Multi-level De-anonymisation for Initially Anonymous Discussion Systems in a Self-regulated Learning Environment

Tenshi Hara[1], Anne Schumacher[2,*], Karina Hara[3], Iris Braun[1], Felix Kapp[4] and Alexander Schill[1]

[1]*Chair of Computer Networks, School of Engineering Sciences, Technische Universität Dresden, Dresden, Germany*

[2]*Technische Universität Dresden, Dresden, Germany*

[3]*Blue Pumpkin LLC, Kailua-Kona, HI, U.S.A.*

[4]*Chair of Learning and Instruction, School of Science, Technische Universität Dresden, Dresden, Germany*

Keywords: Self-regulated Learning, Discussion Systems, Forums, Anonymity, De-anonymisation.

Abstract: Discussion systems are a valuable asset in attaining self-regulated learning. Beyond the limitations of on-campus classroom settings, they enable internal feedback, peer feedback as well as external feedback. A motivating factor to continued and frequent utilisation of such systems is anonymity. However, anonymity is a double-edged sword. On one side, it provides strong incentives to use discussion systems, on the other side it invites destructive behaviour such as trolling. Furthermore, strong students are discouraged from continued utilisation if they cannot attribute their contribution to themselves. We propose an initially anonymous discussion system, which enables retroactive de-anonymisation on multiple levels, namely with respect to the identity degree as well as the attribution dimension.

1 INTRODUCTION

Modern teaching methods such as Peer Instruction (Mazur, 1999; Crouch and Mazur, 2001; Mazur, 2017) aim at helping students achieve self-regulated learning (Zimmerman et al., 2000). Commonly, such methods rely on peer interactions amongst the students. For example, Peer Instruction includes an explicit Peer Discussion phase. In general, these peer interactions provide students with valuable feedback on their knowledge and learning progress. They can ask questions, provide answers, discuss their understanding of knowledge, or simply receive feedback from their teachers. Of course, these aspects involve self reflection (internal feedback), comprehension of others' concepts (peer feedback) as well as corrections (external feedback).

Peer interactions can be transferred outside of the classroom into online media such as forums and discussion systems. One example for that is our graphical discussion system *Graphicuss* which we presented at last year's CSEdu (Chen, 2016; Hara et al., 2017). This allows students to continue learning activities outside of the classroom. However, interactions are accessible to all peers online, rather than the few directly involved in the classroom. Therefore, anonymity is an imperative motivating factor for utilisation of online discussion systems (summarised in e.g., (Hara, 2016)), allowing students to make mistakes or ask 'stupid questions' without fear of exposure due to the entirely documented and available history of interactions. Nevertheless, at some point in time, motivation may be further fostered by attribution rather than anonymity. In general, anonymous systems do not have a de-anonymisation concept allowing students to attribute their contributions to themselves retroactively. For example, if students provide a very good contribution to a discussion which is well received by their peers, the students may want to attribute the contribution by attaching their name to it. This in return allows their peers to actively seek the students in the classroom (i.e., offline).

In this position paper, we present a multi-level de-anonymisation concept for discussion systems. Students' contributions are initially entirely anonymous. Peers are unable to correlate two contributions of the same student to each other. In a first de-anonymisation step, pseudonyms are attached to a student's contribution. Such pseudonyms can be either the same system-wide, or attached to a discussion thread. In a final step, the pseudonyms can be

[*] Graduate student

Hara, T., Schumacher, A., Hara, K., Braun, I., Kapp, F. and Schill, A.
Multi-level De-anonymisation for Initially Anonymous Discussion Systems in a Self-regulated Learning Environment.
DOI: 10.5220/0006760403020307
In *Proceedings of the 10th International Conference on Computer Supported Education (CSEDU 2018)*, pages 302-307
ISBN: 978-989-758-291-2

replaced by the student's actual identity; e.g., their name. These three levels of identity obfuscation are met by three degrees of attribution, namely individual contributions, threads, and topics (sets of threads).

