

Design Principles for Competence-based Recommender Systems

Valerio Bellandi, Paolo Ceravolo, Fulvio Frati, Jonatan Maggesi
SESAR Lab - Dipartimento di Informatica
Università degli Studi di Milano, Italy
Email: {name.surname}@unimi.it

Gabriela Waldhart, Isabella Seeber
Dep. of Information Systems, Production and Logistics
Leopold-Franzens-University of Innsbruck, Austria
Email: {name.surname}@uibk.ac.at

Abstract—In this paper we analyze the principal preconditions and limitations for designing a competence-based Recommender System. In detail this analysis is contextualized in the ARISTOTELE European project. In the second part of the paper an architectural view is proposed taking in consideration the objective to propose standard and non-standard suggestions. This solution will permit to insert serendipity approaches into classical solutions.

Keywords-Recommender System; Competence Management; Suggestions;

I. INTRODUCTION

Most popular Recommender Systems (RS) are based on the assumption that similar users have similar interests. However, in the context of Competence Management System (CompMS), this assumption does not hold directly. In fact, the proposed suggestion of how a user can improve his curriculum cannot be directly referred to users having similar curricula. Typically, the process guiding a CompMS starts with identification of the tasks to be performed in the organization; then, it performs a “gap analysis” to identify which competences are currently available and which competences the users should acquire; finally, it uses the results of the gap analysis to identify the competence development to be implemented in the organization to fill the learning gaps [DM06]. In other words, CompMS are required to consider the organization as a whole, comparing global objectives with individual competences. On the other hand, a RS makes in general personalized suggestions by extracting knowledge from the previous user interactions with the system [RRS11], considering the history of all users and eventually local contexts [AT11], but without involving information related to the goals orienting the system. This difference suggests that integrating a RS in a CompMS requires the definition of specific design principles to be considered in constructing the algorithms guiding the recommendation process. Within this context, this paper discusses preconditions and limitations for methodologies exploiting collaborative networks within organizations to enable competence-based recommendations for learning measures. It analyzes organizational scenarios developed in two organizations emerging from the ARISTOTELE [ARI] project with the goal to derive design principles guiding the definition of a new algorithm for competence-based recommender system.

A. The ARISTOTELE Project

In the modern so called “knowledge society” organizations are step by step realizing that shifting the relevance from tangible to intangible assets increase their competitiveness on the market. Information technology is deemed of being important to deliver strategic value [SL02], however their adoption and assimilation into the organization might inhibit positive and negative consequences that need to be considered carefully [Pha04]. It is important to enlarge the focus by considering complex environments that integrate models, processes and technologies with organizational aspects in a more systemic approach. The ARISTOTELE research project [ARI] aims at relating the learning process to the organizational ones as well as to the innovation process management. In particular, in the organizational contexts three kinds of processes are traditionally identified: organizational processes (marketing & communication, human resources management, business), learning processes (group training sessions), and social collaboration processes (spontaneous formation of groups within the organization).

II. RELATED WORK

RS reuse knowledge that was made available by other users and by basing suggestions on users’ preferences. Relying on the identification of small subsets of items that can have potential of being of interest for users, RS can decrease information overload.

In literature, four basic types of recommender systems have been identified: content-based, collaborative, knowledge-based, and hybrid recommender systems [JZFF11]. Moreover, in literature four constructs that are relevant for the design of recommender systems have been distinguished [AKS09]: *i) Homophily*: similarity between sources and recipients, *ii) Tie strength*: intensity of the relationship between the recipient and source, *iii) Trust*: cognitive and affective dimension of a trust relationship between recipient and source, and *iv) Social capital*: a source’s reputation or opinion leadership.

Homophily describes the tendency of individuals to associate and bond with others that are similar to them. It can be achieved by recording a group’s consumption patterns and comparing these patterns. However, cases have shown that recommendations are not purposeful when recommendations

are too similar (e.g., recommendation of similar news articles covering the same story) [AT05].

Tie strength can be calculated in different ways such as relationship's duration, interaction frequency, or feeling of closeness. The investigation of interpersonal ties also highlights strength of weak ties, as studied by Granovetter in [Gra73]. In his study, he pointed out that weak ties generate opportunities for individuals and support their integration into communities.

Trust in a recommended source represents a further indicator that determines the likelihood of accepting a recommendation or not [AKS09].

Finally, social capital refers in this context to the user's position in the social network and the user's ability to influence others. Social network analysis (SNA) techniques can be applied to determine e.g., centrality of users [AKS09].

