Ontologies: A Wish List for the Humanities

Robert C. Kahlert (rck@cyc.com)
Cycorp Inc.
3721 Executive Center Drive
Austin, TX 78731

Jennifer Sullivan
(jsullivan@converra.com)
Convera Inc.

Abstract
In this paper, we analyze the problems the Humanities currently face and discuss how ontologies could help them – provided the ontology toolkits had support for additional capabilities. We enumerate these capabilities and describe their necessity for research in the Humanities.
Keywords: logical contexts, ontologies, argumentation, test corpora, CYC.

1 Introduction

The Humanities are a late-comer to the ontology game. As much as we are encouraged by signs of growing interest, we will argue that it is essential for the Humanities to leverage fully the research possibilities that ontologies offer. We will justify this claim through an analysis of the problems that the Humanities currently face.
However, as a result of the community's delayed involvement, we perceive the playing field to be shaped by the needs of the early adopters. The capabilities of the ontology toolkits currently available are dominated by the needs of the Natural Sciences – esp. the medical[1] and the pharmaceutical disciplines – and the Engineering Sciences.
We propose that the different research needs of the Humanities emphasize different capabilities and identify five areas of modeling capability that the Humanities need: ontological plurality, contexts, argumentation systems, question corpora management systems and representational expressivity. These capabilities together form a wish list addressed to the ontology toolkit development community. We conclude by showing that the Humanities have allies in requesting these features and make preliminary suggestions for the order in which the features should be pursued.

2 The urgent Need for Ontologies

A deluge of Literature

The basic task of the Humanities is to use texts to generate new texts. The researcher draws upon texts to build up a mental model of a specific subject matter and focus and produces a new text that exposes both the construction process and the result to peer review. The resulting text is fed back into the cycle of interpretation and becomes basis for new considerations, ideas and models, for the researcher as well as the community.
With the productivity of modern technology and the professionalization of the Humanities in the universities since the 19th century, this model of working has become untenable. *Niebelungenlied* scholar Siegfried Grossen sums it up nicely when he writes:

> The literature on … the *Niebelungenlied* … [is] unbelievably huge and cannot be surveyed by an individual, much less evaluated. The plentitude becomes visible, for example, when comparing the bibliographies included with various essays: They always differ from each other and have only partial overlap.[2]

The quality of the mental model depends critically on having considered as many different arguments as possible in its construction; yet sheer quantity of texts (and thereby potentially different arguments)

[1] For example, the well-known ontology toolkit Protégé came out of the medical and bio-informatics community at Stanford.
[2] (Grossen, 2002), p.719: "Die Literaturmenge ... des Niebelungenliedes (NL) ... ist unvorstellbar groß und vom einzelnen nicht zu überschauen, geschweige denn zu werten. Die Fülle wird z.B. beim Vergleich von Literaturverzeichnissen sichtbar, die den Abhandlungen beiliegen: denn sie weichen stets von einander ab und sind nur partiell deckungsgleich." (Translation rck)

exceeds the reconstructive capabilities of human researchers. What ends up getting published involves slices of the text space only and is therefore itself fragmentary.

The Text as Part of the Problem

Part of the problem is the text as the research product. Western style education amply prepares researchers for reading, analyzing and creating texts. But a text is the end-product of a laborious process, while it is the steps of the process that interest during peer-review and criticism. Each reviewer afresh has to rebuild the model mentally, that is, reverse-engineer the argument structure, check the supporting sources, review the classifications and usage of terminology, validate that the conclusions follow from the premises or detect omissions in terms of known sources or prior research.

The Promise of Ontologies

Much of this checking is rote symbolic reasoning, the type of well-understood symbolic busy work that automatic theorem provers (ATPs) and expert systems were designed to handle. One is tempted to observe that a text might be no more than a print-out generated by an expert system that we can no longer interact with.[3] Ontologies in the broad sense of the term, as rich semantic networks of logical terms and relations that are actionable by symbolic reasoning mechanisms, hold the promise of redefining the research product of the Humanities.