2 RELATED WORK

It is imperative for students to receive feedback on their understanding of knowledge as well as their learning progress. This can be achieved through formative assessment in means of internal feedback, peer feedback and external feedback (Peters et al., 2017). This in return helps students achieve self-regulation (Winne and Hadwin, 1998; Zimmerman et al., 2000). The process of internal feedback can be fostered by asking students to reflect on their own understanding of knowledge. Explicitly asking them to write a summary or explain a topic to their peers forces them to evaluate and order their own knowledge. Also, asking questions in an understandable fashion requires some preceding internal feedback. Similarly, it should be easy for lecturers to provide students with external feedback[1]. Therefore, we argue that the challenge lies in providing a system with attributes of useful, feasible and suitable peer feedback. The interesting question then is, how can peer feedback be fostered and perceived as a motivating experience for feedback providers and receivers?

2.1 Self-regulated Learning

Self-regulation unfolds over four more or less flexibly sequenced phases of recursive cognition (Winne and Hadwin, 1998), namely *task perception*, *planning*, *enacting*, and *adaptation*. During task perception, students gather information about the task and classify it based on their motivational state, their self-efficacy, as well as their environment, especially in means of peer performance. Next, during planning, students define goals to be achieved and rewards to be obtained by fulfilling the task. For that, they plan how to achieve the goals and determine a reward importance. The strategy can vary depending on explicit behaviours, cognitive engagement, and motivation. After surpassing a threshold volition, students enact their plan. Finally, students evaluate their performance and adapt their strategy for increased (or at least maintained) success in future repetitions of the same or similar tasks. The success in achieving the defined goals as well as the perceived value of the obtained rewards gravely influence the adaptation process.

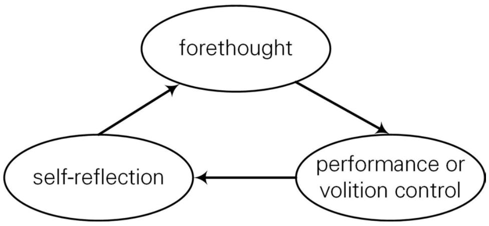


Figure 1: Cyclic transition states of self-regulation.

The four phases defined by (Winne and Hadwin, 1998) are inter-connected by three transition states as defined by (Zimmerman et al., 2000), namely *forethought*, *performance or volition control*, and *self-reflection* (cf. Figure 1). The latter two are influenced by the learner's peers, especially by assessing the own performance in comparison to the peers' performance as well as by reflecting on own mistakes in comparison to the peers' mistakes and (potentially better/worse) solutions. Thus, supporting students in these two transition states immediately enhances their success in the corresponding self-regulation phases connected through these states.

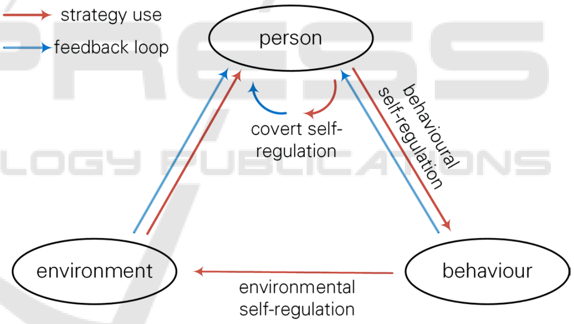


Figure 2: Triadic self-regulation.

Of course, peers' influences on a learner's perception of their own learning progress are also important from a triadic self-regulation perspective (Zimmerman et al., 2000) (cf. Figure 2). From the three aspects of triadic self-regulation, *behaviour* and *environment* are directly influenced by a learner's peers, especially when taking the before-mentioned transitions states into consideration. Simply, peer influences have a strong impact self-regulation strategies: interactions with peers are an imperative aspect in the process of achieving self-regulation. Therefore, peer interactions are a highly important part of modern learning and teaching concepts such as Peer Instruction (Mazur, 1999; Mazur, 2017).

[1] Providing external feedback individually and at the best moment is a topic of its own and shall not be discussed in this position paper.

2.2 Discussion Systems and Forums

Forum systems such as auditorium (Beier et al., 2014) or our Graphicuss (Chen, 2016; Hara et al., 2017) provide users with the possibility to discuss topics in a linear, chronologically ordered fashion. An underlying tree structure consisting of the posts (partially) ordered within threads, which again are organised by topics or categories[2] allows for intuitive cognition of posts belonging together as a discussion, as well as related or unrelated post, threads, and topics. Within a thread, users can discuss the main subject as well as answers other users have posted before. This tree structure can have a large number of levels if the creation of sub-threads within threads is allowed.

