
Semantic Tagging of Scientific Articles

13th International Conference on Electronic Publishing

DC-Social Tagging community workshop
ELPUB 2009, Milan, Italy

June 10, 2009

Sudeshna Das ^{1,2,3}

¹Initiative in Innovative Computing, Harvard University

²MIND, Massachusetts General Hospital

³Harvard Medical School

Content

- What is semantic tagging?
- Enabling semantic tagging
- Demo
- Use of semantic tagging

What is semantic tagging?

- Tagging with a term or resource
 - belongs to a defined class
 - that has a URI (Uniform Resource Identifier)
 - controlled vocabularies or ontologies
- The tag has meaning!
 - Defined relationship between document and tag
 - Document “discusses” PhenomenonA
 - Document “majorTopic” Gene123
 - DocumentA “cites” DocumentB
- Provenance and status
- Tag - annotation

Enabling semantic annotation

- Semi-automatic text-mining
 - Use machines to “suggest terms”
 - Use editors to refine choice
- Currently mining for
 - Gene names
 - Gene Ontology terms

Annotation of articles



- Semi-automatic text-mining
 - Use machines to “suggest terms”
 - Use editors to refine choice
 - Currently mining for
 - Gene names
 - Gene Ontology terms
- ✓ High recall
- ✓ High precision

StemBook

www.stembook.org


[Login or Register](#)


StemBook





[Home](#) [About](#) [Contents](#) [Contributor Info](#) [Resources](#) [eAlerts](#)


STEMBOOK IS A COMPREHENSIVE, OPEN-ACCESS COLLECTION OF ORIGINAL, PEER-REVIEWED CHAPTERS COVERING TOPICS RELATED TO STEM CELL BIOLOGY. [Read More](#)


[Cellular and nuclear reprogramming](#)


[Genomics and proteomics](#)


[Renewal](#)


[Ectoderm specification and differentiation](#)


[Germ cell and somatic stem cell biology in reproduction](#)


[Stem cell immunology](#)


[Endoderm specification and differentiation](#)

[Mesoderm specification and differentiation](#)

[Therapeutic prospects](#)

[Epigenetics](#)

[Niche biology, homing, and migration](#)

[Tissue engineering](#)

News

Obama Ends Ban on Stem Cell Research

News, Washington Post 9 March 2009

President Obama lifted restrictions on funding for human embryonic stem cell research this morning and issued a presidential memorandum aimed at insulating scientific decisions across the federal government from political influence. To read more, [click here](#)

[Login or register](#) to post comments


Commentaries

July 31, 2008
[Random versus deterministic input into stem cell lineage choice](#)

June 26, 2008
[Stem cell maintenance factors represent potential therapeutic targets for cancer](#)


May 30, 2008
[Genomic approaches provide insights into the molecular basis of pluripotency](#)

[more](#)



The Science Collaboration Framework is a project of the Initiative in Innovative Computing at Harvard University in collaboration with the Harvard Stem Cell Institute, based on the Drupal open source content management system.

Copyright 2008-2009 by The President and Fellows of Harvard College. Drupal is copyright by the individual Drupal contributors. SCF and Drupal are both licensed under the GPL version 2 software license.