In the following, we will take this – admittedly speculative – premise and derive from it a wish list of the capabilities that ontology construction and reasoning tools have to provide in order to assist the Humanities in producing ontologies. These capabilities follow from the structure of the research objects in the Humanities and may differ decisively from the needs of prior champions of ontology research, such as the medical or pharmaceutical community.

We believe that these "wishes" are reasonable; after all, at least one ontology toolkit, CYC (Lenat, 1995), already implements significant parts of all of these features. We will therefore employ its knowledge representation language, CycL, to provide examples that illustrate our argument.[4] As our guiding exemplar we will use a historical reconstruction of aspects of the Salem Witch trials.[5]

3 An Ontological Wish List

Wish #1: Ontological Plurality

If every text reflects the underlying mental ontology of its author(s), then it is the duty of the researcher to attempt to reconstruct the text in terms of that ontology. This means that ontological toolkits for Humanities research need to be able to deal with a minimum of two ontologies at the same time, namely the mental ontology of the author and the mental ontology of the reader. In some sense, textual interpretation is the act of proposing a formal mapping between the mental ontology of the author and the mental ontology of the reader that explains the differences.

Ontological plurality is also important for the process of reconstructing the ontology of the author from the sources. There may be reasons for and against classifying terms in a specific way, and the researcher needs the ability to try multiple ontological assignments simultaneously and inspect their consequences to determine which one best fits the sources.

Finally, the number of concurrent ontologies increases further once researchers begin to include the prior work of the discipline, as this will reflect the ontologies their colleagues constructed when interpreting the texts.

---

[3] The large index card systems that scholars such as Adolf von Harnack or Franz Overbeck were constructing in the late 19th century are paper expert systems without any symbolic reasoning apparatus – other than their maintainers!

[4] The literature on CycL is sparse at the moment; reasonable documentation can be found on the OpenCYC website at http://www.opencyc.org/doc/#ke.

[5] The literature for the Salem Witch trials is no less than it is for the song of the Niebelungs; see (Kahlert et al, 2006) for a small selection of literature highly relevant to this research problem.

Example: Robert Calef *versus* Cotton Mather

One interesting aspect of the Salem Witch trials is the literary "quarrel" between the Bostonian writer Robert Calef and the Bostonian clergyman Cotton Mather. Both were contemporary witnesses, involved to various degrees, and published accounts and interpretations of the events during their lifetimes. However the presentations they give of each other are difficult to reconcile. There are basically two positions: Either Cotton Mather was a rabid witch hunter who exploited the hands-on healing of young Margareth Rule for his lecherous interests and Calef the astute witness who exposed the hypocrite; or Robert Calef was a libelous self-styled wit who enlarged his social standing from weaver to merchant and purposefully employed ambiguous constructions in his narratives to accuse the Mathers of smutty things – possibly with the intent of discrediting them and thereby curtailing their influence on the proceedings. Deciding this question is crucial for weighing all of the other contributions that these two make to the Salem Witch trials.[6]

```
;; Mather as lecher in disguise
Constant: CottonMather-BostonClergyman.
isa: MalePerson, Lecher.
Constant: MargarethRule-Bostonian.
isa: FemalePerson, WantonWoman.
Constant: RobertCalef-Bostonian.
isa: MalePerson, Merchant.
Constant: HealingOfMargarethRule.
isa: HumanSexualBehavior, ImmoralAct.
objectActedOn: (AnatomicalPartOfFn Breast MargarethRule).
actors: CottonMather-BostonClergyman, IncreaseMather-BostonClergyman.
observers: RobertCalef-Bostonian.

;; Robert Calef as paparazzi
Constant: CottonMather-BostonClergyman.
isa: MalePerson, (FrequentPerformerFn Healing).
Constant: MargarethRule-Bostonian.
isa: FemalePerson.
Constant: RobertCalef-Bostonian.
isa: MalePerson, Weaver.
Constant: HealingOfMargarethRule.
isa: HealingEvent.
objectActedOn: (AnatomicalPartOfFn Stomach MargarethRule).
actors: CottonMather-BostonClergyman, IncreaseMather-BostonClergyman.
observers: RobertCalef-Bostonian.
```