Forums commonly require users to register an identity/pseudonym under which their posts are distinguished from others'. Participation of anonymous users is generally not desired, as this invites spam and various forms of vandalism, such as trolling.

Forum systems that support anonymous participation attribute all anonymous posts to one indistinguishable pseudonym, e.g. 'anonymous', making it difficult to follow individual contributors' argumentations. Attribution to distinguishable anonymous users, as available in other collaboration systems like Google Docs[3] is is not implemented in any forum system to the authors' knowledge.

Verified identities, as implemented in popular social media platforms like Facebook[4] and Twitter[5], are uncommon. Nevertheless, verified identities, or to the least verified pseudonyms, are imperative for a well-structured and believable discussion culture. Access control must be established in a privacy-respecting fashion (Pötzsch and Borcea-Pfitzmann, 2009) while enabling effective deflection of spam and vandalism, or sh%t-storms[6] (Rost et al., 2016). At the same time, users' contributions must remain trustworthy and believable, especially allowing other users to assess the value and truthfulness of a contribution (Kartal et al., 2011). Accordingly, product recension platforms often only allow negative feedback if the users attach their name to their recension.

2.3 Anonymity

As discussed in the previous sections, self-assessment and peer influences are an important factor for students' learning success. Therefore, these are amongst the most important influence factors identified in the Visible Learning meta-studies (Hattie, 2009; Hattie, 2013). Nevertheless, exposure should also be considered, especially if it leads to students being forced to concede to a lack of knowledge or wrong understanding thereof. We identified anonymity as a strong motivating factor to free students of fears of exposure (Hara, 2016).

Giving students the option to contribute anonymously can be desirable to lower users' inhibitions to contribute. However, in common systems anonymity is generally only provided in regard to what other users are presented.

The problem with anonymity is its double-edged nature. On on hand, it provides strong incentives for students to contribute without fear of exposure, on the other hand it inhibits strong students from contributing, as they cannot attribute their contributions to themselves. The strong motivation for attribution is 'showing off': at some point, students want their peers to know who has authored well-received contributions. Prestige is a strong motivation after all.

Another problem with anonymity is system susceptibility to trolling and other malicious activities. Under the cover of anonymity, users loose restraint to negatively or destructively contribute to the system. Therefore, systems commonly store identifying metadata (e.g., IP address, e-mail address) in order to retroactively reprimand users or revoke system access. This in return leads to an underlying fear of exposure as users must trust that the system does not divulge their identifying information unnecessarily.

Regardless the problems, anonymity remains a strong motivating factor for system utilisation. Allowing students to remove the constraints of anonymity retro-actively can address strong students and motivate them to contribute even more. Trust must be established between the system and the students, only providing (desirably strong) anonymity amongst the students themselves.

Similar considerations apply to the relations between students and teaching staff. Knowing that lecturers are unable to identify students within a system can foster 'stupid questions'[7], which students would normally not dare to ask. In reverse, lecturers might want to be able to identify weaker students

[2] In general, a post is a single contribution of a user. Multiple posts, e.g. a question and corresponding answers, are organised in threads. Finally, multiple threads can be attached to a topic spanning multiple discussions on a broader subject.

[3] https://docs.google.com/ – They use animals; e.g., 'anonymous elephant' or 'anonymous cat'.

[4] https://facebook.com/

[5] https://twitter.com/

[6] The four-letter word was censored at an offended reviewer's request. We wish to emphasise that it is a common, properly cited term in this context.

[7] Questions students perceive to be stupid. However, this is a common misconception: any question helps strengthen a correct understanding of knowledge.

in order to provide targeted help. Based on our research, we strongly believe that a well designed system should allow lecturers to provide such support even in an anonymous setting (Hara, 2016).