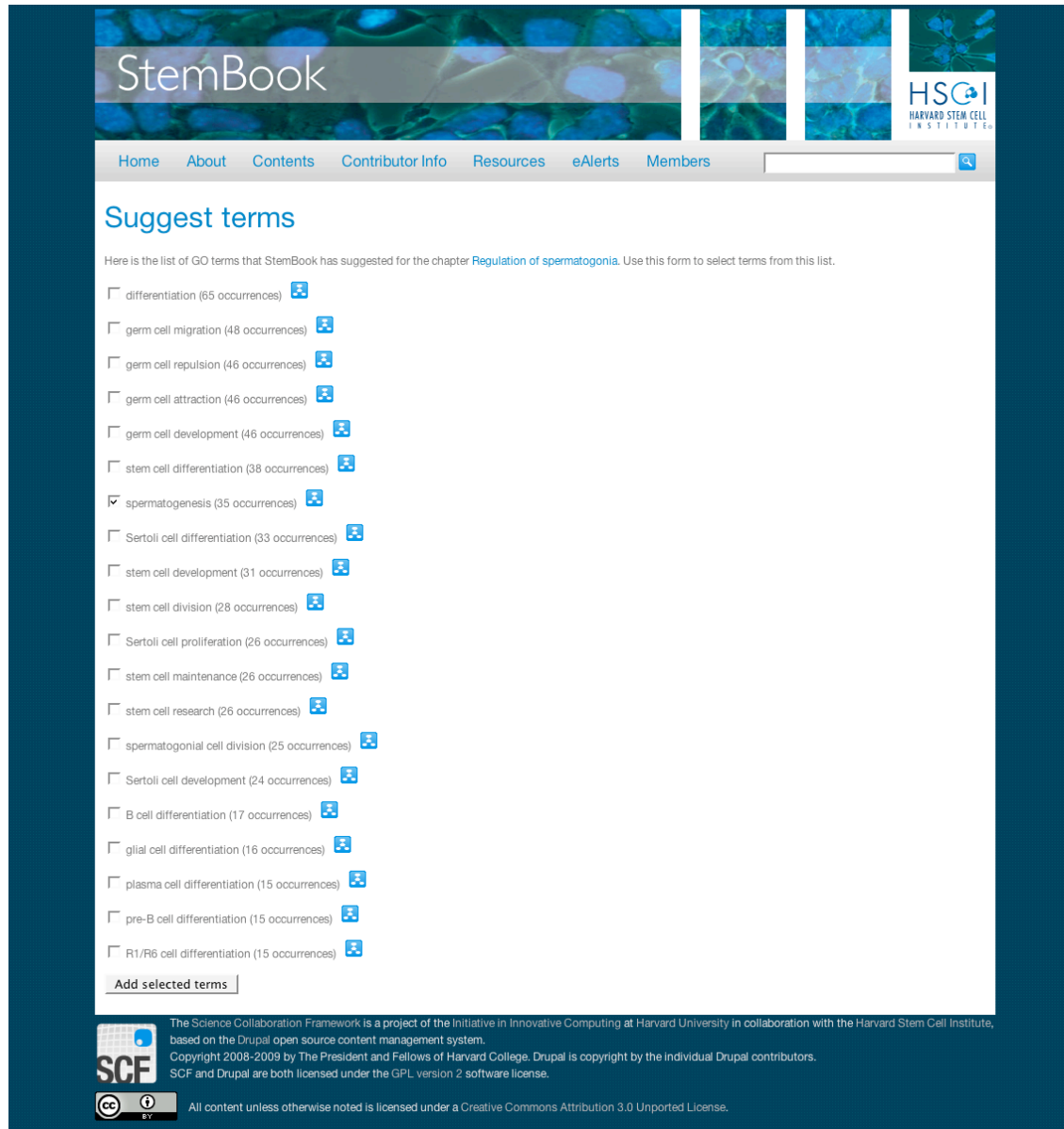


All content unless otherwise noted is licensed under a Creative Commons Attribution 3.0 Unported License.

StemBook

- Online, open access, review of stem cell biology
- Collaborative effort with HSCI
 - Lisa Girard PhD (editor)
- Editorial review
 - 14 Harvard-affiliated stem cell experts and 13 non-Harvard
- Launched in Sep 2008
 - 39 articles and 89 commissioned in 12 sections
 - 14557 unique visitors
 - ~150 visitors/day

Text mining - suggested terms



StemBook

Home About Contents Contributor Info Resources eAlerts Members


Suggest terms

Here is the list of GO terms that StemBook has suggested for the chapter [Regulation of spermatogonia](#). Use this form to select terms from this list.