Wish #2: Contexts

One way to implement ontological plurality is by supporting contexts, both for ontology building and for reasoning (Guha, 1991). In terms of reasoning, contexts provide a way of compartmentalizing mutually contradictory information without inferring a contradiction (from which logically any fact could follow).[7]

Contexts provide a straight-forward way to simulate the mental models that researchers construct when they work with texts. At a very coarse level, each of the texts and each of the research positions – both own and others – can be assigned their own context and reconstructed autonomously within.

Contexts can however be applied in more fine-grained manners, for example to model the individual steps of the reconstructive process. By assigning separate contexts to the textual reconstruction, the translation, the analysis of the textual form, the analysis of how the text was edited and the investigations into the traditions and motives leveraged by as well as inspired by the text, the researcher can mix and match possible models for each of these steps simply by bundling contexts

[6] This presentation is a simplification for pedagogical purposes – the actual quarrel was more subtle than that. The material follows (Hansen 1969, 190ff).

[7] For a more detailed discussion of the use of contexts see our other submission to this workshop, *Microtheories*.

together. Furthermore, prior research can be leveraged in equivalent ways, giving a building block quality to the investigation of alternative interpretations.
Finally, contexts make for handy units of information exchange between researchers who want to collaboratively construct sets of ontologies.

Example (cont.): Putting Calef vs. Mather in Contexts

The contextualizing approach to maintaining contradictory descriptions of an event such as the healing of Margareth Rule effectively places the two descriptions in separate ontologies. In CycL this is done via microtheories:

```
In Mt: MatherTheLecher-InterpretationMt.
;; as lecher in disguise
;; … the healing was faked and the proceedings were smutty …

In Mt: CalefTheLiar-InterpretationMt.
;; Robert Calef as paparazzi
;; … the healing was successful and the reporting was smutty …
```

However, notice that these two interpretations, despite their contradictory conclusions, share several assumptions. In a contextualizing approach, these would be factored out and inherited. The following snippet of CycL represents the shared assumptions of the two points of view:

```
In Mt: HealingOfMargarethRule-BackgroundMt.
Constant: CottonMather-BostonClergyman.
isa: MalePerson.
Constant: MargarethRule-Bostonian.
isa: FemalePerson.
Constant: RobertCalef-Bostonian.
isa: MalePerson, PersonWithOccupation.
Constant: HealingOfMargarethRule.
isa: Event.
objectActedOn: (AnatomicalPartOfFn HumanBody MargarethRule).
actors: CottonMather-BostonClergyman, IncreaseMather-BostonClergyman.
observers: RobertCalef-Bostonian.
```

Context inheritance makes this information accessible for the interpretation viewpoints.

```
In Mt: BaseKB.
f: (genlMt MatherTheLecher-InterpretationMt
      HealingOfMargarethRule-BackgroundMt).
f: (genlMt CalefTheLiar-InterpretationMt
      HealingOfMargarethRule-BackgroundMt).
```

In the more specific contexts, the information that was already provided can now be augmented in ways that expresses the biases of the positions. Sometimes the additional information augments the inherited information; sometimes more precise restatements override the inherited.

```
In Mt: MatherTheLecher-InterpretationMt.
f: (isa CottonMather-BostonClergyman Lecher).
f: (isa MargarethRule-Bostonian WantonWoman).
f: (isa RobertCalef-Bostonian Merchant).
f: (isa HealingOfMargarethRule HumanSexualBehavior).
f: (isa HealingOfMargarethRule ImmoralAct).
f: (objectActedOn HealingOfMargarethRule
     (AnatomicalPartOfFn Breast MargarethRule)).

In Mt: CalefTheLiar-InterpretationMt.
f: (isa CottonMather-BostonClergyman (FrequentPerformerFn Healing)).
f: (isa RobertCalef-Bostonian Weaver).
```

```
f: (objectActedOn HealingOfMargarethRule
     (AnatomicalPartOfFn Stomach MargarethRule).
```

