

---

# An Architecture for the Automated Detection of Textual Indicators of Reflection

Thomas Daniel Ullmann

Knowledge Media Institute, The Open University, UK  
t.ullmann@open.ac.uk

**Abstract:** Manual annotation of evidence of reflection expressed in texts is time consuming, especially as fine-grained models of reflection require extensive training of coders, otherwise resulting in low inter-coder reliability. Automated reflection detection provides a solution to this problem. Within this paper, a new basic architecture for detecting evidence of reflection is proposed that allows for automated marking up of written accounts of certain, observable elements of reflection. Furthermore, three promising example annotators of elements of reflection are identified, implemented, and demonstrated: detecting reflective keywords, premise and conclusions of arguments, and questions. It appears that automated detection of reflections bears the potential to support learning with technology at least on three levels: it can foster creating awareness of the reflectivity of own writings, it can help in becoming aware of reflective writings of others, and it can make visible reflective writings of learning networks as a whole.

**Keywords:** reflection detection, learning networks, awareness

## 1 Introduction

Agreement of human coders about levels of written reflection seems to be a difficult task. Wong et al. [1] report regarding inter-coder reliability that the use of a fine-grained categorization schema with six categories was problematic and less reliable. They state that in the literature is more discussion about the concept of reflection, than research on how to assess reflection. Sumsion and Fleet [2] report 50% inter-coder reliability for a three-stage categorization system for reflection (highly reflective, moderately reflective, not reflective). On the other hand trained coders can achieve high reliability in assessing journal writings using a three-category schema (non-reflector, reflector and critical reflector) [1].

Although these methods provide valuable insights about reflections in writings, their application is time-consuming and results are usually available only far after the act of reflective writing.

Methods for the automated annotation of writings regarding elements of reflection can be a promising technology to raise awareness about levels of own reflective writings, to find reflective writings of others, or to get an overview of reflective writings

of the whole learning network, immediate and independent of the daily performance of the evaluators.

This research describes a first approach of how to automatically detect accounts of reflection with natural language processing techniques with the goal to make visible traces of reflection in online learning networks with large number of participants.

To fulfil this vision several steps have to be taken. The following text focuses on the central ideas and exemplifies it with three annotators, which are derived from theories of reflection. The annotators build the fundament of the software architecture. The related approach section show successful implementations of automated content classifiers in an educational setting. Then the architecture of the reflection detector is outlined and its main strength described. The concrete implementation of the core building blocks of the architecture targets reflection. Therefore definitions of reflection and definitions, which contain reflection, are described to explore the manifold facets of reflection. Three elements of reflection are highlighted, which later will be used as examples of automated detection. After the theoretical foundations, Subsequently, three annotators based on the identified elements of reflection are presented and demonstrated with the help of an example of reflective writings and an encyclopaedic text.

## **2 Related Approaches - Automated Classification of Content**

One approach in the area of automated content analysis focused especially on the related concept of reflection – critical thinking. High inter-coder reliability was reported between the computerized content analysis system and human codes of 0.65 and 0.71, for four categories of critical thinking [3]. The classified contents were discussion forum posts. The automated essay scoring was based on Bayesian networks.

Another system reports about the automatic analysis of collaborative learning processes [4]. According to the study a novel algorithm for the automated classification of content was used – the confidence restricted cascaded binary classification approach. For each of the seven classification categories they calculated Cohen's Kappa. Values for their baseline approach ranged from 0.49 to 0.91. Only two categories, the epistemic and the social modes of co-construction dimensions, were below 0.7.

Both approaches are based on pre-annotated data sets, which are used to train a classifier. The high reliability scores seem encouraging for further research in this area. Especially the first approach is more similar to the domain of reflection, while the second one uses a more sophisticated approach, applied however in another domain than reflection.

The use case of both approaches is to help researchers with the annotations of texts and text segments for content analysis. The ACAT (Automated Content Analysis Tool) system for example contains a quantitative content analysis (QCA) training module, with which users can train a model for their content analysis. This model has to be created beforehand with a model management tool [3]. The mentioned approaches are static insofar as the process foresees to import into the system a defined set of documents, which are then annotated by the system.