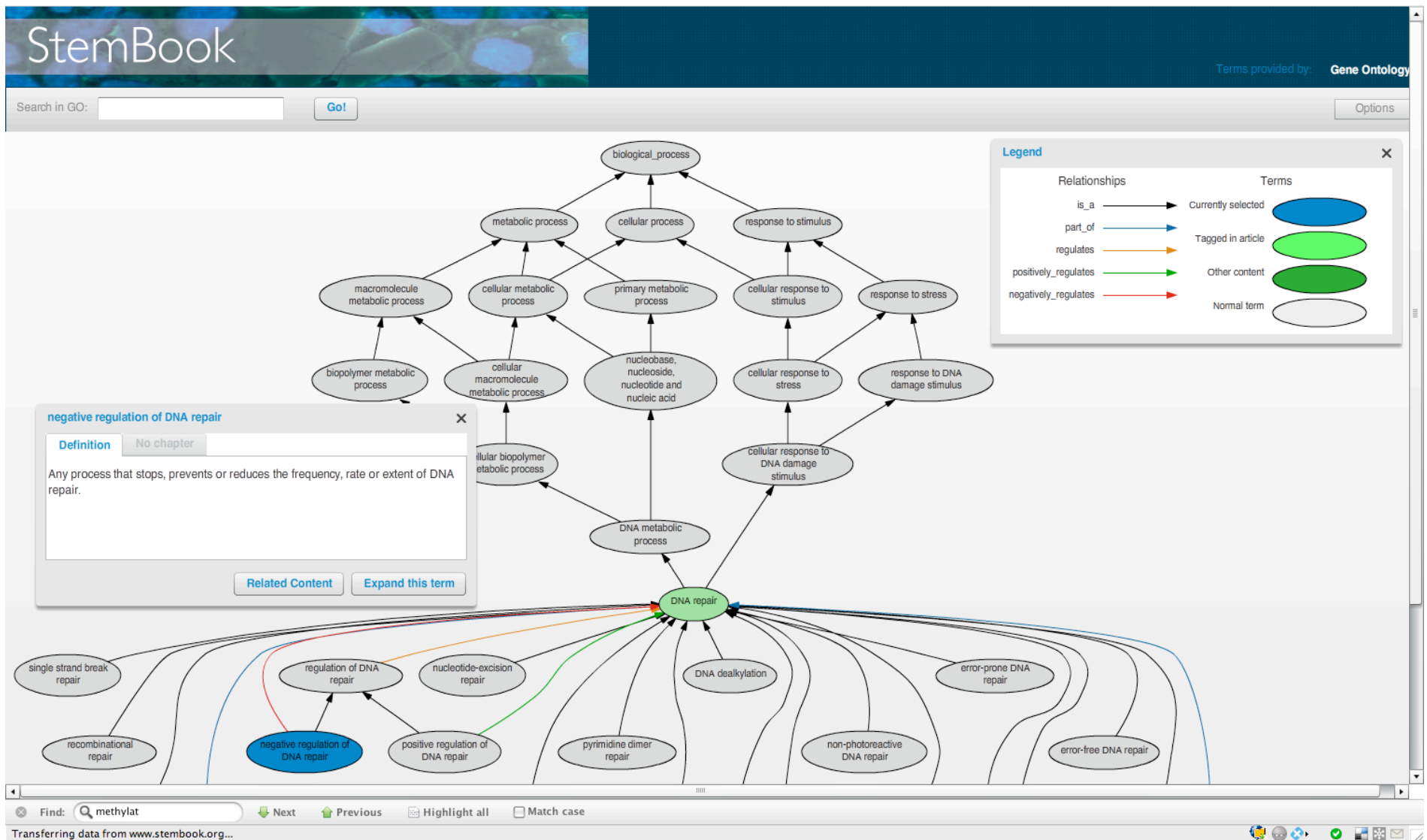
- ☐ differentiation (65 occurrences)
- ☐ germ cell migration (48 occurrences)
- ☐ germ cell repulsion (46 occurrences)
- ☐ germ cell attraction (46 occurrences)
- ☐ germ cell development (46 occurrences)
- ☐ stem cell differentiation (38 occurrences)
- ☒ spermatogenesis (35 occurrences)
- ☐ Sertoli cell differentiation (33 occurrences)
- ☐ stem cell development (31 occurrences)
- ☐ stem cell division (28 occurrences)
- ☐ Sertoli cell proliferation (26 occurrences)
- ☐ stem cell maintenance (26 occurrences)
- ☐ stem cell research (26 occurrences)
- ☐ spermatogonial cell division (25 occurrences)
- ☐ Sertoli cell development (24 occurrences)
- ☐ B cell differentiation (17 occurrences)
- ☐ glial cell differentiation (16 occurrences)
- ☐ plasma cell differentiation (15 occurrences)
- ☐ pre-B cell differentiation (15 occurrences)
- ☐ R1/R6 cell differentiation (15 occurrences)