Arazy et al. [AKS10] have shown in their work that homophily positively influences the users' willingness to accept an advice or recommendation. This relationship is mediated by tie strength and trustworthiness. Further, competence-based networks (social networks based on domain expertise) have more positive influence on willingness to accept advice than benevolence-based social networks (social networks based on friendship). The study however, could not show a substantial positive correlation between tie strength (measured by duration and frequency) and willingness to accept advice [AKS10]. In order to give recommendations, the system needs to have some kind of data input. Felfernig defined in his work [FB08] four typical information sources, comprising *i*) the user herself, *ii*) other peer users of the system *iii*) information about the items being recommended, and *iv*) information about the domain of recommendation.

A. Diversity, Novelty and Serendipity

RS have reached in the last years a good level of accuracy; in particular, recommender systems adopting collaborative filtering have succeeded in exploiting the huge number of subscribers, providing (after a training period) recommendations in line with the users' interests, based on the ratings (explicit or implicit) of other like-minded users [AT05]. RSs, however, still fail in discovering users latent interests [LCX⁺12]: they often suggest items that, although accurately tailored on the users' past behavior, do not consider appropriately how the users' preferences can evolve in the future. When a user profile is very focused in terms of content experience, the user is provided with recommendations that she likes, but to which she would be anyway exposed through other channels.

It is important to state the difference between diversity, novelty and serendipity in the context of recommender systems. The term diversity can have several meanings, but probably the most common one refers to how diverse recommendations proposed by a recommender system are:

for example, if a movie recommender system proposes just one genre of movie (e.g. "action movies"), the diversity of the suggestions provided is very low, regardless to how such recommendations are new for the user. In particular, novelty refers to how original is a recommendation respect to what the user does not deem to be obvious recommendation; if the user already knows the proposed items (for example she already knows the movie proposed) the recommendation is not novel. Serendipity, on the other side, adds another element: it represents the ability of a RS to propose non-obvious, surprising, suggestions. An implicit difference between novelty and serendipity is the success of the recommendation: in literature the concept of serendipity is in fact usually associated to a surprising and interesting suggestion, in other words, not only the suggestion should be apparently far from the current user's preferences, but the user should also like it; the concept of novelty, instead, does not usually include this last assumption.

B. Knowledge-sharing and Solidity

Michael Polanyi stated that a person can know more than she can tell. This quote and Polanyi's research underline the concept of tacit (context-specific, personal, hard to formalize and communicate) and explicit (codified, transmittable in formal, systematic language) knowledge [Pol66]. Nonaka and Takeuchi (1995) extended Polanyi's work by including distinctive features that are typical for either tacit or explicit knowledge. For example, tacit knowledge is subjective, created from experience and needs the "here and now" for sharing [NT95].

Communities are believed to be able to create channels for the transfer of tacit knowledge (personalization) and explicit knowledge (codification) among individuals within or across organizations. Also, through constant interaction, communities generate knowledge and reinforce and renew themselves [WS00]. Communities provide a way for managing knowledge as an asset by allowing free and creative flow of knowledge and experiences and hence have positive effects on the overall organizational performance [WS00]. The existence of communities is especially worthwhile when members can draw expertise from different sources, even outside their, for instance, project team. Hence, the exchange of knowledge of critical areas can be improved by organization as well as maintaining or even improving the competitive position in the market.

A further implication for knowledge sharing is related to the phenomenon of organizational change based on altered knowledge. This means that a human mind or a collective can reconfigure meaning of things within minutes, days, or months. In contrast to that, explicit knowledge stocks within large organizational information systems take much more time to be updated, since semantics relations need to be reconfigured as well. Hence, the question arises on how to pick stable knowledge, e.g., in order to make it explicit in a

Table I
RECOMMENDER SYSTEMS REQUIREMENTS

ID	Functionality	ID	Functionality
UR01	Focus on work	UR12	Sharing objectives
UR02	Assign objectives	UR13	Goals and needed competences
UR03	Support for Collaborative Network	UR14	Suggesting people
UR04	Performance and Collaboration Indicator	UR15	Update of organizational knowledge
UR05	Common layer for semantic correlation	UR16	Classifying Organizational knowledge
UR06	Knowledge collection, tracking, and presentation	UR17	Correlating content of organizational knowledge
UR07	Tracking tasks	UR18	Expressing doubts and questions
UR08	Improved Search	UR19	Agile activities planning
UR09	Learning	UR20	Workbench
UR10	Exploiting successful experience	UR21	Suggestions and problem solving
UR11	Sharing successful experience	UR22	Instant questions