The proposed architecture however targets the dynamics of content creation in online learning networks, in which learners create content on the Web with tools of their choice (for example blogs or wikis), while others can subscribe or follow the work and contribute back. The goal of the architecture is to automatically retrieve these dynamically growing contents, annotate them according to their reflective elements and provide an interface to retrieve reflective documents. The architecture will be therefore web-based compared to the above outlined desktop-centered approaches.

### **3 Architecture for Detection of Reflection**

The core of the reflection detection architecture consists of analysis engines, which analyze textual artefacts. They add structure to unstructured data. An analysis engine consists of annotators, which add metadata to the artefact. If an analysis engine consists of only one annotator it is called a primitive analysis engine if it consists of more than one annotator it is an aggregated analysis engine. The outcome of an annotator results in a common analysis structure (including the type, features of the annotation, and the position in the text).

The architecture envisions a web-service taking as input (web-)documents and returning either an annotated version of the document or statistics about the document. Several document formats like html, feeds, rich text format, and PDF should be supported. A mime type detector assigns each document type to a specific parser to extract the content and metadata.

In the case of webpages and other documents usually only the content is important and not so much the information about the navigation or other decorating elements. A cleaner therefore removes the boilerplate of such documents.

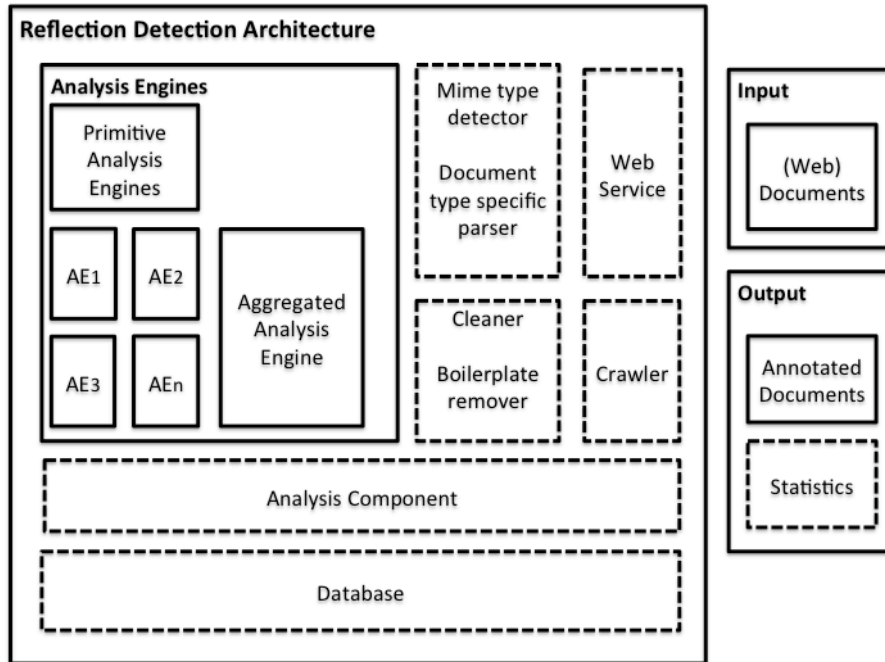


Fig. 1. Reflection Detection Architecture (solid lines: core building blocks; dotted lines: optional building blocks).

An extended analysis component would be necessary to return more elaborate statistics than counts of each annotation type per document. This could be a rule engine combining the information of each annotator based on rules with the goal of making statements about the depth of reflection.

The documents, annotation, and metadata, can be indexed and stored in a database. Optionally a web crawler could help to spider documents starting with seed URLs. This could be helpful to compare a set of webpages over time, according their reflective accounts. The following figure shows the reflection detection architecture.

The current implementation of the reflection detector is based on Apache Unstructured Information Management Architecture UIMA<sup>1</sup>, a framework architecture to gain structured information analyzing unstructured data. The framework provides a standardized environment for developing components, which can be shared and plugged in other infrastructures, which adopt this standard.

As mentioned, the core of the architecture are the analysis engines, which analyze and annotate documents. The goal is to develop engines, which are tailored to detect indicators of reflection in writings. Before I outline three concrete implementations of these core elements, the following section serves as a short overview of reflection theory with the goal to show the variety of elements, which make up reflection.

<sup>1</sup> <http://uima.apache.org/>

## **4 Elements of Reflective Thinking**

Several concepts are connected to reflection, which some authors in the literature subsumed as reflective thinking, others highlight as a related but different thinking skills. Examples include strategic thinking, meta-cognition, critical thinking, and logic.