[Add selected terms](#)

The Science Collaboration Framework is a project of the Initiative in Innovative Computing at Harvard University in collaboration with the Harvard Stem Cell Institute, based on the Drupal open source content management system.
Copyright 2008-2009 by The President and Fellows of Harvard College. Drupal is copyright by the individual Drupal contributors.
SCF and Drupal are both licensed under the GPL version 2 software license.

 All content unless otherwise noted is licensed under a Creative Commons Attribution 3.0 Unported License.

Editorial review



Semi-automatic Annotation

Neuroinflammation

By [Chris Scientist](#), Clinician, Washington University



The brain has generally been considered to be immune privileged and thus largely protected from immune factors. However, it is now recognized that the brain can initiate injury responses including neuroinflammation as a means to reduce injury and clean up injured brain tissue. The main cellular responders to brain injury are microglia, which produce a number of factors to regulate the injury response including cytokines and protective factors. Although neuroinflammation may have initial protective effects, prolonged and chronic

neuroinflammation may lead to prolonged neurodegeneration such as that seen in PD. Whether neuroinflammation acts as an initial trigger of PD pathogenesis or is a downstream result of another triggering pathogenic event is unclear. Neuroinflammation may lead to increased oxidative stress. Evidence for neuroinflammation in PD includes presence of activated microglia, increased expression of cytokines and pro-inflammatory signaling cascades (e.g., NF- κ B). Furthermore, epidemiological data supports a reduce risk of PD in users of anti-inflammatory drugs (e.g., NSAIDS).

References

Tansey MG, Frank-Cannon TC, McCoy MK, Lee JK, Martinez TN, McAlpine FE, Ruhn KA, Tran TA. **Neuroinflammation in Parkinson's disease: is there sufficient evidence for mechanism-based interventional therapy?** *Front Biosci.* 2008 Jan 1;13:709-17. [PubMed](#)

Wilms H, Zecca L, Rosenstiel P, Sievers J, Deuschl G, Lucius R. **Inflammation in Parkinson's diseases and other neurodegenerative diseases: cause and therapeutic implications.** *Curr Pharm Des.* 2007;13(18):1925-8. [PubMed](#)

Contributions


[Cannabinoids and neuroprotection in basal ganglia disorders](#)
J. Fernandez-Ruiz, Ciudad Universitaria, Madrid

6 May 2008

Semi-automatic Annotation

NF-κB

Neuroinflammation
By [Chris Scientist](#), Clinician, Washington University



The brain has generally been considered to be immune privileged and thus largely protected from immune factors. However, it is now recognized that the brain can initiate injury responses textmining: genes as a means to reduce injury and clean up injured brain tissue. NF-κB factors to brain injury are microglia, which produce a number of factors to regulate the injury response including cytokines and protective factors. Although neuroinflammation may have initial protective effects, prolonged and chronic neuroinflammation may lead to prolonged neurodegeneration such as that seen in PD. Whether neuroinflammation acts as an initial trigger of PD pathogenesis or is a downstream result of another triggering pathogenic event is unclear. Neuroinflammation may lead to increased oxidative stress. Evidence for neuroinflammation in PD includes presence of activated microglia, increased expression of cytokines and pro-inflammatory signaling cascades (e.g. NF-κB). Furthermore, epidemiological data supports a reduce risk of PD in users of anti-inflammatory drugs (e.g., NSAIDS).

