#### Association for Information Systems

# AIS Electronic Library (AISeL)

Proceedings of the 2018 AIS SIGED International Conference on Information Systems Education and Research

SIGED: IAIM Conference

2018

# DESIGNING CHATBOTS FOR HIGHER EDUCATION PRACTICE

Jonas Sjöström Uppsala Universitet, jonas.sjostrom@im.uu.se

Nam Aghaee Department of Informatics Linnaeus University, nam.aghaee@lnu.se

Maritha Dahlin Department of Informatics and Media Uppsala University, maritha.dahlin@im.uu.se

Pär J. Ågerfalk Uppsala University, Sweden, par.agerfalk@im.uu.se

Follow this and additional works at: https://aisel.aisnet.org/siged2018

#### **Recommended Citation**

Sjöström, Jonas; Aghaee, Nam; Dahlin, Maritha; and Ågerfalk, Pär J., "DESIGNING CHATBOTS FOR HIGHER EDUCATION PRACTICE" (2018). *Proceedings of the 2018 AIS SIGED International Conference on Information Systems Education and Research*. 4. https://aisel.aisnet.org/siged2018/4

This material is brought to you by the SIGED: IAIM Conference at AIS Electronic Library (AISeL). It has been accepted for inclusion in Proceedings of the 2018 AIS SIGED International Conference on Information Systems Education and Research by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

## DESIGNING CHATBOTS FOR HIGHER EDUCATION PRACTICE

Jonas Sjöström Department of Informatics and Media Uppsala University jonas.sjostrom@im.uu.se

Nam Aghaee Department of Informatics Linnaeus University nam.aghaee@Inu.se

Maritha Dahlin Department of Informatics and Media Uppsala University maritha.dahlin@im.uu.se

Pär J. Ågerfalk Department of Informatics and Media Uppsala University par.agerfalk@im.uu.se

#### Abstract:

In this research–in–progress paper, we employ design science research to articulate design knowledge for chatbots in higher education practice. We conducted a literature review to factor previous research into the design process. In addition, we performed a content analysis of student e-mails and forum posts from four instances of a basic Java programming course. Drawing from literature and data, we present a conceptual architecture for chatbots in higher education, discuss its rationale, and provide a proof-of-concept implementation. We conclude with a discussion including tentative design recommendations and a plan for continued research.

Keywords: chatbot, conversational agent, learning, higher education, practice, design science research

### I. INTRODUCTION

Conversational agents – or *chatbots* – are now prevalent on the Internet. Chatbots.org report the existence of several hundred chatbots, tailored for different countries, languages, concepts, and use cases. There are various chatbot hosting companies specializing in, for instance, web-based bots, messenger bots, and application-based bots. Ismail (2017) accounts for the top 14 chatbot platforms of 2017, including Motion.ai, Converse.ai, QnA Maker, Octane AI, IBM Watson, Botsify, Chatfuel, Pandorabots, Microsoft Bot Framework, Wit.ai, and Semantic Machines. Several platforms support the Artificial Intelligence Markup Language (AIML) to store input patterns and output templates to be used by the bot's natural language processing (NLP) algorithms to infer responses from input statements. Chatbots serve various practical purposes, for instance, in education and entertainment, and as a means to interact with a software system or operating system. Known examples of chatbots include Microsoft's Cortana, Amazon's Alexa, Apple's Siri, and Google's Assistant. These bots are all designed to provide additional and richer ways for humans to interact with technology.

While chatbot technology has matured over time, there is still need for research on how chatbot technology can appropriately add value to human practice, including challenges in desiging effective dialogue between humans and bots (Fryer, Ainley, Thompson, Gibson, & Sherlock, 2017; Leonhardt, Tarouco, Vicari, Santos, & da Silva, 2007; Neves, Barros, & Hodges, 2006; Picard et al., 2004).