Dewey, for example, describes reflective thinking as an “active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusion to which it tends” [5]. Reflection is seen as the critical evaluation of own assumptions and conclusions.

Halpern especially focuses on critical thinking: “Critical thinking is the use of those cognitive skills or strategies that increase the probability of a desirable outcome. It is used to describe thinking that is purposeful, reasoned, and goal directed – the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions, when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task” [6]. Reflection and critical thinking are seen as highly connected.

The model of Pintrich [7] deals amongst others with meta-cognitive skills in the context of self-regulated learning, which he defines as “an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate and control their cognition, motivation and behaviour, guided and constrained by their goals and the contextual features in the environment”. This theory emphasizes the close relation of self-regulation and reflective thinking.

Emotions, critical thinking, and the change in mindset are seen as crucial elements of reflection in the theory of Atkins & Murphy [8]. They conclude in their initial review that despite the differences between the accounts of authors describing their model of the reflective process, three key stages can be identified:

1. Awareness of uncomfortable feelings and thoughts. The skills needed are seen in self-awareness and the ability to describe feelings and thoughts.
2. Critical analysis of the situation, feelings, and knowledge, which according to the author needs the skill critical analysis
3. Development of new perspectives on the situation, which needs the skill set of synthesis and evaluation.

Already these four definitions suggest that reflection is an important part in several theories and has many facets. This character of reflection makes it an interesting area of research. Each element of reflection bears its own research problem regarding the automated detection. Three elements of reflective critical thinking, which are seen as a starting point for an automated reflection detector, are outlined in depth to illustrate this point.

### **4.1 Reflective Keywords**

Reflections can be expressed in many ways. Some nouns, verbs, adjectives, and adverbs however are semantically connected to the concept of reflection. An example for a close semantic relation of words is “to reflect about something” and “to muse

about something”. Musing means to reflect deeply on a subject. While reflecting about something and to praise something is still an act of thought, they have a very different character. It is reasonable to assume that people writing reflectively will use these keywords that are semantically connected with reflection more frequently when writing non-reflective.

#### 4.2 Premise and Conclusion

One of the skills involved in critical and reflective thinking is the ability to craft correct and convincing arguments. The study of logic is conducted in several disciplines, for example in computer science, maths, philosophy, and argumentation theory.

The critical analysis of the situation can be expressed with arguments. One type of arguments is sentences that relate premises to conclusions. To prove a conclusion the premises have to be true and the argument valid. The writer has to think about the situations under which the conclusion is valid thereby making explicit the underlying logic of its arguments.

To state a premise and conclusion several indicator words exists. According to Halpern [6] the following premise and conclusion indicators can be found:

**Table 1.** Premise and conclusion indicators, according to Halpern [6].

<b>Premise indicators</b>	<b>Conclusion indicators</b>
Because, for, since (in the meaning of because and not related to time), if, given that, being that, as shown by, as indicated by, the reasons are, it may be inferred (or deduced) from, the evidence consists of, in the first place (suggests that a list of premises will follow), secondly, seeing that, it follows from, whereas	Therefore, hence, so, thus, consequently, then, shows that (we can see that), accordingly, if follows that, we may infer (conclude) (deduce) that, in summary, as a result, for all the reasons, it is clear that

#### 4.3 Thought provoking Questioning

Questions are one of the most important techniques to engage people in thinking about the answer of the question or to create new questions. In a learning scenario a teacher can use questions for guiding learning. And the ability of students to ask own questions and to find answers is a highly desirable skill for learners. Self-questioning is one of the success criteria of a highly reflective accounts [9].

Questions can be either closed questions, which can be answered with yes or no, or open questions, which need a longer argument to answer. Latter ones bear more potential to provoke reflective and critical thinking.

