

University of Cologne
Faculty of Management, Economics and Social Sciences
Information Systems Area



Department of Information Systems and Information Management

Bachelor Wirtschaftsinformatik Bachelorseminar

Summer Term 2017

Organizing department:
Department of Information Systems and Information Management (Prof. Dr. Schoder)

Contact person:
Laura Menth (menth@wim.uni-koeln.de)
Office: Pohlighaus, Room 405; Office hours: by appointment (via E-Mail)

Please register for the course via ILIAS. Registrations can be made until **02 April 2017** (https://www.ilias.uni-koeln.de/ilias/goto_uk_crs_2095975.html). Please respond to the survey in the ILIAS course to submit your application.

1. Course Description

The major goals of the seminar are to get a glimpse of problems and possible solution approaches in the respective subject area, the independent development of a solution for a specific assignment, to get familiar with academic research including literature search, selection and administration and to learn about presentation techniques. Accordingly, the students learn how to conduct scientific research and how to present the results. Evaluation: students submit a scientific work on one of the topics from subject areas below (approx. 15 pages and subject to further requirements) and present the results in a concluding presentation.

We offer interesting and relevant problems from the following subject areas and research projects:

- *Digital Radio Customized*: Like other industries, radio broadcasting is affected by a digital transformation and is currently undergoing major changes. Whereas music and film industry are already advanced in the process of digital transformation due to pent-up pressure to rearrange their business models, radio broadcasters had experienced less pressure so far, but now try to leap up by providing their content on mobile devices as a first step towards digital transformation. Broadcasting organizations are facing increasing competition for listeners' attention by new music streaming services and new multimedia platforms. Spotify had 100 million active users in mid 2016. The more popular music and video streaming services become, the less popular radio consumption becomes in relation, as listeners' time and attention are limited resources. Broadcasters react by inventing new channels to distribute their content, and one of the most charming approaches is to automatically play recommended items that reflect the interests of the listener, just like Spotify does. A good example of an individualized radio app is NPR One. However, radio recommendation is different from music recommendation, as radio is a mixture of diverse formats. Our research goal is to understand what makes a good customized radio program, how radio recommender systems should be built, and how radio broadcaster should position themselves in the future.

- *Enriching product search with information from reviews:* Search functionality in web shops is limited today. For example, Katie is in search of a laptop. She has some requirements – it should be 13”, with a 1TB hard disk, but most of all she wants a laptop which does not make any noise, as the fan of the previous notebook was annoying her. In the web shop, she can select the size of the laptop and the storage, but she cannot search for laptops with a silent fan. For all laptops that match her criteria of size and storage, she has to read all the reviews and find out what users wrote about the fan, which is very time-consuming. Current search functionalities allow users to search and filter for a limited set of facts about the product, such as storage size. All other aspects, among them decision-critical ones, are not accessible in the product search. Whereas techniques exist to extract product features from textual data (product feature extraction, sentiment analysis, opinion mining), little is known about how to integrate these extracted features into a product search. We assess how such an enriched product search could look like by considering data quality problems, user interaction challenges, and other effects like negativity bias etc.
- *Enriching product search with information from reviews:* Search functionality in web shops is limited today. For example, Katie is in search of a laptop. She has some requirements – it should be 13”, with a 1TB hard disk, but most of all she wants a laptop which does not make any noise, as the fan of the previous notebook was annoying her. In the web shop, she can select the size of the laptop and the storage, but she cannot search for laptops with a silent fan. For all laptops that match her criteria of size and storage, she has to read all the reviews and find out what users wrote about the fan, which is very time-consuming. Current search functionalities allow users to search and filter for a limited set of facts about the product, such as storage size. All other aspects, among them decision-critical ones, are not accessible in the product search. Whereas techniques exist to extract product features from textual data (product feature extraction, sentiment analysis, opinion mining), little is known about how to integrate these extracted features into a product search. We assess how such an enriched product search could look like by considering data quality problems, user interaction challenges, and other effects like negativity bias etc.
- *Media Mass Customization and Electronic Commerce:* Individualized customer communication has well-known, measureable benefits for both customers and companies. In the field of Media Mass Customization and Electronic Commerce our department does research on methods and develops information systems, which can help assemble and distribute individualized information and documents. The developed methods and information systems can generate - among others - customized catalogues and newspapers. Underlying processes shall proceed fully automated and cost-efficient for a high number of users.
- *Information Quality in Social Information Systems.* Social Information Systems (Social IS), also referred to as “social media” or “social software”, have become extremely popular over the last decade, both within and outside of formal organizations. Broadly speaking, constitutive characteristics of this type of IS are the ability of users to produce and consume content (hence the term “prosumers”); a focus on digitally enabling social activities; a higher degree of openness towards voluntary users, compared to traditional IS, and hence subscribed user bases of thousands or millions of users; allowing for co-governance, to a greater or lesser extent. Given these conditions, the quality of the information in so-called user-generated content (UGC) in social IS has often been questioned. One can find examples for social IS with relatively high quality of content and with low quality of content. There are also differences in information quality (IQ) within the same social IS, for instance, between different articles on Wikipedia. Among others, the research project aims to understand how different levels of IQ emerge in social IS, that is, under which sociotechnical circumstances content of high IQ is created and when not; what communities and providers of social IS do, for instance, in terms of technical features to improve IQ; and what consequences of different levels of IQ are.