In higher education (HE), chatbots may facilitate better interactivity, sociability, and information acquisition by influencing educational flow to be more interactive and dynamic (AbuShawar & Atwell, 2015; Griol, Manuel Molina, & de Miguel, 2014). A potential advantage of using a chatbot in educational settings is the facilitation of instant retrieval of information for learners (Ghose & Barua, 2013). Chatbots have also been proposed as a means to estimate learning styles (Latham, Crockett, McLean, Edmonds, & O'Shea, 2010; Yun & Cho, 2003), and to harvest feedback in elearning environments (Lundqvist, Pursey, & Williams, 2013). Chatbots may even be part of the motivation for continued communication for educational purposes (Fryer et al., 2017). However, there is a need to factor in expectations from teachers and other stakeholders when designing bot technology in an education setting (Tamayo-Moreno & Perez-Marin, 2016). For instance, a chatbot should improve the communication between learner and teacher, rather than replace the teacher.

While there is a lot of research on chatbots in education, research that adopts a broader educational practice perspective (Orlikowski, 2007) is scarce. A practice approach entails taking into account multiple stakeholders in the learning situation, and investigating the emergence of social practices and stakeholder interactions when introducing chatbot technology. The aim of this study is therefore to explore the design and role of chatbots in HE practice. The contribution is a theoretically and empirically grounded conceptual architecture for chatbots in HE practice. We provide a set of tentative design considerations for chatbots in HE practice based on the process of designing the architecture and through a proof–of–concept implementation.

## II. RESEARCH APPROACH

We employ a design science research (DSR) approach (Hevner, 2007; Hevner, March, Park, & Ram, 2004). DSR entails designing, building and evaluating artifacts. The process consists of three interdependent cycles: design, rigor, and relevance. There is still a vivid discussion in the information systems field about how to abstract knowledge from DSR. In keeping with Hevner et al. (2004) and Gregor and Hevner (2013), our process results in an abstract model of an architecture for chatbots in HE practice, as well as an instantiation of that architecture; that is, a working piece of software that conforms to the abstract model. By evaluating our instantiation and documenting the design process, we aim at assessing our abstract model through an informed argument. We elaborate more on the rigor cycle, the relevance cycle, and evaluation below.

The DSR rigor cycle aims at drawing from the knowledge base to synthesize both foundations and methodologies. Previous research on designing and using chatbots in a learning context constitute the primary foundation in the current case. As the design process unfolded, additional foundations were identified, primarily regarding microservices architectures (Dragoni et al. 2017; Newman 2015). Additionally, general design thinking (e.g., Krippendorff, 2006) ingrained the design. Methodologies include both the DSR approach employed here, and specific methodological concepts that make sense in the current DSR context, such as, evaluation of chatbots.

To understand the foundations of HE chatbot design, we searched the Web of Science core collection for publications where either topic or title contained the word "chatbot" or the word "conversational agent", rendering 374 hits. A refined search where the title or topic also contained the words "education" or "learning" rendered 99 results. We read through each abstract to further narrow down our search. After the reading of abstracts, we ended up with 50 articles considered relevant to inform this work. In addition to those 50 articles, we identified another 13 articles while reading or from suggestions by peers. In total, 63 articles from the years 2001–2017 were identified – informing the rigor cycle both with foundations and contextually relevant methodologies. Furthermore, we factored in the literature on general design topics including software architecture and communication as shown later in the paper.

The DSR relevance cycle consisted of collecting and analyzing recorded queries (N=369) from student e-mails and discussion fora. We obtained the data from a programming course given four times 2015–2017. We conducted a content analysis of the collected data to articulate a learner-centric concept of conversational agents in HE. The authors are all active teachers with

substantial experience in teaching, providing the design process with a teacher perspective on chatbots in HE practice.

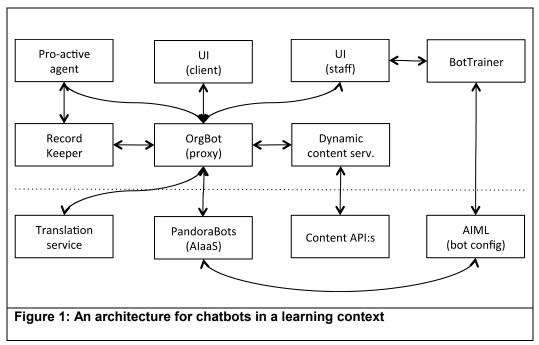
Design principles were articulated on basis of the outcomes from the rigor and relevance cycles. Drawing from the design principles, we identified the need for various subsystems and their interactions in the design. Those subsystems and their interactions resonated well with the design of a microservices architecture (Dragoni et al., 2017; Newman, 2015); that is, small, loosely coupled components that interact over the Internet via a lightweight protocol. The process was iterative, and this paper conveys the outcomes of the process so far articulated as design principles and a corresponding conceptual architecture for bots in a learning context.