3 DE-ANONYMISATION

Based on the existing ideas in combination with the goals of self-regulation, we propose a multi-level de-anonymisation approach. Initially, students shall be totally anonymous, with only system administrators or law enforcement being able to identify students. Hence, even teachers are initially unaware of their students' identities as system users. Then, it shall allow students to ask questions, provide answers, and discuss topics in three *identity degrees*, namely *totally anonymous* (`ID-0`), *pseudonymised* (`ID-1`), and *identity-attributed* (`ID-2`; also *identified*), as well as three *attribution dimensions*, namely *per post* (`Dim-0`), *per thread* (`Dim-1`), and *per topic* (`Dim-2`). The combination of identity degrees and attribution dimensions can be described by an *attribution tuple* $(\texttt{Dim},\texttt{ID})$; e.g., $(0,1)$ indicates that a student contributes a single post under their pseudonym. Both sets are ordered descending: $\texttt{ID-2} > \texttt{ID-1} > \texttt{ID-0}$ and $\texttt{Dim-2} > \texttt{Dim-1} > \texttt{Dim-0}$. Furthermore, attribution tuples shall hold

$$\forall i,j \in \{1,2,3\} : \begin{pmatrix}\texttt{Dim}_1\\ \texttt{ID}_i\end{pmatrix} > \begin{pmatrix}\texttt{Dim}_2\\ \texttt{ID}_j\end{pmatrix} \text{ iff } \texttt{Dim}_1 > \texttt{Dim}_2$$

and

$$\begin{pmatrix}0\\ \texttt{ID}_1\end{pmatrix} > \begin{pmatrix}0\\ \texttt{ID}_2\end{pmatrix} \text{ iff } \texttt{ID}_1 \geq \texttt{ID}_2 \,.$$

A higher attribution dimension always overwrites a lower, independent of the value of the identity degree. However, this condition shall only hold for future contribution in the discussion system. Even though the identity degree is set in the higher attribution dimension and applies to all lower attribution dimensions, a student can always change settings for individual posts/threads in lower attribution dimensions. For example, if a thread is set to `ID-1`, all future contributions of the same student in that thread also default to `ID-1`. Nevertheless, the student can choose `ID-0` or `ID-2` for individual posts. Therefore, the above defined order only serves the purpose of simplifying interactions: if a student wants to be predominantly anonymous in a thread, they can set the attribution tuple $(1,0)$ rather than having to choose the identity degree each time they post within the thread.

In order to have a feasible identity degree system, the discussion system has to ensure that pseudonyms and identities can only be used by the authorised users. We suggest allowing anonymous registrations be limited to `ID-0` independent of the attribution dimension. That way, malicious users cannot incarnate pseudonyms or identities. Pseudonyms should be enabled for verified user accounts; e.g., after e-mail verification. Then, `ID-0` and `ID-1` are available to the users in all attribution dimensions. Users should be able to choose their own pseudonyms or get random pseudonyms assigned. In both cases, pseudonyms used shall only be re-usable for the users who have first used the pseudonym. Last, `ID-2` should be made available only to users with verified identities. For this, a trusted identity provider or identity management system (IMS) should be used. For example, a university could provide a student's identity based on the matriculation records. Secure access to these sensitive data could be realised through an authentication service like Shibboleth[8], which is summarised in Table 1.

Table 1: Available identity degrees based on verification.

available default ID	allowed upgrades	verfication
$\{0\}$	$\emptyset$	none
$\{0,1\}$	$\{1\}$	E-Mail
$\{0,1,2\}$	$\{1,2\}$	IMS

As mentioned above, students should not be limited to a single pseudonym. They might want to use a pseudonym in a discussion they are sure to contribute positively, but use a different pseudonym in a discussion they are unsure and which might reflect negatively on them. However motivating, such flexibility may open the discussion system to 'adverse whitewashing'. Students can simply choose a new pseudonym as soon as they feel a pseudonym has be attributed to too many negatively perceived contributions. Additionally, 'trolls' could constantly use new pseudonyms for their *trolling* activities. Hence, a suitable compromise must be found to allow flexible utilisation of pseudonyms on one hand, and troll control on the other. Likewise, pseudonym management efforts must be taken into consideration. Answering questions / quoting posts (in full as well as partially) requires further considerations with respect to the `ID` of the quoted post; e.g., which `ID` is the answer supposed to have, or what happens to quoted content as soon as the `ID` of the original posting has been altered? Describing all considerations would break the confines of this position paper; thus, we wish to allude the reader to Table 2 for further details. Beyond that scope, further functions of discus-