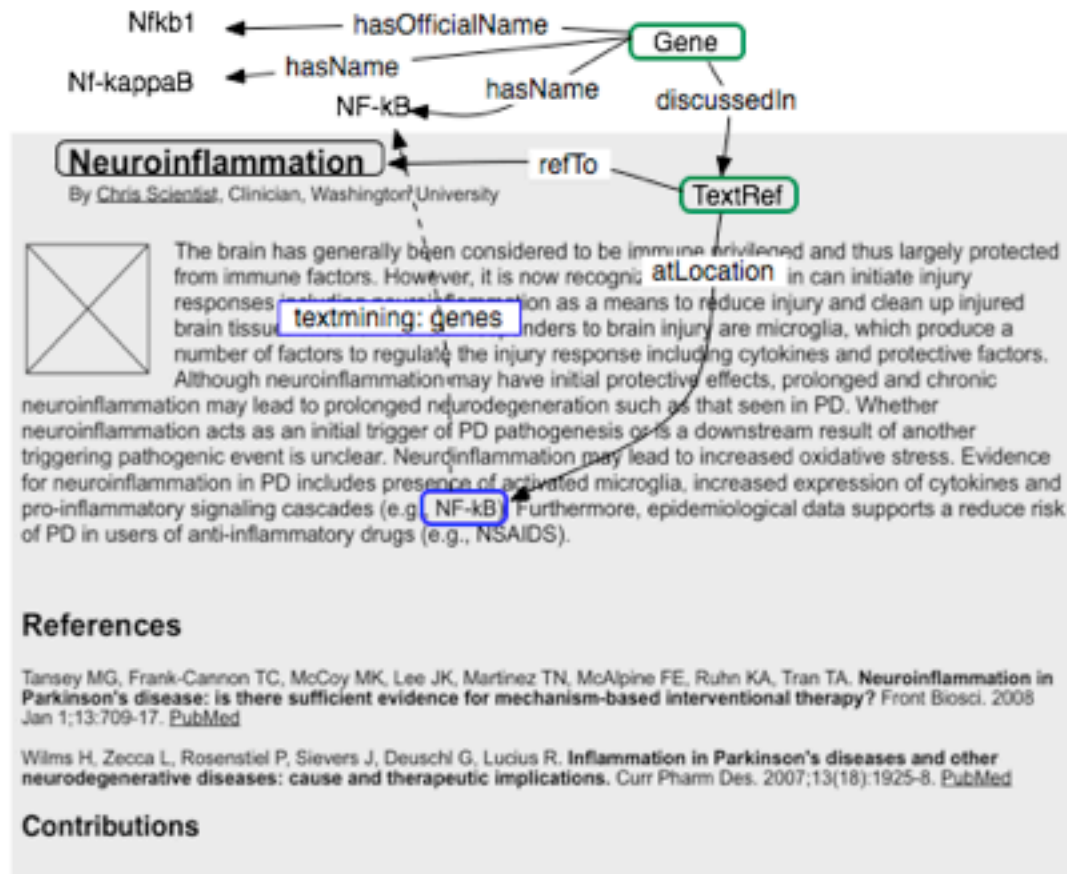
References

Tansey MG, Frank-Cannon TC, McCoy MK, Lee JK, Martinez TN, McAlpine FE, Ruhn KA, Tran TA. **Neuroinflammation in Parkinson's disease: is there sufficient evidence for mechanism-based interventional therapy?** *Front Biosci.* 2008 Jan 1;13:709-17. [PubMed](#)