**Table 2.** Thought provoking questions, according to King [10].

<b>Thinking Skill</b>	<b>Question</b>
-----------------------	-----------------

Application	What is a new example of...?
Application	How could ... be used to...?
Prediction/hypothesizing	What would happen if...?
Analysis/inference	What are the implications of...?
Analysis/inference	What are the strengths and weaknesses of...?
Identification and creation of analogies and metaphors	What is ... analogous to...?
Activation of prior knowledge	What do we already know about...?
Activation of relationship (cause-effects)	How does...affect...?
Activation of prior knowledge	How does ... tie in with what we learned before?
Analysis	Explain why...
Analysis	Explain how ...
Analysis	What is the meaning of... ?
Analysis of significance	Why is ... important?
Comparison-contrast	What is the difference between...and...?
Comparison-contrast	How are ... and ... similar?
Application - to the real world	How does ... apply to everyday life?
Rebuttal argument	What is the counterargument for...?
Evaluation and provision of evidence	What is the best ... and why?
Synthesis of ideas	What are some possible solutions to the problem of...?
Comparison - contrast	Compare ... and ... with regard to...
Analysis of relationship (cause-effect)	What do you think causes...? Why?
Evaluation and provision of evidence	Do you agree or disagree with this statement:...?
Evaluation and provision of evidence	What evidence is there to support your answer?
Taking other perspectives	How do you think ... would see the issue of...?

King [10] outlines that when questions are factual, the tendency is that only facts are recalled (see Table 2). If questions are thought provoking critical thinking is more likely to occur. As a guideline for students a table of question stems could help students to formulate their own thought provoking questions, based on these generic questions.

The description of these three elements focused especially on a word and sentence level. These textual cues can be matched in texts with regular expressions, which was mainly used for the annotators in the following example section. Regular expressions are seen as one of many possible methods to detect reflection in writings. One of their benefits is that no model has to be trained in advance and the time to parse content is relatively short compared to more complex methods. I will now outline the translation of each of the three elements of reflection into an analysis engine.

#### 4.4 Annotator for Reflective Keywords

For the annotator of words, which are semantically related to reflection WordNet 3.0<sup>2</sup>, a lexical database of English was used in combination with the Java WordNet Library<sup>3</sup> to automatically retrieve these words for their latter use in the annotator. WordNet organizes words into so called synsets, which is a set of synonym words. These synsets are linked to other synsets, which have semantic relations like hyponyms, hypernyms, and entailment relations, etc.

WordNet contains relatively few adverbs, however, adjectives can be in most cases converted in adverbs by adding an affix (-ly) to the end. Most connections between synsets are made within the same part of speech (nouns, verb, adjectives, adverbs), with some links between different parts of speech (POS).

POS	Unique Strings	Synsets	Total Word-Sense Pairs
Noun	117798	82115	146312
Verb	11529	13767	25047
Adjective	21479	18156	30002
Adverb	4481	3621	5580
<b>Totals</b>	<b>155287</b>	<b>117659</b>	<b>206941</b>

Fig. 2. WordNet Statistics<sup>4</sup>

To find all related words the starting point is to determine the appropriate synsets. Reflection for example, has eight senses in WordNet. One sense is described as “a calm, lengthy, intent consideration”, while another sense is about “the phenomenon of a propagating wave”. Only senses, which are related to cognition are kept for the further analysis. This process is repeated for the verb “reflect”, the adjective “reflective”, and the adverb “reflectively”. The filtered senses serve as seeds for the next step, which finds all associated synsets of the seed synsets. This extended set of synsets is then enriched with related synsets. The following relations are considered:

- Hypernyms: Y is a hypernym of X if every X is a Y: “Consider” and “think about” are hypernyms of contemplate. Only hypernyms up to the second degree were considered.
- Hyponyms: Y is a hyponym if every Y is an X: “Introspect” is a hypernym of “soul-searching”, “self-analysis” and “examination”. Only hyponyms up to the second degree were considered.
- Coordinate Terms: The synset siblings (hyponyms of its hypernyms): “puzzle over”, “rationalize”, “think”, “philosophize”, “brainstorm”, etc.
- Verb groups: A group a verb belongs to.
- Synonyms: X is synonym of Y.

---

<sup>2</sup> <http://wordnet.princeton.edu/>

<sup>3</sup> <http://sourceforge.net/projects/jwordnet/>

<sup>4</sup> <http://wordnet.princeton.edu/wordnet/man/wnstats.7WN.html>



Based on four seed words, this process generates 194 synsets containing 416 words (nouns, verbs, adjectives, and adverbs). Positive examples of this list with their WordNet glossary are:

- Think (noun): an instance of deliberate thinking; "I need to give it a good think".
- To muse (verb): reflect deeply on a subject; "I mulled over the events of the afternoon"; "philosophers have speculated on the question of God for thousands of years"; "The scientist must stop to observe and start to excogitate".
- Wondering (adjective): showing curiosity; "if someone saw a man climbing a light post they might get inquisitive"; "raised a speculative eyebrow".
- Reflectively (adverb): in a reflective manner: "he watched her reflectively".

