Detecting Knowledge-Level Claims in Research Articles

Ágnes Sándor
Xerox Research Centre Europe
Agnes.Sandor@xrce.xerox.com

special credit to
Frédérique Lisacek (Genebio)
Simon Buckingham Shum (KMI, OU)
Anna de Liddo (KMI, OU)

OAI8 20. June 2013

xerox

Knowledge-level claims

In contrast with previous hypotheses, **compact plaques form before significant deposition of diffuse A beta …**

The WFS1 protein is a glycoprotein located in the endoplasmic reticulum (ER) membrane but its function is poorly understood.

Ex vivo gene therapy is emerging as a promising approach **for the treatment of neurodegenerative diseases and central nervous system (CNS) trauma.**

Knowledge-level claims

In contrast with previous hypotheses, compact plaques form before significant deposition of diffuse A beta … contrast

The WFS1 protein is a glycoprotein located in the endoplasmic reticulum (ER) membrane but its function is poorly understood.

Ex vivo gene therapy is emerging as a promising approach **for the treatment of neurodegenerative diseases and central nervous system (CNS) trauma.**

Knowledge-level claims

In contrast with previous hypotheses, compact plaques form before significant deposition of diffuse A beta … *contrast*

The WFS1 protein is a glycoprotein located in the endoplasmic reticulum (ER) membrane but its function is poorly understood.

open question

Ex vivo gene therapy is emerging as a promising approach **for the treatment of neurodegenerative diseases and central nervous system (CNS) trauma.**

Knowledge-level claims

rhetorical formulas scientific facts

In contrast with previous hypotheses, compact plaques form before significant deposition of diffuse A beta …

contrast

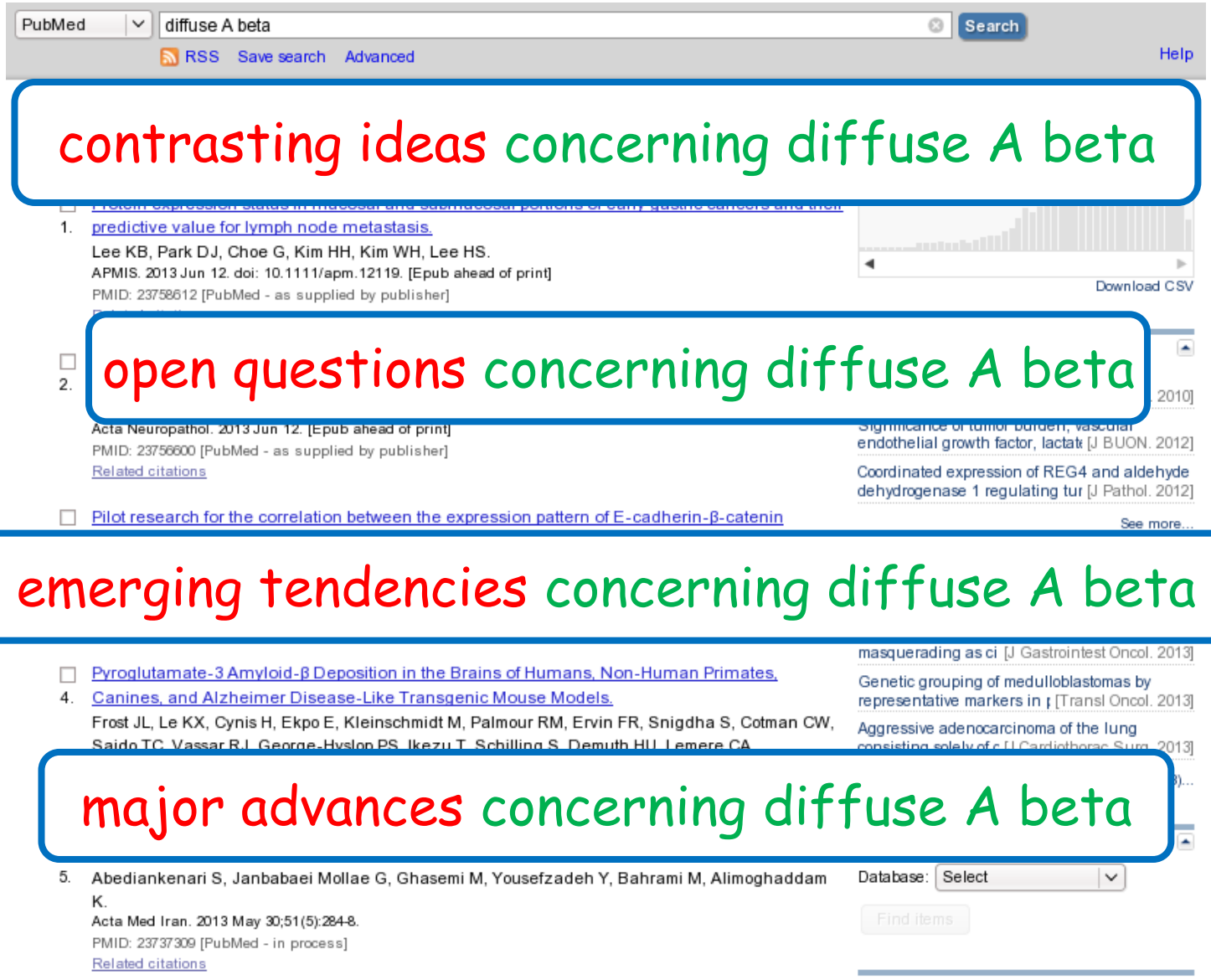
The WFS1 protein is a glycoprotein located in the endoplasmic reticulum (ER) membrane but its function is poorly understood.