Notice that we called the contexts *interpretation* microtheories. Indeed, one important problem to consider is that the beliefs, desires and interpretations are forms of modal statements. That means they do not obey the normal forms of deductive inference. For example, Robert Calef might have believed that the healing of Margareth Rule was a sexual activity, and Calef might also have thought that watching others partake in sexual activity was voyeurism – but from this it does not follow that Calef believed himself to be a voyeur. At the same time, for the purposes of historical reconstruction it is important to perform deductive inference on the contents of the beliefs held by the agents under study. The proper way to deal with this distinction is to employ two contexts, one context that captures the beliefs and one context that captures all of the consistent worlds that follow from the contents of the beliefs.[8]

```
In Mt: BaseKB.
f: (consistentWorldProjectionMt RobertCalef-BeliefMt
      RobertCalef-ConsistentWorldViewMt).

In Mt: RobertCalef-BeliefMt.
f: (beliefs RobertCalef-Bostonian
      (isa CottonMather-BostonClergyman Lecher)).
f: (beliefs RobertCalef-Bostonian
      (isa HealingOfMargarethRule HumanSexualActivity)).
;; … etc …

In Mt: RobertCalef-ConsistentWorldViewMt.
f: (isa CottonMather-BostonClergyman Lecher).
f: (isa HealingOfMargarethRule HumanSexualActivity).
;; … etc …
```

Calef's interpretation of the healing of Margareth Rule simply inherits this consistent view of the world according to Calef. This nicely models the way human presuppositions color their interpretations of events.

```
In Mt: BaseKB.
f: (genlMt MatherTheLecher-InterpretationMt
      RobertCalef-ConsistentWorldViewMt).
```

Finally, this model agrees with our intuition that in the `RobertCalef-ConsistentWorldViewMt`, one might indeed infer that Calef was acting voyeuristically in watching a sexual activity out of his own free will, while this fact cannot be proven in the model of his belief, the context `RobertCalef-BeliefMt` due to the modal representation.

Wish #3: Argumentation Systems

The reconstruction of a text is an argument that assigns meanings to the linguistic elements – sentences, phrases, words – identified within the text, drawing upon lexical knowledge, common sense and the text itself.[9] In order to expose such a reconstruction to peer criticism, the argumentative links between the elements have to be made explicit in the ontology. The ontological toolkits need to support qualifying argumentative links, for example as an application of the rule-of-thumb of *lectio brevior* or *lectio difficilior*.[10]

---

[8] This approach of interpreting modal operators is inspired by the work of Jaako Hintikka and Robert Stalnaker.

[9] This is under the assumption that the paleographical analysis of the text has been completed; if the deciphering of the manuscripts is still part of the scholarly debate – cf. the 1995 discovery of apparently text-critical "umlauts" in the *Codex Vaticanus* by Philip Payne – then there is a prior step, which establishes a mapping of graphemes to linguistic elements (or meta-elements, as in the case of the text-critical "umlauts").