knowledge base [Tuo99]. When stable knowledge stocks get further constructed and extended with additional knowledge, this knowledge needs to be valid and justified in order to be interpreted as stable. This phenomenon has been denominated as knowledge sedimentation. Knowledge maturing, however, starts off from a collective, where individuals are seen as participants that allow advancing, or knowledge maturing, on the collective level in a goal-oriented way. Moreover, also artifacts (e.g., documents, videos, wiki pages, etc.) and social facts (e.g., rules, processes, collaborative practices that have been evolved in an organization) are regarded as manifestations of knowledge and are therefore also in focus of knowledge maturing.

Literature indicates that information sources which are considered for recommendations should also contribute to novelty, diversity or serendipity. Moreover, RSs need to consider that knowledge is not necessarily stable and might evolve over time. Hence, some kind of evaluation, possibly implementable with a knowledge assessment service, is required which improves the accuracy of recommendations.

III. ANALYSIS OF RECOMMENDER SYSTEMS SCENARIOS

This Section presents the results of the analysis of the descriptions of requirements for the ARISTOTELE prototype platform. In this regard, the platform should be capable of supporting *i*) creation and execution of personalized learning activities, *ii*) innovation building and management processes, and *iii*) collaborative knowledge exploitation and acquisition during daily worker activities [ARI], taking into consideration the concepts described in Section II.

Table I exposes the User Requirements (UR) developed and defined in the ARISTOTELE project.

In order to derive requirements for RS, the user requirements defined were analyzed and summarized in Table I. In this context, inductively defined codes dimensions of *Competence Based Management*, *Activity and Task Management*, *Collaboration*, and *Knowledge* have been developed. They refer to recommending functionalities and situations on which the RS focus. Relevance to RS was determined by assessing whether a UR corresponds to a methodology question and can therefore fully or partly answer the functionality request. Methodology questions in this regard relate to recommendation results, which the ARISTOTELE RS will generate. Each dimension will be described with the

support of information derived from coded scenarios. Based on this information it was possible to generate possible preconditions and limitations. Therefore, each dimension will include a short description as well as preconditions and limitations. Results will then be presented and categorized along the above mentioned dimensions.

A. Competence Based Management

The dimension refers to the understanding that ARISTOTELE allows the representation and reasoning of related concepts stored in ARISTOTELE models. This means that contextual information, such as a person who has a set of competences and works on a specific task, can be reused to express relations. However, also users can extend the knowledge base by instances (e.g., rate a document with five stars) of concepts by rating, tagging, commenting contents or people. Hence, it is possible to recommend expert advice or make informed team creation decisions by using competence profiles but also suggest relevant and accurate items (e.g., a document).

Precondition 1) Competence Profiles. One central requirement for companies that will adopt ARISTOTELE is the availability of competence profiles and therefore the awareness on which competences are required for organizational tasks. Furthermore, competence profiles need to be accessible and updatable by the ARISTOTELE platform. Also, information is required to what extent a person is capable of fulfilling the competence, which relates to the notion of proficiency. This means that a beginner in Java programming, who has a Java competence level of 1, requires an expert on the topic to improve her level, who relates to each Java programmer with a competence of, e.g., 5. *Limitation:* In case competence profiles and related competence levels are not available, the recommender will not be able to suggest people related information (i.e. competences, experts).

Precondition 2) Tagging/Annotating. The ARISTOTELE platform needs to provide tagging or annotation functionalities for artifacts, people, and processes. Additionally, users need to have the corresponding rights and willingness to annotate content that is available. Based on the availability of tags or keywords, content-based recommendations are technically possible. *Limitation:* In the case that users cannot tag or annotate contents or other types of resources, it is likely that recommender results will not reach high accuracy. This will also be dependent on the availability of other contextual information.

B. Activity and Task Management

The management of performed activities and tasks on an individual or team level was frequently expressed in ARISTOTELE scenarios. In this regard, ARISTOTELE allows a comprehensive workflow management, where people can define and assign objects, which can further be refined by recipients and be monitored. In this context, system

notifications have been mentioned as a means to inform status of work progress and monitoring. Moreover, workers talked about daily challenges at the work place (e.g., information overflow, time coordination, etc.) that can be improved with the help of accurate resource recommendations. Communication within ARISTOTELE tools needs to support community work by recommending appropriate groups that might hold necessary expertise or resources. Again, recommendations will be widely based on content available in the ARISTOTELE models, which are rdf-based databases holding any kind of information deemed important for ARISTOTELE tools and services, and hence will comprise contextual information to a situation. Furthermore, information on activities and related sources needs to be stored in activity logs or history logs so that it is possible to answer problems that were successfully answered earlier. The ARISTOTELE platform needs to be able to merge information from various systems, such as social network, ERP, CRM, etc. including the visualization of the progress of a specific activity or task. In this context, it is also important that information on successfully finished tasks can be expressed in the ARISTOTELE models and exploited by the RS.