open question

Ex vivo gene therapy is emerging as a promising approach for the treatment of neurodegenerative diseases and central nervous system (CNS) trauma.

emerging tendency

Knowledge-level claims



Knowledge-level claims

and loss of its function disrupts axonal extension and pathfinding. SMN is known to associate with the RNA-binding protein hnRNP-R, and together they are responsible for the transport and/or local translation of β-actin mRNA in the gro s. **However, the full complement of SMN-interacting proteins in neurons remains unknown**. Here we used mass spectrometry to identify HuD as a novel neuronal SMN-interacting partner. HuD is a neuron-specific RNA-binding protein that interacts with mRNAs, including candidate plasticity-related gene 15 (cpg15). We show that SMN and HuD form a complex in spinal motor axons, and that both interact with cpg15 mRNA in neurons. CPG15 is highly expressed in the developing ventral spinal cord and can promote motor axon branching and neuromuscular synapse formation, suggesting a crucial role in the development of motor axons and neuromuscular junctions. Cpg15 mRNA previously has been shown to localize into axonal processes. Here we show that SMN deficiency reduces cpg15 mRNA levels in neurons, and, more importantly, cpg15 overexpression partially rescues the SMN-deficiency phenotype in zebrafish. **Our results provide insight into the function of SMN protein in axons** a potential targets for the study of mechanisms that lead to the SMA pathology a muscular diseases.

open question

advance

xerox

Categories

open question

contrasting ideas

tendency

novelty

significance

surprise

background knowledge

summarizing

Xerox Incremental Parser

CONTRASTING IDEAS

… unorthodox view resolves … paradoxes …

In contrast with previous hypotheses ...

... inconsistent with past findings

OPEN QUESTION

… little is known …

… role … has been elusive

Current data is insufficient …

Applications

Use-case types:

- Text-mining
- Information retrieval
- Support for peer-reviewing
- Support for human annotation
- Visualisation of research literature

Domains:

- Bio-medicine
- Educational science

Genres:

- Research articles
- Project reports

Detecting "paradigm shifts"

Claimed Knowledge Updates*

interact with cpg15 mRNA in neurons. CPG15 is highly expressed in the developing ventral spinal cord and can promote motor axon branching and neuromuscular synapse formation, suggesting a crucial role in the development of motor axons and neuromuscular junctions. Cpg15 mRNA previously has been shown to localize into axonal processes. Here we show that SMN deficiency reduces cpg15 mRNA levels in neurons, and, more importantly, cpg15 overexpression partially rescues the SMN-deficiency phenotype in zebrafish. Our results provide insight into the function of SMN protein in axons and also identify potential targets for the study of mechanisms that lead to the SMA pathology and related neuromuscular diseases.