Regarding evaluation, a fundamental aspect of DSR (Hevner et al., 2004; Venable, Pries-Heje, & Baskerville, 2016), we only provide limited results in this work–in–progress paper. The concluding discussion accounts for the evaluation conducted so far and planned future evaluation efforts.

## **III. ARCHITECTING CHATBOTS IN HIGHER EDUCATION**

Our literature review showed previous research focusing on architectures for chatbots in learning situations. For example, the Basilica architecture to support learning in collaborative learning environments; that is, settings with multiple learners and chatbots (Kumar & Rose, 2011). An additional example is the more technical approach to design architectures to support chatbots that evolve autonomously using fuzzy logic (Hassani, Nahvi, & Ahmadi, 2013). We did, however, not encounter any previous work addressing architectural design based on a view of HE as a learning practice, i.e., building on the situations and social interactions between teachers and learners.

**Figure 1** shows a conceptual architecture for chatbots in HE, drawing from the literature review as well as from the content analysis. The core of the architecture is the *OrgBot*, acting as a proxy with built-in logic for every interaction in the system. The OrgBot receives a question from learner through the *client UI* and passes it on to the AI as a service (AlaaS). The AlaaS reads the AIML configuration configuration to find a suitable response for the question. If there is no response, a 'fallback' feature is activated, asking the learner if the question should be sent to the teacher. If yes, the OrgBot forwards the question to the *staff UI*. All the steps in the process are recorded by the *RecordKeeper* component, which logs data for future analysis (e.g., for research or education improvement purposes).



The OrgBot and its interaction with the AlaaS correspond to the basic idea of chatbots: responding to questions. The content analysis showed that the chatbot should be able to answer common questions from students, for example, regarding course content, prerequisites and requirements for a course, and information regarding exams deadlines. See Table 1 for the types of interactions conceptualized based on the content analysis. From a student point of view, the chatbot may simplify access to important information. From a teacher's point of view, the chatbot's ability to answer such questions could reduce administrative overhead. Furthermore, the content analysis supports that we need to view the bot as support and supplement, rather than a substitute for the teacher or peer. The fallback function makes the bot act as a door to the teacher when needed. The chatbot should promote interaction between teachers and students, and between students. While many studies of bots focus on the interaction between a learner and a bot, we are interested in the bot as a part of a learning practice. There is a risk that bots decrease interaction between students and teachers, which may cause negative feelings and consequently affect learning negatively.

Category	Description			
Factual and general information	General queries regarding the examination, attendance and course literature. This category of the queries was explicitly and predominantly a single dialogue; i.e., the student asks a question and teacher responds with an answer.			
Course content information and peer interaction	Students need further guidance to proceed with the course content. A dialogue between student and teacher to get further information about course content.			
Self-statement or one-way information	One-way information given by students. The student does not require a reply from the teacher; i.e., they just need to notify the teacher about something (that may be inaccessible, incorrect, etc.). In this case the teacher might need to take action to correct it or to provide further information.			
Appeal	Student asks, e.g., an extended deadline, permission to skip a mandatory seminar, etc. Information exchange between student and teacher, which the teacher may confirm or manage by forwarding the information to a third party, e.g., course administrator.			

Table 1:	Student of	question	categories	generated	through	content analy	ysis

Sometimes, the AlaaS returns control codes to the OrgBot. If a return message includes such a control code, it is passed along to the dynamic content service, which fetches data using external content APIs, and injects them into the response message. This mechanism allows for control codes in the AIML definitions that translate into dynamic content at run-time. Returning to Table 1, the design supports fetching data about the syllabus, schedule, assignments, etc. By providing such content dynamically, the AIML definitions remain useful over time.