Precondition 3) Activity/Task Monitoring. In order to derive activity/task information from other enterprise systems, these systems need to provide interfaces to the ARISTOTELE platform to extract necessary information about tasks/activities the users are involved in. Also organizations need to confirm to reuse this information to improve their business processes. Moreover, the definition of activity/task objectives needs to be expressed in the system by people, in order to reason upon successfully performed tasks based on strategy statements. Other important information delivered in activity or task logs related to timestamp, actor(s), resource, channel used, etc. *Limitation:* In the case that activity/task information is not available for recommendation, methodologies might be restricted, since main behavior of set actions by the user is difficult to analyze automatically (e.g., process mining). Moreover, strategy or goal definitions of employees are important to relate a strategy with a task and receive whether the task was accomplished successfully.

C. Collaboration

The dimension Collaboration provides in most cases the basic layer of communication and exchange of content or knowledge. Scenarios have included possible channels such as instant messaging, private message in collaborative systems, discussion boards, shared workspaces, forums, meetings arrangement, or workflow management systems. As already mentioned in Section III-A, the ARISTOTELE platform needs to allow the use of multiple communication channels (e.g., instant messaging, private messages, discussion boards, forum) and annotation of used content by rating, tagging, or giving feedback or comments.

Precondition 4) Willingness to change communication practices. Due to the fact that ARISTOTELE will integrate various possibilities within the platform to communicate with others (e.g., messaging, forum, discussion groups, mailing, etc.) it is necessary that the organizational culture, and therefore people, accept and are willing to use ARISTOTELE communication offers. *Limitation:* Even though ARISTOTELE strives to offer interfaces to common communication tools, there exists a risk that in some cases the exploitation of communication information is not possible. This might endanger the necessary amount of context information required to perform accurate recommendations.

D. Knowledge

Of central importance is the handling of explicit/codified knowledge, as well as the appropriate support of people to exchange implicit knowledge successfully. Moreover, the ARISTOTELE platform needs to support activities that are related to the update and sharing of (organizational) knowledge which might also be stored outside of ARISTOTELE (e.g., process description). The platform strives to deal with information overload by recommending relevant sources (e.g., people, documents, activities, etc.) during work activities. A challenge in this regard is that sources, comprising codified knowledge, are diverse and difficult to merge within the ARISTOTELE models. Hence, it is important to find ways how to determine trustable and justifiable knowledge sources. Moreover, ARISTOTELE can only be successfully implemented if it achieves critical mass, i.e. a sufficient number of people participate and a large enough network is available.

Precondition 5) Migration of existing knowledge base. In order to query knowledge assets such as e-resources, people, processes, there is a need that the existing knowledge base gets migrated into ARISTOTELE or that employees of an organization are tasked to create an initial knowledge base on which ARISTOTELE can work on. In the first case the concept of ontology matching could improve that same things from diverse systems are described as one thing. *Limitation:* In the case ARISTOTELE does not have a knowledge base, which also includes existing information on people, resources, and processes, the platform might be deemed to not be able to solve the start-up problem and create sufficient positive network effects.

Precondition 6) Context Information. In order to derive contextual information (e.g., customer number, purchased items, etc.) ARISTOTELE needs to have interfaces to other enterprise systems, such as CRM, HRM, ERP or other applications that hold information that relates to process information, customer and business information, communication events, or the like. It is likely that information needs to be mapped among the different sources (as mentioned above) in the case that they deliver the same semantic information (e.g., worker, client, user, etc.). *Limitation:* In the case that

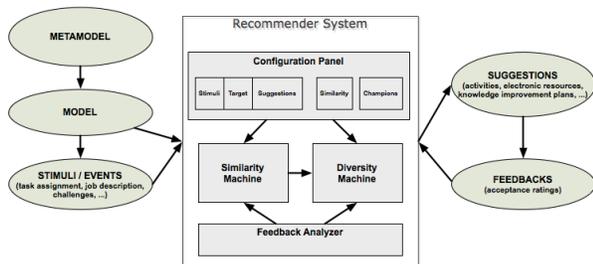


Figure 1. RS Components and sub-services.

