could and should be added to the curriculum and what needs to be reduced or even removed.

Additionally, we established a yearly workshop to bring together industrial partners, students and us both as teachers and researchers. Each present their work and their interest. This platform allows us to get and to give an overview on today's topics, but also as researchers to interact with industrial partners and establish research cooperation.

It becomes evident that information scientists can be employed in very different industries. These are of course museums and libraries, but also airlines, pharmaceutical companies, marketing agencies, online shops, information service providers, software companies, automotive companies, technical documentation, film institutes, travel agencies, chemical companies, start-up companies, railway companies, telecommunications companies and others. In all these cases, technical knowledge such as databases, programming skills, search engine technology and semantic data models are at the forefront of our students' practical tasks. In addition, there is knowledge of modern project management methods.

The graduates of information science accompany the digitalization of society and economy. Topics that are emerging in the economy, such as information in the digitalized industry 4.0 and the natural language interface to information, must be integrated into the curriculum flexibly and quickly.

Industrial cooperation is also ensured by the fact that many students write their final theses in companies and contribute information science content there. In practice-oriented research projects between companies and the university, innovative topics are tested for practical relevance.

Innovative Teaching Methods

Innovative content also requires innovative teaching methods that integrate society and industry. The traditional lecture is not the only useful method when it comes to preparing students for digitization processes in business and society. Here are some examples of innovative teaching methods in information science.

We have replaced some of the lectures with a blended learning. Here, the lecture content has been recorded on video. Documents are made available to the students for reading. Students are given multiple-choice tests for a self-check whether they have understood the content. Every week there are also exercises that are solved in working groups. The working groups are supervised by student tutors. In the plenary session, which replaces the lecture, students' questions on the respective topic are discussed. This allows to integrate basic knowledge with current issues relevant for the topics discussed, combining the grounding education with state-of-the-art information.

In addition to classical seminars and lectures, there are projects. Here, students work in working groups on projects on a topic that is also relevant in research and industry (see also Trinh, Nguyen, and Minh 2017). An example: In the project "Opinion Mining" there were working groups on these topics: systematic evaluation of sentiment analysis systems, implementation of a system for automatic trend detection in Twitter, implementation of a system for the sentiment analysis of Amazon reviews of

technical products, automatic creation of a lexicon for sentiment analysis of the German language, creation of a research map on the topic, and automatic detection of opinion spam in Amazon reviews.

The form of examination in these projects is the learning portfolio. Although they work in workgroups and produce a workgroup report, each student writes an individual report about her learning success and contribution to the project. A special project is the participation in a research competition, such as a Shared Task in NLP. This is where students come into contact with research groups, where they can compete and exchange ideas with them. At the same time, they not only learn to master a complex programming task, but also to document their work professionally and scientifically.

The teaching conference is another teaching method in information science: The task for the students is to write a scientific paper on an information science topic and to present it at the end of the semester in a "conference". The first version of the scientific paper is uploaded anonymously via Moodle, just like in a scientific conference. Each student then writes a peer review report of three different papers, not knowing the author. On the basis of the reports, the papers are then revised and presented at the end of the semester. The examination of the course is made up of the assessment of the reports written by the student, the paper in the updated version, and the presentation.

Conclusion

Information science is a discipline that cannot be explained in a few words, especially because of its interdisciplinary nature. However, this supposed weakness of the subject is actually its strength: the students receive a broad basic education with a technological focus and can be employed in many different ways. The subject itself can and must regularly revise its contents and adapt them to the requirements of society and business. This (and a strong focus on innovative technology) creates a need for innovation and constant discussion. Visionary scientific work is necessary in order to advance the subject and to be able to survive in the future.

References

Agogo D., and Hess T.J. (2018). Scale Development Using Twitter Data: Applying Contemporary Natural Language Processing Methods in IS Research. In: Deokar A., Gupta A., Iyer L., Jones M. (eds) *Analytics and Data Science. Annals of Information Systems*. Springer, Cham

Erdmann, E., Johnson Coenen, K., Link, L., Longbotham, J., Nickl, M., Reuther, U., Schöffer, A., and Siegel, M. (2017). *Regelbasiertes Schreiben. Englisch für deutschsprachige Autoren,* 2. edition; Tekom; Stuttgart.

Lee, Roger (ed.) (2018). *Computer and Information Science*. Studies in Computational Intelligence, 719.

Muthig, J., & Schäflein-Armbruster, R. (2008). *Funktionsdesign® – methodische Entwicklung von Standards. Standardisierungsmethoden für die Technische Dokumentation*, tekom-Hochschulschriften, 16.

Porsiel, Jörg (2017): *Machine Translation. What Language Professionals Need to Know*. Berlin: BDÜ Weiterbildungs- und Fachverlagsgesellschaft mbH.

Siegel, Melanie und Christian Lieske (2015). Beitrag der Sprachtechnologie zur Barrierefreiheit: Unterstützung für Leichte Sprache. In: *Trans-kom* 8(1), 40–78.

Trinh, Son, Luu Nguyen, and Minh Vo (2017). Combining Lexicon-Based and Learning-Based Methods for Sentiment Analysis for Product Reviews in Vietnamese Language. In: (ed. R. Lee) *Proceedings of the proceedings of the 16th IEEE/ACIS International Conference on Computer and Information Science (ICIS).* Springer, Cham.