- *Social Media Usage*: Digitalization is increasingly changing user behavior of Web Services. Social Media Platforms, such as Twitter, Facebook, Instagram or Snapchat are widely used. Therefore it is necessary to understand why users act like they do and which factors affect that behavior and usage. Different parameters are responsible for user behavior and acceptance: intentions, motives and functions. Platforms and Services change and are updated all the time. Instagram for example issues a kind of latest updates like InstaStorys or InstaLive. Snapchat published the Spectacles Sun and Snap Glasses. Social Media is habitualized in our lives. A view at the platform evolution helps to understand the development of usage. Furthermore, it is examined whether users have variable Social Media roles for different services. So the focus of this subject area is how Social Media Services change user behavior and vice versa.
- *Social Media based (Computational) Social Sciences*: Social Media Content is a growing phenomenon. The number of user's expands every day. The content produced and published by users offers new opportunities to study human behavior and development. Hedonic Media products like Movies, TV-Series and Books as a mirror of human behavior, culture and social structures reflect sociocultural developments. User generated content focused on hedonic Media products, like Wikipedia articles about books movie or TV series as well as summary's, reviews and recommendations on webpages (like e.g. imdb), or forum or blog discusses about hedonic media products offer therefore (in combination with computational text mining and network analysis methods) new opportunities to analyses social phenomena and developments. So the focus of this subject area is what Social Media content particularly about hedonic Media products (in combination with new computational methods) can tell us about sociocultural phenomena's and developments?
- *Social Media Analysis and Datafication*: Social Media Content is a growing phenomenon. The number of users expands every day. This has been attracting the attention of different communities interested in analyzing its content. Companies and media organizations are increasingly seeking ways to mine User Generated Content for information about what people need and feel about their products and services. Text Mining is popular, but a qualitative content analysis with big data is challenging. While there has been a lot of research on how sentiments are evaluate, to examine customers' needs has been much less studied. Conventional research methods are time and cost intensive. Big Data has the potential to supplement classic methods. The focus of this subject area is Needmining.
- *Predicting churners using data mining / machine learning*: Data mining / machine learning is used to obtain knowledge from data. This helps to solve different problems. One interesting problem that occurs in different industries, is to predict customers which will most likely end their contract in the near future – this is also called “to churn”. There are different techniques used to do this in practice and research. Most commonly used are logistic regression, decision trees, neural networks, or survival analysis. In our project we are predicting churners in the energy sector using these four techniques. Here it is interesting to find out if the results can be generalized over different industries, why different techniques lead to different results on different data sets, or which techniques perform best.
- *Automation of data mining processes*: There are different models which describe the process of data mining. One of them is the CRISP-DM model which consists of 6 phases. First of all it is important to become familiar with the business context and the available data. Secondly, the data preparation phase is necessary to create a data mining table which can then be used for modeling using e.g. a logistic regression or a survival analysis. Then, the model has to be evaluated before it can finally be deployed. For some of these steps or sub-steps a level of automation is already possible but for others human interaction is still required. All these steps have to be performed by data scientists which are recently highly demanded. When looking at the diverse knowledge which is required to perform all data mining tasks, it becomes obvious that this is difficult to find united in one person. Therefore, it becomes interesting how much of the data mining process can already be done by machines and which knowledge is then still be needed to perform the complete data mining process? Do these people have to be data scientists or are less skilled workers good enough?

- *Human-Machine Symbiosis*. Latest developments in artificial intelligence, big data, and machine learning produce new opportunities for full or partial automation through information technology (IT) that match or even exceed humans' performance in a range of activities, including ones that require cognitive capabilities. Recent examples include financial high-frequency trading, healthcare, or human resource management. Nevertheless, although replacement of human work is to some extent technically feasible, contexts exist in which regulatory constraints, social factors, or complexity inhibit full automation and, thus, require human involvement to some extent. Especially, cases in which human's and machine's distinct capabilities in combination perform operations more effectively than in isolation, require a careful design of IT artifacts that consider the interplay of human and machine and its symbiotic co-evolution over time. Focusing on the development of an intelligent decision support system for ULD resource allocation and its organizational introduction at a leading service provider for ULD management, this research area aims to distill and formulate effective design knowledge for Human-Machine Symbiosis.

2. Additional Information

Additional information regarding this course can be found on the homepages of the Department of Information Systems and Information Management (www.wim.uni-koeln.de) and in ILIAS (https://www.ilias.uni-koeln.de/ilias/goto_uk_crs_1920134.html).