ARISTOTELE cannot draw on information stored elsewhere in the organization, content-based recommendation will be highly limited.

The stated preconditions and limitations were input for the development of the ARISTOTELE recommender system, the competence-based recommender. This RS differs from other common RS by taking as central concept the competences and working experiences of all members of an organizations. To this aim, the ARISTOTELE RS was designed to work as a *competence-based* RS, including it in the more general family of hybrid recommenders [JZFF11].

IV. ARISTOTELE COMPETENCE-BASED RECOMMENDER SYSTEM

The RS is an important core service of the ARISTOTELE framework; it has been designed to be triggered by a specific *stimulus*, giving not only suggestions on activities and learning plans correlated to the subject, but also suggesting a set of alternative objects that, at a first sight, could seem completely unrelated, but, looking at past experiences, they could have a positive influence to the knowledge improvement process. This way the RS can introduce novel and unexpected knowledge fields in the ordinary business process. Furthermore, users can specify a specific *target* over which the RS will be configured and returned learning suggestions should be applied.

As stated above, the RS takes as inputs the knowledge base, the target, and the stimuli, and it returns a set of suggested activities that could improve the overall knowledge level [DCBF11].

The RS is organized in four main sub services (see Fig. 1). The four sub services can work independently, but need to cooperate in order to achieve the maximum reliability of suggestions given, and better results in terms of knowledge improvement. They rely on the *RS metamodel* depicted in Fig. 2 and described in [DCBF11]. It is important to note that the four components share the same source data space, clustered exploiting specific similarity functions.

The *Configuration Panel* defines the specific settings of any single instantiation of the RS. The objects analyzed within the RS are identified in the metamodel but the specific modeling of these concepts, with properties and relations

expressed, depends on the model adopted. The information processed by the RS involves all information available in the ARISTOTELE models but the input triggering the analysis is given by the internal or external stimuli generated within the process execution. From the point of view of the implemented data structure, a Stimulus is represented by a data object or by a pattern of objects of the metamodel. These stimuli represent new events or knowledge that encourage the organization must face. The structure of the stimuli that are accepted by the RS can be configured in the Configuration Panel. The notion of Target is also important to configure the RS. The Target defines the data objects to be compared with the stimuli. Using this notion it is possible to exploit the RS ranging from the single user to a team or to the entire enterprise. The model is also required in order to define the similarity functions implemented. The sub-services of the RS generate suggestions that are lists of items (represented in term of data objects of the metamodel) that are considered *close* to the stimulus under analysis. Suggestions can typically be activities, knowledge improvement plans, or human resources, according to the information in the Configuration Panel of a tool. In particular, it is necessary to define how to rank suggestions computing the similarity measure with respect to the stimulus. The *Similarity Machine (SM)* organizes the objects in the model in a similarity space. When a stimulus triggers the RS the SM can identify objects related to this stimulus and that, by consequence, the organization can use for reacting. The profile of the user (her history) is also adopted in order to cluster the similarity space. The *Diversity Machine (DM)* is searching out of the boundaries of the similarity space, in order to increase the variability of the suggestions proposed. The DM is strictly related to the notion of expert to be followed in order to identify sequences of objects that are not similar to stimuli. In fact, because of their description, they can be related since they distinguish the activity of an expert that succeeded in achieving goals related to the stimuli. The idea is that if an expert was successful in reacting to similar stimuli, it is meaningful to analyze the activities she performed, even if they are not similar to the stimuli under analysis. Finally, the *Feedback Analyzer* is tracking the acceptance of the suggestions proposed by the SM and the DM. Suggestions refused, not adopted, or not providing successful results are signaled to the SM and DM that use this information to refine its definitions of similarity.