Keywords: neuritin, embryonic lethal abnormal vision Drosophila-like 4 (ELAV-L4), local protein synthesis

Other Sections▾

Spinal muscular atrophy (SMA) is a devastating genetic disease leading to infant mortality, due mainly to the loss of α-motor neurons of the spinal cord and brainstem nuclei. SMA occurs due to depletion of a ubiquitously expressed protein, SMN, which in all cells regulates RNA biogenesis and splicing through its role in the assembly of small nuclear ribonucleoprotein (snRNP) complexes (1). Despite the well-characterized association of SMN with the snRNP complex in both the nucleus and cytoplasm of motor neurons, in the axons SMN associates with mobile ribonucleoprotein (RNP) particles that are free of the core snRNP complex proteins (2). Thus, it is hypothesized that SMN may function in the assembly of axonal RNPs to regulate axonal mRNA transport and/or local protein synthesis (3, 4). Deficits in mRNA transport and local mRNA translation are associated with such neurologic disorders as fragile X syndrome and tuberous sclerosis (5, 6). Therefore, the interaction of SMN complex with other RNPs and their associated mRNAs within the axon may be crucial to understanding the pathophysiology of SMA.

- Nucleotide
- Protein
- PubMed
- Substance
- Taxonomy
- Taxonomy Tree

- *Review* The SMN complex. [Exp Cell Res. 2004]
- Multiprotein complexes of the survival of motor neuron protein SMN with Gemins traffic to neuror [J Neurosci. 2006]
- Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization [J Neurosci. 2003]
- Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and local [J Cell Biol. 2003]
- Dynamic association of the fragile X mental retardation protein as a messenger ribonucleoprote [Mol Biol Cell. 2008]

See more articles cited in this paragraph

▶SMN deficiency reduces cpg15 mRNA levels in neurons.

▶cpg15 overexpression partially rescues the SMN-deficiency phenotype in zebrafish

*Common work with Anita de Waard

xerox

Social sciences: search engine + peer-reviewing

highlighting knowledge-level claims

Summary of the research problem	Description of the research problem
The purpose of this article is to develop the idea that …	My interest of inquiry emerged in 1997 from a new idea …
The perspective I shall use in this essay relies heavily on the view …	… contrast with prior research results …

enhancing document search

reading support for quality judgment

EERQI

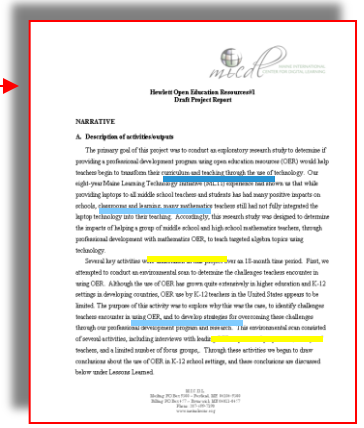
EUROPEAN EDUCATIONAL RESEARCH Quality Indicators

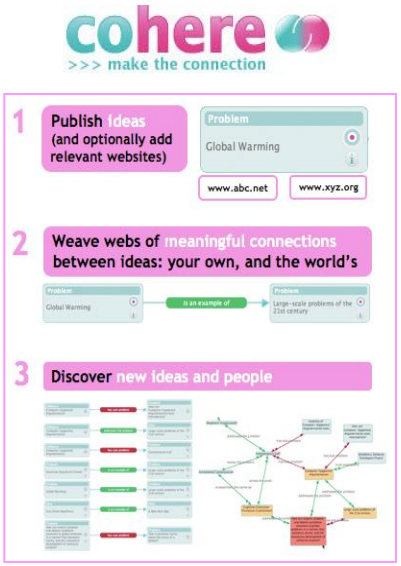
xerox

Human and machine annotation

template

XIP-annotated report





Human and machine annotation

Annotation of an analyst

The primary goal of this project was to conduct an exploratory research study to determine if providing a professional development program using open education resources (OER) would help teachers begin to transform their curriculum and teaching through the use of technology. Our eight-year Maine Learning Technology Initiative (MLTI) experience had shown us that while providing laptops to all middle school teachers and students has had many positive impacts on schools, classrooms and learning, many mathematics teachers still had not fully integrated the laptop technology into their teaching. Accordingly, this research study was designed to determine the impacts of helping a group of middle school and high school mathematics teachers, through professional development with mathematics OER, to teach targeted algebra topics using technology.