Negative examples include:

- wisecrack (noun): witty remark.
- dally (verb): to consider not very seriously; "He is trifling with her"; "She plays with the thought of moving to Tasmania".
- highbrowed (adjective): highly cultured or educated; "highbrow events".

As the goal of this annotator is to find keywords that are related to reflection, this set of words can serve as a starting point, as most words are positive hits. However the word list has to be manually refined to compensate for associations, which are according to WordNet associated with reflection, but seem not suited for the purpose of this annotator.

The annotator uses regular expressions to find sentences, which contain these words. By now the annotator uses the infinitive form of the words and does not take into account inflections.

The annotator can be aggregated with an annotator for self-references (I, my, me, myself, mine, etc.), to annotate sentences, which are referred to own inner reflective thoughts.

#### **4.5 Premise and Conclusion Annotator**

The premise and conclusion annotator takes the above outlined indicator words to match them using regular expressions. Some of the indicator words, however, need special attention. For example the word "since" needs special treatment, as it can be used as an indicator word for a premise, but it can also be used to express a temporal dimension. Another example is the word "for". It can be used in the sense of because, but it can also be used in other senses, e.g. "for the sake of", "to be all for it", "what for", etc.

#### **4.6 Question Annotator**

The question annotator consists of three parts. It uses again regular expressions to identify all sentences ending with a question mark, sentences containing interrogative words (for example why, how, what), and sentences, which follow Kings [10] blue-

print of thought provoking questions. A sentence, which is annotated as thought provoking and contains an interrogative word, and a question mark, is annotated three times. In the analysis process only one of this three annotations can be considered for further use.

## 5 Annotated Example

The first example is taken from the resource chapter of Moon [9], which is highlighted as a “reasonable reflective writing”. From the left to the right it shows the premise and conclusion, the reflective word, and the question annotator in action. The detected parts of the text are highlighted. It shows that the analysis engine detects words and sentences according to the outlined indicator words and can distinguish between the three types. What is not visible in the examples is that every annotator consists of a finer level of detection. For example premise and conclusion are two distinct features, and questions consist of simple questions and thought-provoking questions.

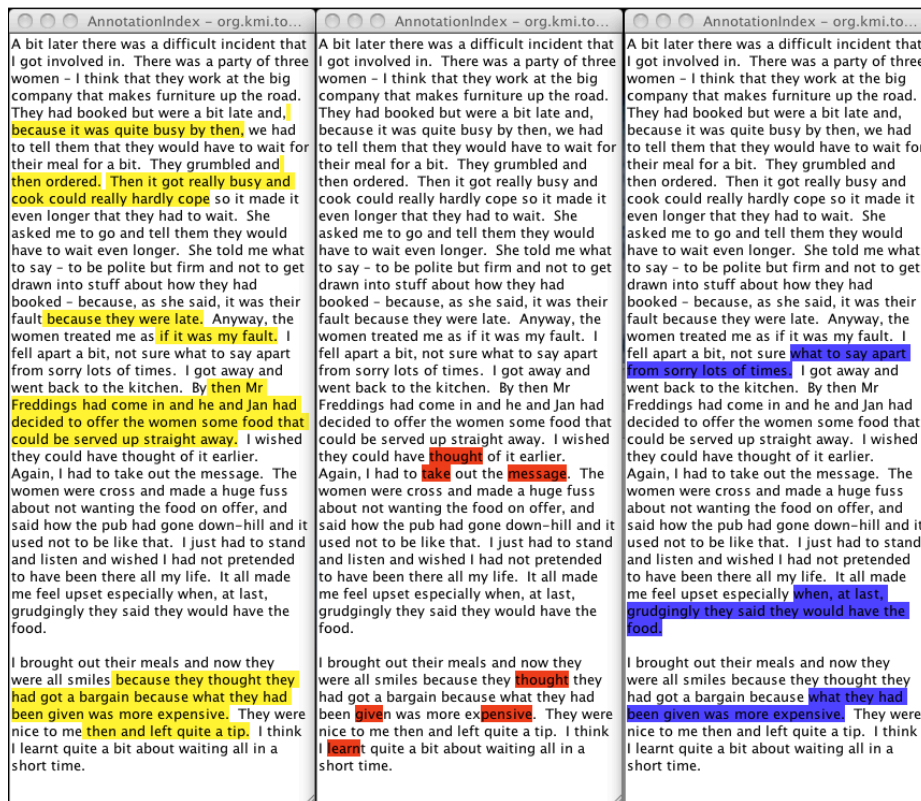


Fig. 3. Annotations of a reflective account (left: premise/conclusion, center: reflective key-words, right: questions).