[8] https://www.shibboleth.net/

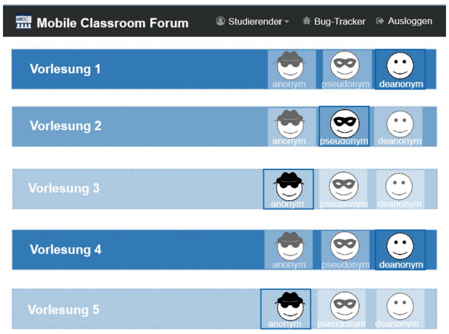
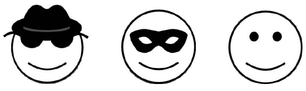


Figure 3: Overview of lectures and their pre-set `ID` levels.

sion systems must also observe the boundaries of attribution. For example, a personal message (PM) system introduces relations between students. If a student follows another student's contributions based on a PM-established relation of `ID-1`, a posting under `ID-2` should not be attributed to the given relationship as the `ID-2` posting would break the pseudo-anonymity of the `ID-1` PM-established relationship (in simple term: as soon as a real identity is attached to an initially pseudonymous post, the other end of the relationship would be able to attribute all previously pseudonymised posts to the student).

Independent of the challenges described above, students require a simple visualisation of their degree of anonymity. Simply showing the attribution tuple is not a suitable solution. Based on user interviews and usability tests, we recommend an icon-based visualisation with 'anonymID smilies' (cf. Figure 4), or something similar with clearly distinguishable visualisation for the `ID` levels.

The anonymID smilies work best in combination with easily comprehensible colour schemes. In our prototype we stayed true to the basic colour scheme of our system (blue) and used shades of the basic highlight colour (by saturation: 75% for `ID-2`, 50% for `ID-1`, 25% for `ID-0`). That way, students are able to immediately identify the pre-set identity degree for lectures[9] (cf. Figure 3) as well as threads (cf. Figure 5). The colour coding did not work as well for individual posts. Instead, textual representations in a designated area should be used; in Figure 5 this can be seen in the upper right corner of each post ('Anonym': an anonymous student; 'Pseudo1': a pseudonymised student; 'Max Mustermann': student's verified identity (name)).



(a) `ID-0` (b) `ID-1` (c) `ID-2`

Figure 4: Suggestion for visualisation of the `ID` levels.

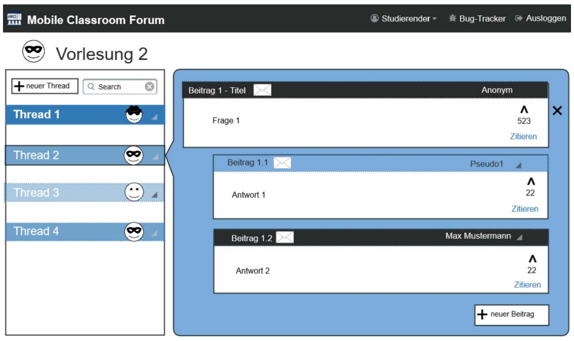


Figure 5: Within a lecture view: different pre-set `ID` levels for threads (left) and individual postings (right).

Of course, this can be attributed the fact that our prototype only used shades of blue rather than different colours. Without the other colours in visible proximity, it is very hard for users to intuitively distinguish the colours.

A justifiable question is how beneficial retroactive de-anonymisation actually can be. Students may decide to de-anonymise a contribution, but this may remain totally unnoticed by their peers if the topic or thread as been closed or concluded its line of discussion. However, students do tend to revisit certain discussions in preparation of their exams. This is especially true for questions on understanding complex course material. In order to quickly revisit how a certain understanding developed, these old discussions are very helpful. This in return brings the argument back full-circle to the beginning: students can attribute their contributions and become valuable learning contacts for their peers. This in return is very motivating, and it fosters an even stronger understanding of and confidence in their knowledge as the students might change into the teacher role when explaining concepts, et cetera to their peers.

4 CONCLUSION

Discussion systems are a valuable asset in attaining self-regulation. They enable internal feedback, peer feedback as well as external feedback beyond the scope of on-campus lectures, namely into the online media. A motivating factor to continued and frequent utilisation of such systems is anonymity. In this position paper we have discussed the pros and cons of anonymity and why it is important to enable retroactive de-anonymisation. We presented a multi-level de-anonymisation concept based on three identity degrees as well as three attribution dimensions, and an attribution tuple $(\mathtt{Dim}, \mathtt{ID})$ to describe anonymity settings within discussion systems. For some aspects we

[9] Topics correspond to lectures in our prototype.