[10] *Lectio brevior* and *lectio difficilior* are two heuristics for textual interpretation when reconstructing the reading of a manuscript. *Lectio brevior* argues that the shorter text is usually the earlier one (and therefore closer to the

Modeling the argumentative network explicitly in the ontology has additional benefits: By making the connections between the facts of the arguments obvious, the ramifications of any modifications to the supporting facts become immediately clear as well. The arguments become accessible to automatic truth revision when assumptions or supports change. For example, when an application of a rule-of-thumb such as *lectio difficilior* becomes implausible due to additional evidence, then all dependant parts of the argument chain can be retracted automatically by the ontology toolkits.
Of course, in an ontology toolkit that supports contexts, it may be part of the exploration of the problem space for the researcher to problematize some of the assumptions, just to observe the effects on the argument chain.
Finally, the assignment of argumentative links between any fact and an interpretation should not be limited to only one link, or even to several links of only one truth value. Indeed, in a heuristic enterprise such as textual interpretation, there may be pro and contra arguments with regards to a specific reading of the text. For example, if the longer reading should also turn out to be the more convoluted one, then this effectively pitches *lectio brevior* against *lectio difficilior*. Therefore ontology toolkits should provide for a way of enumerating the pro and contra supports for a specific step in the reconstructive argument.

Example (cont.): Capturing Calef's reasoning

Stepping back from the stark dichotomy that Hansen suggested, there are two possibilities why Calef recounted the healing of Margareth Rule as he did; he was either convinced that Mather was a lecher and therefore fishing for supporting evidence when Mather attended to Ms Rule. Or he happened to be present at the healing of Ms Rule, interpreted what he saw in sexual terms, and decided that Mather was a lecher who needed to be exposed. In terms of modeling, we are asking which of the following two statements captures best how Calef correlated his convictions:[11]

```
;; Choice #1: fishing for evidence
In Mt: RobertCalef-BeliefMt.
f: (explains-PropProp
      (beliefs RobertCalef-Bostonian
         (isa CottonMather-BostonClergyman Lecher))
      (beliefs RobertCalef-Bostonian
         (isa HealingOfMargarethRule HumanSexualActivity))).

;; Choice #2: exposing the insight
In Mt: RobertCalef-BeliefMt.
f: (explains-PropProp
      (beliefs RobertCalef-Bostonian
         (isa HealingOfMargarethRule HumanSexualActivity))
      (beliefs RobertCalef-Bostonian
         (isa CottonMather-BostonClergyman Lecher))).
```

There is some evidence pointing toward Choice #1; Calef writes that he had purposefully chosen the night at which he visited Ms Rule to observer the Mathers, which suggests that he was fishing for something to expose about them, maybe to curb their influence in the witch hunts.

```
In Mt: RobertCalef-BeliefMt.
f: (explains-PropProp
      (beliefs RobertCalef-Bostonian
         (isa CottonMather-BostonClergyman WitchHunter))
      (ist HealingOfMargarethRule-BackgroundMt
         (observers HealingOfMargarethRule RobertCalef-Bostonian))).
```

---

original and to be preferred), as later writers tend to expand, clarify or embellish; alternatively, the heuristic of *lectio difficilior* argues that the more difficult reading is usually the earlier one (and therefore closer to the original text and to be preferred), as later writers tend to clarify or otherwise remove difficulties for the readers.

[11] For presentation reasons, we are leaving aside the question of how to adequately represent the temporal ordering of these two beliefs for right now.

Wish #4: Question Corpora Management Systems

Research is defined by the questions that the researchers ask. As the researcher reconstructs the meaning of a text or investigates its prior interpretations, the set of research questions does not remain static: Some questions find answers; new questions are identified; old questions are deemed infertile and placed aside.

The ontology toolkits should support the representation of the research questions – presumably in the form of queries against the ontologies – as first-order objects. Furthermore, the toolkits should make it easy to track the stability of acquired answers as the ontology building proceeds. Akin to the way test- and use-cases drive the development of software artifacts, modeling the questions and monitoring the validity of their answer sets can help focus the research effort and ensures that modifications to the theory do not endanger prior gains.