The second text is taken from the Wikipedia article about awareness. As it is an encyclopedic article the purpose of the writing is to present a result and not the reflective process, which lead to this article. Again the three annotations are shown.

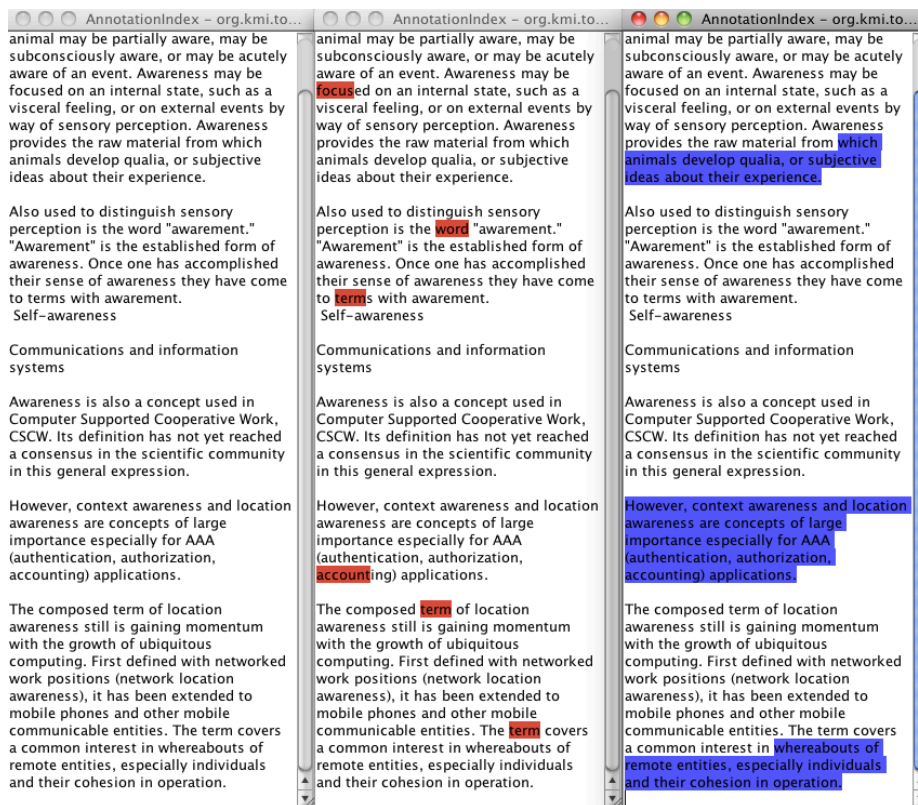


Fig. 4. Annotations of the encyclopedic article (left: premise/conclusion, center: reflective keywords, right: questions).

Besides the correct annotation according to the indicator words it can also be said that this three annotators can be useful on its own, for example to find all questions, or premises, or conclusion of a learning network and to present them in a meaningful way. Used only separate however, they say little about what a reflective writing is and what not.

As can be seen in both examples, the three annotators are not always identifying reflective accounts as intended, leaving room for further fine-tuning. Still, they manage to identify a set of relevant reflective acts expressed in the text. Even if single indicators have lower accuracy, in sum they bear the potential to flag the level of reflectivity to the interested analyzing person.

## **6 Evaluation**

There are several ways of how to evaluate the quality of the detection. It largely depends on the purpose of the reflection detector. In the case of a supporting tool for content analysis, as it was in the case of the described in the related approach section, the goal would be to achieve high agreement between human coders and the automated annotation. The level of detail of the analysis is usually on paragraph level, sentence level and the whole document. The annotations of each level of text can then be compared with the automated annotation to calculate agreement. One of the most common measurements is the inter-coder reliability (inter-rater, or interjudge reliability), which expresses the degree of agreement between independent raters. Cohen's kappa is one of the most common measure for inter-coder reliability. The outcome of Cohen's kappa is between  $<0$  and  $1$ , while a value  $> .7$  is seen as an acceptable agreement between raters (however this value varies in the literature). Precision and recall of human annotations of texts can complement the picture, particularly as the calibration of detection algorithms have to find the best possible equilibrium between high precision and high recall (often complemented by the f-measure, a combination measure of both).