The *BotTrainer* subsystem allows teachers to provide answers to questions from the 'fallback' scenario above, while at the same time allowing for supervised training of the chatbot. The chatbot needs to evolve under the supervision of humans, to align its behavior with institutional norms, and to ascertain quality of responses. Even though it is technologically feasible to automatically train bots (e.g., Hassani et al., 2013), we believe that it is risky to do so in HE. In the Swedish context, for instance, bot actions in a University context are a form of agency comparable to exercising public authority.

The *pro-active agent* subsystem contains the logic to initiate conversations, such as quizzes, course evaluations, and reminders to log on to the course Intranet. The main idea is to facilitate a mechanism to promote student activity, in keeping with the idea of supportive accountability in eHealth (Mohr, Cuijpers, & Lehman, 2011). Conceptually, the subsystem exists in the conceptual architecture to allow for chatbot features beyond question-answer exchanges; essentially design

for mutability as suggested by Gregor & Jones (2007). Continued research may, for instance, include designs where data analytics methods are employed to identify when and how to trigger conversations with learners based on quiz results, inactivity, etc.

Finally, the idea of a *translation service*, which is still in its infancy, may prove very powerful to integrate a cloud translation service to facilitate interaction with international students, and provide them with an automatic translation of essential course information into their native language.

# IV. PROOF-OF-CONCEPT EVALUATION

The architecture discussed above was implemented into software both (1) as a proof-of-concept and (2) in preparation for continued evaluation in a Java course.

The architecture consists of a set of interacting subsystems, which resonates well with the design of a microservices architecture (Dragoni et al., 2017; Newman, 2015); that is, small, loosely coupled components that interact over the Internet via the REST lightweight protocol. A microservices architecture with its loose coupling allows different subsystems to be implemented using different programming languages, and in different server environments.

An essential design decision concerned which AlaaS to use. We decided to go with PandoraBots for two reasons. First, it allows us to define chatbot behavior using AIML, a 'de facto' standard that is reasonably convenient and works well with supervised learning. Second, PandoraBots was available as a cloud service via a REST API, making it easy to integrate into the architecture.

We implemented the software based on the idea that the bot should operate in communication channels that the students already use, to make it easily accessible and to avoid the risk of nonuse. We justify this idea by drawing on design thinking. Krippendorff (2006) refers to an artifact as part of an ecology of artifacts. A chatbot, according to this principle, should be easily accessible and function within the existing ecosystem of its intended users. The principle is also supported by the content analysis, which clearly shows the importance of bot accessibility, both from a student and from a teacher point of view. A survey among the students showed that a vast majority of them use the Facebook Messenger client. Therefore, we decided to implement the chatbot UI (teacher and student) in that environment.

The other services in the architecture were all quite trivial to implement. The dynamic content service and the pro-active agent are only stubs at his point, but they are integrated into the architecture and ready to develop further when needed.

One lesson learned from the implementation work is the need to address privacy issues at an early point. Not only is it necessary *per se* when we utilize educational technology for education and research, but it was also needed in this case due to requirements from the Facebook Messenger API. Facebook requires us to upload a privacy policy to open up the chatbot for other users than invited testers. Also, requirements from the cloud services in use demand encrypted communication channels. The privacy issue may prove problematic in a scenario where this technology is used in a larger scale, still running in a 'cloud architecture'. While legislation differs across different parts of the world, we suspect that student questions may sometimes be rather sensitive in nature, thus not always suitable for cloud processing. Privacy issues needs to be thoroughly factored into the design work, from the very inception of the process.

# V. CONCLUDING DISCUSSION AND FUTURE WORK

In this research–in–progress paper, we have presented a conceptual architecture for chatbots in higher education (HE), drawing from a literature review and a content analysis of student questions in e-mails and discussion fora. The architecture has been implemented into a chatbot software accessed via Facebook Messenger. From the design process so far, in addition to the conceptual architecture, we have identified a set of tentative design considerations:

• A conceptualization of questions from learners – aiding designers in considering what type of questions and answers an HE chatbot should support.