In addition, corpora of questions that a specific reconstruction can answer provide an additional hook for peer criticism. Competing reconstructions now become comparable at the level of the kinds and the number of questions that they can answer. Intuitively, the "simplest" reconstruction that can answer "all" research questions to the "satisfaction" of the research community should be the preferred one.[12] Notice that such an approach will require the construction of shared repositories of research question sets – akin to the Penn Treebank (Marcus et al, 1993) in natural language processing, the TREC competition[13] in information retrieval or the TTPTP repository (Sutcliffe et al, 1998) in the automatic theorem prover (ATP) community.

Finally, there is something pragmatically valuable about such corpora of research questions. There exist many ways to represent any domain ontologically. Many proposals, especially where concerning upper ontology representations, in principle have much to be said for them. However, having an actual application context would allow the research community to determine quickly which proposals merit consideration and which ones are notational possibilities without any practical benefits to recommend them.

Example (cont.): Some Research Questions about Calef and Mather

We finish our discussion with two research (and test) questions about Calef and Mather that could for example be used to track the development of a research question corpus about the Salem Witch trials.

```
In Mt: SalemWitchTrials-ResearchQuestionMt.
Constant: HowDidCalefLibelMather-RQ.
isa: ResearchQuestion.
questionSentence:
(ist RobertCalef-InterpretationMt
  (and
    (isa ?X ImmoralAct)
    (actors ?X CottonMather-BostonClergyman))).

Constant: HowDidMatherInterpretTheseEvents-RQ.
isa: ResearchQuestion.
questionSentence:
(and
  (isa ?X Event)
  (ist RobertCalef-InterpretationMt
    (isa ?X ImmoralAct)
    (actors ?X CottonMather-BostonClergyman))
  (ist CottonMather-InterpretationMt
    (unknownSentence
       (isa ?X ImmoralAct)))).
```

[12] The quotes around words such as "all" and "satisfaction" indicate that this is consciously a radical simplification of the social aspects of scientific research; but the advantages that would accrue if such clear measurements of progress were identified are equally obvious.

[13] TREC Text Retrieval Conference, http://trec.nist.gov/

Wish #5: Representational Expressivity

We argued above that the Humanities cannot perform their research without contexts. Notice that this need entails switching from description or first order predicate logic to modal or higher-order logic – depending on the implementation of the contexts.[14] This means that "computational completeness (all entailments are guaranteed to be computed) and decidability (all computations will finish in finite time)" – cf. (Smith et al, 2004) – are no longer guaranteed.

On the other hand, once incompleteness and undecidability have been incurred – and we believe this step to be unavoidable for the Humanities – one may as well exploit the representational benefits of modal and higher order logic. Modal and higher order representations are usually far more concise, which lessens the burden placed on the researcher.

4 Conclusion

We argued that the Humanities cannot perform their research without ontology toolkits that handle ontological plurality and contexts, provide argumentation and question corpora management systems, as well as support adequate representational expressivity.[15] We documented the possibilities of such an approach with the quarrel between two primary sources in the Salem Witch Trials, Robert Calef and Cotton Mather. We believe that these requests are not unreasonable, because at least one ontology toolkit, CYC (Lenat, 1995), already implements all of these features to a significant extent.

The support in CYC for these five requests is at least partially the result of successfully applying CYC to problems of intelligence analysis, a military discipline that is very proximate to the Humanities in its needs.[16] This in turn suggests that intelligence analysis and its sibling in the business world make for suitable allies of the Humanities in requesting improved support for these capabilities from the ontology toolkit community.

Given a clear set of goals and potential allies, the Humanities are well-positioned to acquire the research infrastructure required to fundamentally change their research culture. Strategically, it seems most appropriate to begin with the preliminary work for question corpora.[17] We predict that such work of identifying research question exemplars and proto-ontologizing these will create the means for communicating the needs of the Humanities to the ontology toolkit community at large.

5 Bibliography

Saša Buvač (1996), *Resolving lexical ambiguity using a formal theory of context*, in: Kees van Deemter, Stanley Peters, (ed.) (1996), *Semantic Ambiguity and Underspecification*. CSLI Lecture Notes, Center for Studies in Language and Information, Stanford, CA.