Another way of evaluating the quality of the detection is to evaluate the usefulness of applications for people in learning networks, which are based on top of the detector architecture. The applications would be tailored to raise awareness about elements of reflections in online learning networks. This could be for example a reflection search engine, or a feed containing only reflective contents, or mash-ups based on reflective contents. Acceptance of the tools and their usefulness would then be the starting point of evaluation. [11].

## **7 Conclusions and Outlook**

One of the benefits of the proposed architecture is that annotators can be independently developed and plugged into the framework to enrich the reflection detector with further elements. As a starting point of the development of the reflection detector a keyword based approach was chosen, however the techniques from the mentioned content classifiers seem to be promising and relevant for the next set of annotators.

The assumption was that there exists a set of words, which reveal reflection. However, we have to consider that a person writes in a reflective manner without using any of these words, or that they use these marker words without being reflective at all. On a general level these words have to be seen as indicators for reflective thinking and not to be mistaken with reflection: it is only possible to inspect evidence of reflection expressed in texts. For assessment of reflective capabilities of humans, this therefore means that it is subject to the assessment set-up, instruction, and method to show that it could validly be used to detect such competence. However, this is not only a problem for automated reflection detection and shared by 'manual' detection. The relatively low inter-coder reliability presented in the introduction expresses the diffi-

culty of human coders to evaluate what reflection is and what not, especially for fine-grained methods.

Another challenge of future work can be seen in the intelligent combination of the detected elements of reflection to ultimately indicate reflection, and furthermore to indicate depth levels of reflection. The architecture therefore foresees an analysis component.

As literature indicates, it would be interesting to investigate emotions and connect this work with sentiment detection in texts, as e.g. feeling of puzzlement seem to be strongly connected to reflection.

As the architecture foresees web-services as the central access point, this open infrastructure will allow integrating the annotated data into mash-ups of learning and research networks. The intended applications are not necessarily dependent on the most accurate reflection detection, as the goal is to support learning networks with awareness indicators of reflection. In this case, having a multitude of indicators could be more rewarding than restricting to the few that yield high precision.

## **8 Acknowledgements**

I want to especially thank my supervisors Fridolin Wild and Peter Scott for their advice and the three reviewers for their recommendations. This work has been funded by the European Union in the projects Stellar under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme under the grant agreement number 231913.

## **9 References**

1. Wong, F.K., Kember, D., Chung, L.Y.F., CertEd, L.Y.: Assessing the level of student reflection from reflective journals. *Journal of Advanced Nursing*. 22, 48-57 (1995).
2. Sumsion, J., Fleet, A.: Reflection: can we assess it? Should we assess it? *Assessment & Evaluation in Higher Education*. 21, 121 (1996).
3. Corich, S., Kinshuk, L.M.: Measuring Critical Thinking within Discussion Forums using a Computerised Content Analysis Tool. *the Proceedings of Networked Learning*. (2006).
4. Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., Fischer, F.: Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*. 3, 237-271 (2008).
5. Dewey, J.: *How we think: A restatement of the relation of reflective thinking to the educative process*. DC Heath Boston (1933).
6. Halpern, D.F.: *Thought and Knowledge. An Introduction to Critical Thinking*. L. Erlbaum Associates (2003).
7. Pintrich, P.R.: The role of goal orientation in self-regulated learning. In: Boekaerts, M., Pintrich, P.R., and Zeidner, M. (eds.) *Handbook of self-regulation*. pp. 451–502 Academic Press, San Diego, CA (2000).

8. Atkins, S., Murphy, K.: Reflection: a review of the literature. *Journal of Advanced Nursing*. 18, 1188-1192 (1993).
9. Moon, J.A.: *A handbook of reflective and experiential learning*. Routledge (2004).
10. King, A.: Inquiry as a tool in critical thinking. *Changing college classrooms: New teaching and learning strategies for an increasingly complex world*. 13-38 (1994).
11. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. *MIS quarterly*. 425-478 (2003).