Comment [Pangiota 1]: Brief summary

Comment [Pangiota 2]: summary

Several key activities were undertaken in this project over an 18-month time period. First, we attempted to conduct an environmental scan to determine the challenges teachers encounter in using OER. Although the use of OER has grown quite extensively in higher education and K-12 settings in developing countries, OER use by K-12 teachers in the United States appears to be limited. The purpose of this activity was to explore why this was the case, to identify challenges teachers encounter in using OER, and to develop strategies for overcoming these challenges through our professional development program and research. This environmental scan consisted of several activities, including interviews with leading OER experts and proponents, surveys of teachers, and a limited number of focus groups. Through these activities we began to draw conclusions about the use of OER in K-12 school settings, and these conclusions are discussed below under Lessons Learned.

Comment [Pangiota 3]: good arguments, rigorous methodology

Comment [Pangiota 4]: multiple methods

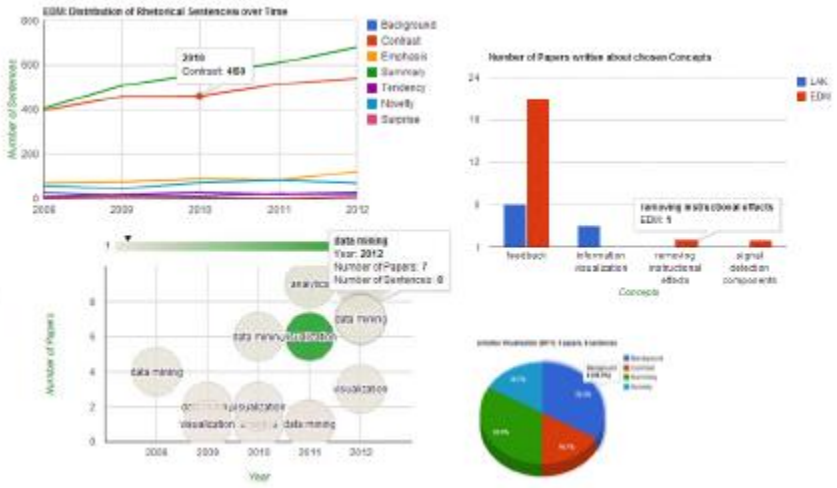
XIP

The primary goal of this project was to conduct an exploratory research study to determine if providing a professional development program using open education resources (OER) would help teachers begin to transform their curriculum and teaching through the use of technology. Our eight-year Maine Learning Technology Initiative (MLTI) experience had shown us that while providing laptops to all middle school teachers and students has had many positive impacts on schools, classrooms and learning, many mathematics teachers still had not fully integrated the laptop technology into their teaching. Accordingly, this research study was designed to determine the impacts of helping a group of middle school and high school mathematics teachers, through professional development with mathematics OER, to teach targeted algebra topics using technology.

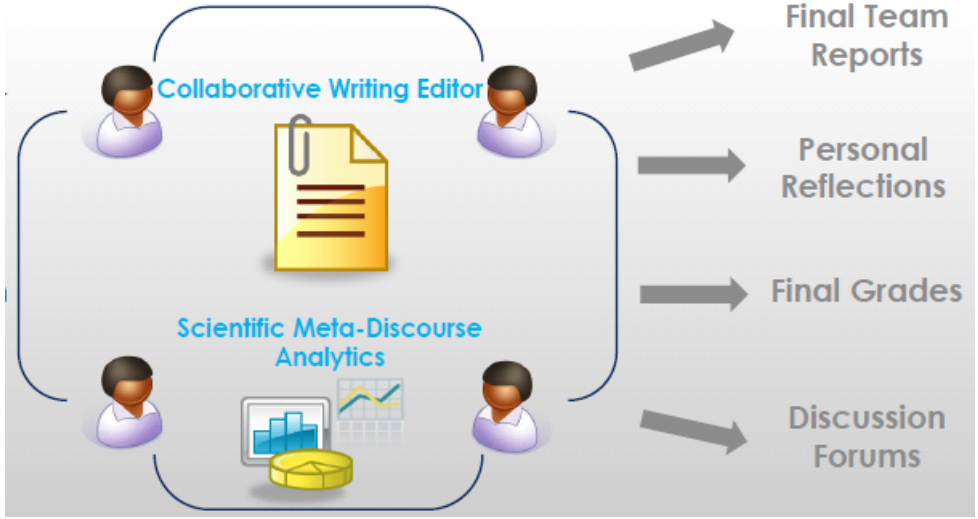
Several key activities were undertaken in this project over an 18-month time period. First, we attempted to conduct an environmental scan to determine the challenges teachers encounter in using OER. Although the use of OER has grown quite extensively in higher education and K-12 settings in developing countries, OER use by K-12 teachers in the United States appears to be limited. The purpose of this activity was to explore why this was the case, to identify challenges teachers encounter in using OER, and to develop strategies for overcoming these challenges through our professional development program and research. This environmental scan consisted of several activities, including interviews with leading OER experts and proponents, surveys of teachers, and a limited number of focus groups,. Through these activities we began to draw conclusions about the use of OER in K-12 school settings, and these conclusions are discussed below under Lessons Learned.