Saša Buvač, Vanja Buvač, Ian Mason (1995), *Metamathematics of Contexts*, in: *Fundamenta Informatica* 23(3), 1995.

Chris Deaton, Blake Shepard, Charles Klein, Corinne Mayans, Brett Summers, Antoine Brusseau, Michael Witbrock (2005), *The Comprehensive Terrorism Knowledge Base in Cyc*, in: *Proceedings of the 2005 International Conference on Intelligence Analysis*, McLean, Virginia, May 2005.

R.V. Guha (1991), *Contexts: A Formalization and Some Applications*, Stanford University, STAN-CS-91-1399-Thesis, 1991.

[14] Cf. (Buvač, 1995); (Buvač et al, 1995); (Ramachandran, 2006).

[15] It may come as a surprise to practicing users of ontology toolkits that better knowledge acquisition tools are not part of the wish list. While we agree that the current knowledge acquisition tools are woefully inadequate, improving them is already a domain of serious research; furthermore, it is not a problem peculiar to the Humanities.

[16] For Lenat's pre-CYC work on intelligence analysis and a good introduction to the domain, cf. (Oresky, 1991). More recently, cf. Deaton et al (2005); Siegel et al (2005); Schneider et al (2005).

[17] See also (Kahlert et al, 2006) and the project it inspired, *Clio knows* (http://clio-knows.sourceforge.net).

Siegfried Grossen (ed.) (2002), *Das Niebelungen-Lied: Mittelhochdeutsch // Neuhochdeutsch*, Stuttgart (Reclam).
Chadwick Hansen (1969), Witchcraft at Salem, Braziller.
Robert C. Kahlert, Jennifer Sullivan (2006), *Toward a Knowledge Representation Corpus of Historical Events*, in: *10th International Conference on Principles of Knowledge Representation (KR'06)*.
Douglas B. Lenat (1995(, *Cyc: A Large-Scale Investment in Knowledge Infrastructure*, in: *Communications of the ACM* 38 (11), 1995.
Douglas B. Lenat (1999), *The Dimensions of Context-Space*, http://www.cyc.com/doc/context-space.pdf.
Mitchell P. Marcus, Beatrice Santorini, Mary Ann Marcinkiewicz (1993), *Building a large annotated corpus of English: The Penn Treebank*, in: *Computational Linguistics*, 19(2):313–330.
Deepak Ramachandran, Eyal Amir, Pace Reagan, Keith Goolsbey, K. Keefe (2006), *Inference Friendly Translations of ResearchCyc to First-Order Logic*, in: *10th International Conference on Principles of Knowledge Representation (KR'06)*.
Dave Schneider, Cynthia. Matuszek, Purvesh. Shah, Robert C. Kahlert, David Baxter, John Cabral, Michael. Witbrock, Doug. Lenat (2005), *Gathering and Managing Facts for Intelligence Analysis*, in: *Proceedings of the 2005 International Conference on Intelligence Analysis*, McLean, Virginia, May 2005.
Nick Siegel, Blake Shepard, John Cabral, Michael Witbrock (2005), *Hypothesis Generation and Evidence Assembly for Intelligence Analysis: Cycorp's Nooscape Application*, in: *Proceedings of the 2005 International Conference on Intelligence Analysis, McLean*, Virginia, May 2005.
Michael K. Smith, Chris Welty, Deborah L. McGuinness (2004), *OWL Web Ontology Language Guide*, http://www.w3.org/TR/owl-guide/, 2004.
TREC Text Retrieval Conference, http://trec.nist.gov/ (via http://web.archive.org/web/20041128090645/)
G. Sutcliffe, C.B. Suttner (1995), *The TPTP Problem Library: CNF Release v1.2.1*, in: *Journal of Automated Reasoning*, 21(2):177–203.