Table 2: Comparison of pseudonym counts.

pseudonym availability	**pseudonyms utilised**	**advantages**	**disadvantages**
exactly one	1	• prevents ‘adverse whitewashing’ • easy pseudonym management	• only allows `ID-0`→`ID-1` and `ID-0`→`ID-2`, but not `ID-0`→`ID-1`→`ID-2` • cannot retroactively modify `ID` (`ID-1`↛`ID-2`)
limited	`max`	• flexibly usable	• possibility of ‘adverse whitewashing’ • no best-practice for suitable `max` value
	1/Topic	• no direct linking of pseudonyms to topics • good degree of ‘adverse whitewashing’ prevention	• lack of Variability/flexibility • no control of / influence over pseudonyms
	1/Thread	• prevents ‘adverse whitewashing’ within threads • allows retroactive `ID` modification (`ID-1`→`ID-2` per thread)	• lack of variability/flexibility
unlimited	∞	• flexibility • own arbitrament • allows retroactive `ID` modification (`ID-1`→`ID-2` per post)	• simplifies ‘adverse whitewashing’ • challenging pseudonym management

outlined the utility and limitations of the so designed anonymity management.

In the future, we want to test our de-anonymisation concept outside the constraints of a simple forum prototype: we plan to implement it into the next version of *Graphicuss*in order to investigate the limitations of our concept; we are sure that different classroom and off-campus online settings require different attribution settings, or to the very least some nifty fine-tuning. Further, we want to investigate the influence of pseudonyms externally provided through an identity management system such as Shibboleth[8].

Simple metrics should be ascertainable through targeted interviews with students and teachers (e.g., System Usability Scale, NASA Task Load Index). Using the prototype in actual lectures should enable us to determine a relation between the students' willingness to de-anonymise, and their actual utilisation of the de-anonymisation feature. This should also help identify potential trade-offs.

REFERENCES

Beier, L., Braun, I., and Hara, T. (2014). auditorium - Frage, Diskutiere und Teile Dein Wissen! In *GeNeMe 2014 - Gemeinschaften in Neuen Medien*. GeNeMe.

Chen, K. (2016). Graphical Discussion System. Master's thesis, Technische Universität Dresden.

Crouch, C. H. and Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American journal of physics*, 69(9):970–977.

Hara, T. (2016). *Analyses on tech-enhanced and anonymous Peer Discussion as well as anonymous Control Facilities for tech-enhanced Learning*. PhD thesis, Technische Universität Dresden. https://katalogbeta.slub-dresden.de/id/0017472650/.

Hara, T., Chen, K., Braun, I., and Kapp, F. (2017). Graphicuss - Proposing Graphical Discussion System. In *Proceedings of the 9th International Conference on Computer Supported Education (CSEdu 2017)*.

Hattie, J. (2009). Visible learning: A synthesis of meta-analyses in education.

Hattie, J. (2013). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Kartal, A., Doerfel, S., Roßnagel, A., and Stumme, G. (2011). Privatsphären-und Datenschutz in Community-Plattformen: Gestaltung von Online-Bewertungsportalen. *Informatik*, 412:1–15.

Mazur, E. (1999). Peer instruction: A users manual.

Mazur, E. (2017). Peer instruction. In *Peer Instruction*, pages 9–19. Springer.

Peters, O., Körndle, H., and Narciss, S. (2017). Effects of a formative assessment script on how vocational students generate formative feedback to a peers or their own performance. *European Journal of Psychology of Education*.

Pötzsch, S. and Borcea-Pfitzmann, K. (2009). Privacy-Respecting Access Control in Collaborative Workspaces. In *PrimeLife*, pages 102–111. Springer.

Rost, K., Stahel, L., and Frey, B. S. (2016). Digital social norm enforcement: Online firestorms in social media. *PloS one*, 11(6):e0155923.

Winne, P. H. and Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in educational theory and practice*, 93:27–30.

Zimmerman, B. J., Boekarts, M., Pintrich, P., and Zeidner, M. (2000). Attaining self-regulation: a social cognitive perspective. *Handbook of self-regulation*, 13.

All URLs within this paper were last successfully accessed on 19 November 2017.

Parts of this position paper are based on Anne Schumacher's Master's thesis.