- A recommendation to deliver HE chatbot functionality within the existing "student ecosystem" of applications to promote accessibility and ease-of-use.
- A call to build HE chatbot technology that *promotes* interaction between humans in the learning context rather than considering it a substitute to human interaction.
- An argument for supervised learning due to the demand for quality controlled responses from the chatbot, and a potential role of the chatbot as an agent exercising public authority.
- A reflection about the multitude of privacy norms that govern design in this case both educational norms, research ethics, and regulations to comply with third party cloud services regulations.

The design considerations above have been articulated through experiences in the design process but have still not been rigorously evaluated. Future work includes the process of defining chatbot behavior for a particular course (Basic java programming). The behavior will be defined through the creation of AIML documents drawing from both literature and the conducted content analysis of student questions from previous instances of the course. Clearly, the behavior of the chatbot is an essential part of design. Previous research, insofar as feasible, will be factored in to the AIML design process, such as the concept of academically productive talk (e.g., Tegos, Demetriadis, Papadopoulos, & Weinberger, 2016).

The next step will be to use the bot in the Java course and promote students to use it as a 'first resort' when asking questions. Data will be collected both through the RecordKepper; the chatbot logging feature in the architecture. The log data will be used to produce descriptive statistics of the use of the chatbot. In addition, we will conduct interviews with students and teachers to obtain qualitative (and possibly quantitative) data about their experience of using the chatbot. A mixed-methods approach (Venkatesh, Brown, & Bala, 2013; Ågerfalk, 2013) will thus be used to evaluate the implications of the chatbot from a multiple-stakeholder perspective.

It still remains to articulate a detailed plan for evaluation; that is, how to 'measure' or assess how the chatbot impacts student learning and other possible qualities in the learning practice. One of our main interests is how to align this type of technology to become part of the learning practice. Therefore, we conceive of the first 'naturalistic evaluation' in the Java course as a first step to produce narratives of the learning practice once the chatbot is introduced. Based on the first course, we will reflect about the results, (possibly) re-design the architecture and the AIML config, and test the chatbot again in a subsequent course.

Depending on the outcomes of our evaluations, we see a really interesting future field of research in combining the practice-oriented approach suggested here with collaborative learning environments based on multiple interacting agents (Hayashi, 2014). In such a setting, various agents with different roles would intervene in discussions among learners and teachers. Great opportunities, and indeed challenges, lie ahead for information systems research to understand how to design and employ conversational agents to facilitate an education practice effectively supporting students' learning.

### VI. REFERENCES

- AbuShawar, B., & Atwell, E. (2015). ALICE chatbot: trials and outputs. *Computación y Sistemas*, *19*(4), 625–632.
- Dragoni, N., Giallorenzo, S., Lafuente, A. L., Mazzara, M., Montesi, F., Mustafin, R., & Safina, L. (2017). Microservices: yesterday, today, and tomorrow. In *Present and Ulterior Software Engineering* (pp. 195–216). Springer.
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, 461–468. http://doi.org/10.1016/j.chb.2017.05.045