Collaborative Writing Editor
Student / Educator / Researcher

Extract from the PhD plan of Duygu Simsek



XIP dashboard: visual analytics of concepts as they appear in knowledge-level claims



KMi

xerox

Wrap-up

SCIENTIFIC FACTS

concerning
entities, relationships, correlations,
events etc.

xerox

Wrap-up

KNOWLEDGE-LEVEL CLAIMS

RHETORIC

contrasting ideas

open questions

findings

tendencies

major advances

SCIENTIFIC FACTS

concerning
entities, relationships, correlations, events etc.

xerox

A lot of plans

- detection of Claimed Knowledge Updates
- user interfaces for complementing human annotation
- rhetorical content + scientific facts
- educational applications
- web-service in Open Xerox: http://open.xerox.com/
- … open to any suggestion!

xerox

References

Lisacek, F., Chichester, C., Kaplan, A. & Sándor, Á. (2005). Discovering paradigm shift patterns in biomedical abstracts: application to neurodegenerative diseases. *First International Symposium on Semantic Mining in Biomedicine*, Cambridge, UK, April 11-13, 2005.

Sándor, Á., Kaplan, A. & Rondeau, G.. (2006). Discourse and citation analysis with concept-matching. *International Symposium: Discourse and document (ISDD)*, Caen, France, June 15-16, 2006.

Sándor, Á. (2006). Using the author's comments for knowledge discovery. *Semaine de la connaissance, Atelier texte et connaissance*, Nantes, June 29, 2006.

Sándor, Á. (2007). Modeling metadiscourse conveying the author's rhetorical strategy in biomedical research abstracts. *Revue Française de Linguistique Appliquée* 200(2), pp. 97--109.

De Waard, A., Buckingham Shum, S., Carusi, A., Park, J., Samwald, M., Sándor , Á. (2009). Hypotheses, Evidence and Relationships: The HypER Approach for Representing Scientific Knowledge Claims. *ISWC 2009, the 8th International Semantic Web Conference, Westfields Conference Center near Washington, DC., USA*, 25-29 October 2009.

Sándor, Á., Vorndran, A. (2009). Detecting key sentences for automatic assistance in peer reviewing research articles in educational sciences. In *Proceedings of the 2009 Workshop on Text and Citation Analysis for Scholarly Digital Libraries, ACL-IJCNLP 2009, Suntec, Singapore*, 7 August 2009 Singapore (2009), pp. 36--44.

xerox

References

Astrom, F., Sándor, Á. (2009). Models of Scholarly Communication and Citation Analysis. *ISSI 2009, 12th International Conference on Scientometrics and Informetrics*, Rio de Janeiro, Brazil, July 14-17, 2009

Sándor, Á., Vorndran, A. (2010). The detection of salient messages from social science research papers and its application in document search. Workshop Natural Language Processing in Social Sciences, May 10-14. Buenos Aires.

Sándor, Á., De Waard, A. (2012). Identifying Claimed Knowledge Updates in Biomedical Research Articles. In *Proceedings of the 2012 Workshop on Detecting Structure in Scholarly Discourse, ACL 2012, Jeju Island, Korea.*

De Liddo, A., Sándor, Á. and Buckingham Shum, S. (2012). Contested Collective Intelligence: rationale, technologies, and a human-machine annotation study. Computer Supported Cooperative Work (CSCW), 21(4-5), pp.

Simsek,D.,Buckingham Shum,S.,Sándor, Á.,de Liddo,A.,Ferguson,R. (2013). XIP Dashboard: visual analytics from automated rhetorical parsing of scientific metadiscourse. 3rd Conference on Learning Analytics and Knowledge, Leuven, Belgium, 8-12 April, 2013.

xerox

for your attention!