V. CONCLUSION

In this paper we investigated important preconditions and limitations that are considered to have impacts on the design and success of a competence-based Recommender System, taking into consideration a specific scenario expressed in the European project ARISTOTELE. On the one hand, the discussed preconditions, can be directly used as input factors for the recommender system and therefore represent

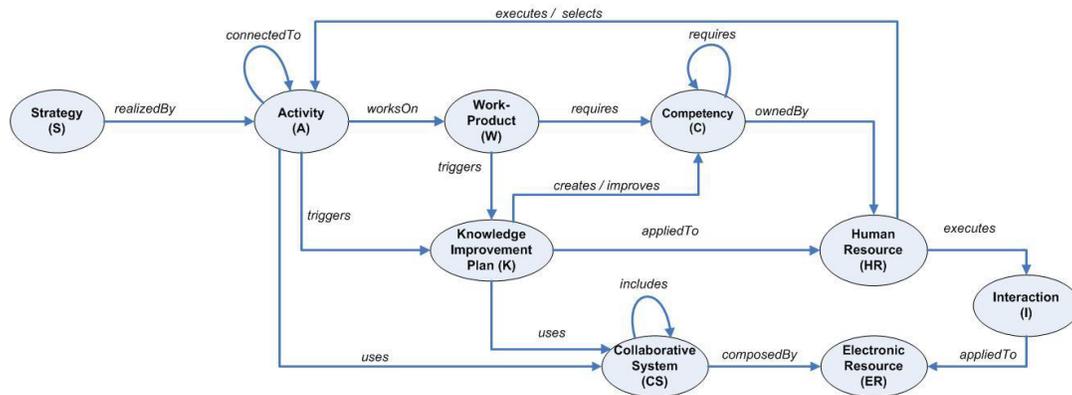


Figure 2. RS Metamodel.

conditions. On the other hand, preconditions can inform the design of recommender systems from a theoretical point of view with the goal to improve the acceptance of recommender results by the user. The transformation of presented requirements, preconditions and limitations was then demonstrated along the design of the structure of the ARISTOTELE Recommender System and its components, the similarity machine, diversity machine and feedback analyzer.

ACKNOWLEDGMENT

This research is partially funded by the EC under the Project ARISTOTELE (Grant Agreement n. 257886).

REFERENCES

- [AKS09] O. Arazy, N. Kumar, and B. Shapira. Improving Social Recommender Systems. *IT Professional*, 11(4):38–44, 2009.
- [AKS10] O. Arazy, N. Kumar, and B. Shapira. A Theory-Driven Design Framework for Social Recommender Systems. *Information Systems*, 11(9):455–490, 2010.
- [ARI] ARISTOTELE project. <http://www.aristotele-ip.eu>.
- [AT05] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowledge and Data Engineering*, 17(6):734–749, 2005.
- [AT11] G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. *Recommender Systems Handbook*, pages 217–253, 2011.
- [DCBF11] E. Damiani, P. Ceravolo, V. Bellandi, and F. Frati. *CR2S : competency roadmap to strategy*, pages 115 – 122. Lecture notes in informatics ; P-182. Gesellschaft fr informatik e.V., Bonn, 2011.
- [DM06] F. Draganidis and G. Mentzas. Competency based management: a review of systems and approaches. *Information management & computer security*, 14(1):51–64, 2006.
- [FB08] A. Felfernig and R. Burke. Constraint-based recommender systems: technologies and research issues. In *Proceedings of the 10th international conference on Electronic commerce, ICEC '08*, pages 3:1–3:10, New York, NY, USA, 2008. ACM.
- [Gra73] M.S. Granovetter. The strength of weak ties. *American journal of sociology*, 78(6):1360–1380, 1973.
- [JZFF11] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich. *Recommender Systems: An Introduction*. Cambridge University Press, 2011.
- [LCX⁺12] Q. Liu, E. Chen, H. Xiong, C.H.Q Ding, and J. Chen. Enhancing Collaborative Filtering by User Interest Expansion via Personalized Ranking. *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(1):218, 2012.
- [NT95] I. Nonaka and H. Takeuchi. *The Knowledge-Creating Company: How Japanese Companies Create The Dynamics Of Innovation*. Oxford University Press, 1995.
- [Pha04] R. Phaál. Technology roadmapping - A planning framework for evolution and revolution. *Technological Forecasting and Social Change*, 71(1-2):5, 2004.
- [Pol66] M. Polanyi. *The Tacit Dimension*. Doubleday, 1966.
- [RRS11] F. Ricci, L. Rokach, and B. Shapira. Introduction to recommender systems handbook. *Recommender Systems Handbook*, pages 1–35, 2011.
- [SL02] U. Schultze and D.E. Leidner. Studying Knowledge Management in IS Research. *Management Information Systems*, 26(3):213, 2002.
- [Tuo99] I. Tuomi. Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory. *Journal of Management Information Systems*, 16(3), 1999.
- [WS00] E.C. Wenger and W.M. Snyder. Communities of practice: The organizational frontier. *Communities of practice: The organizational frontier*, 78(1):139–146, 2000.