- Ghose, S., & Barua, J. J. (2013). Toward the implementation of a topic specific dialogue based natural language chatbot as an undergraduate advisor. In 2013 International Conference on Informatics, Electronics and Vision (ICIEV) (pp. 1–5). http://doi.org/10.1109/ICIEV.2013.6572650
- Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *Mis Quarterly*, 37(2), 337–355.
- Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems*, 8(5), 312–335.
- Griol, D., Manuel Molina, J., & de Miguel, A. (2014). The Geranium System: Multimodal Conversational Agents for E-learning. In E. Omatu, S and Bersini, H and Corchado, JM and Rodriguez, S and Pawlewski, P and Bucciarelli (Ed.), *Distributed Computing and Artificial Intelligence, 11th International Conference* (Vol. 290, pp. 219–226). Berlin: SPRINGER-VERLAG BERLIN. http://doi.org/10.1007/978-3-319-07593-8\_26
- Hassani, K., Nahvi, A., & Ahmadi, A. (2013). Architectural design and implementation of intelligent embodied conversational agents using fuzzy knowledge base. *Journal of Intelligent and Fuzzy Systems*, *25*(3), 811–823. http://doi.org/10.3233/IFS-120687
- Hayashi, Y. (2014). Togetherness: Multiple Pedagogical Conversational Agents as Companions in Collaborative Learning. In K. TrausanMatu, S and Boyer, KE and Crosby, M and Panourgia (Ed.), *Intelligent Tutoring Systems (ITS 2014)* (Vol. 8474, pp. 114–123). Berlin: SPRINGER-VERLAG BERLIN.
- Hevner, A. R. (2007). A Three-Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, *19*(2), 87–92.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in Information Systems research. *Mis Quarterly*, 28(1), 75–105.
- Krippendorff, K. (2006). The Semantic Turn A new Foundation for Design. CRC Press.
- Kumar, R., & Rose, C. P. (2011). Architecture for Building Conversational Agents that Support Collaborative Learning. *IEEE Transactions on Learning Technologies*, 4(1), 21–34. http://doi.org/10.1109/TLT.2010.41
- Latham, A. M., Crockett, K. A., McLean, D. A., Edmonds, B., & O'Shea, K. (2010). Oscar: An Intelligent Conversational Agent Tutor to Estimate Learning Styles. In 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2010). 345 E 47TH ST, NEW YORK, NY 10017 USA: IEEE.
- Leonhardt, M. D., Tarouco, L., Vicari, R. M., Santos, E. R., & da Silva, M. dos S. (2007). Using Chatbots for network management training through Problem-Based Oriented education. In A. Spector, JM and Sampson, DG and Okamoto, T and Kinshuk, X and Cerri, SA and Ueno, M and Kashihara (Ed.), 7TH IEEE International Conference on Advanced Learning Technologies (p. 845+). Los Alamitos, CA: IEEE COMPUTER SOC. http://doi.org/10.1109/ICALT.2007.275
- Lundqvist, K. O., Pursey, G., & Williams, S. (2013). Design and Implementation of Conversational Agents for Harvesting Feedback in eLearning Systems. In A. HernandezLeo, D and Ley, T and Klamma, R and Harrer (Ed.), SCALING UP LEARNING FOR SUSTAINED IMPACT (Vol. 8095, pp. 617–618). Berlin: SPRINGER-VERLAG BERLIN.
- Mohr, D. C., Cuijpers, P., & Lehman, K. (2011). Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research*, *13*(1).
- Neves, A. M. M., Barros, F. A., & Hodges, C. (2006). IAIML: A mechanism to treat intentionality in AIML chatterbots. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, 225–231. http://doi.org/10.1109/ICTAI.2006.64

- Newman, S. (2015). *Building microservices: designing fine-grained systems*. " O'Reilly Media, Inc."
- Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. Organization Studies, 28(9), 1435–1448.
- Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., ... Strohecker, C. (2004). Affective learning—a manifesto. *BT Technology Journal*, 22(4), 253–269.
- Tamayo-Moreno, S., & Perez-Marin, D. (2016). Adapting the design and the use methodology of a Pedagogical Conversational Agent of Secondary Education to Childhood Education. In A. GarciaPenalvo, FJ and Mendes (Ed.), 2016 International Symposium on Computers in Education (SIIE). New York: IEEE.
- Tegos, S., Demetriadis, S., Papadopoulos, P. M., & Weinberger, A. (2016). Conversational agents for academically productive talk: a comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, *11*(4), 417–440. http://doi.org/10.1007/s11412-016-9246-2
- Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European Journal of Information Systems*, *25*(1), 77–89.
- Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quarterly*, 37(1), 21–54.
- Yun, E. K., & Cho, S. B. (2003). Learning classifier system for generating various types of dialogues in conversational agent. In H. Liu, J and Cheung, YM and Yin (Ed.), *Intelligent Data Engineering and Automated Learning* (Vol. 2690, pp. 84–88). Berlin: SPRINGER-VERLAG BERLIN.
- Ågerfalk, P. J. (2013). Embracing diversity through mixed methods research. *European Journal of Information Systems*, 22(3), 251–256. http://doi.org/10.1057/ejis.2013.6