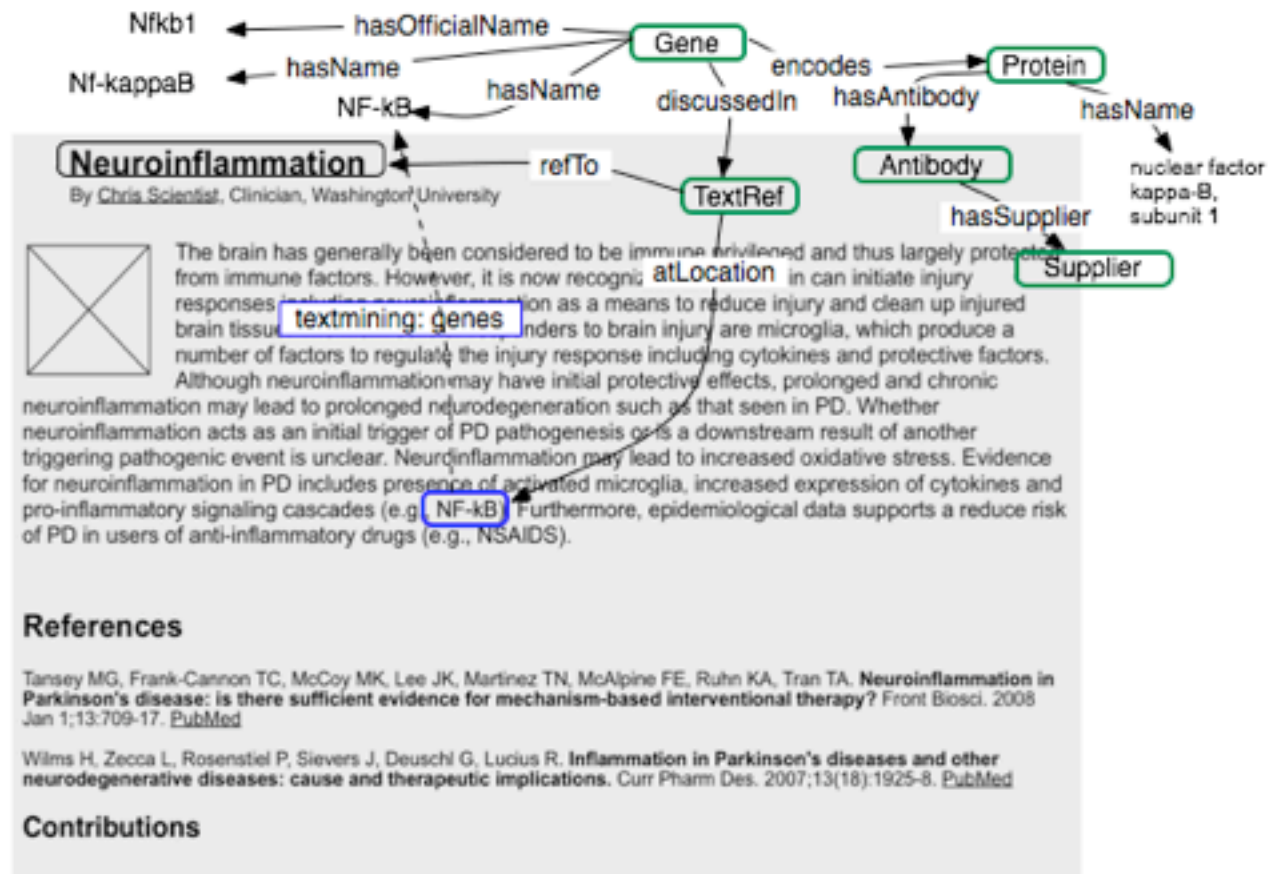
Wilms H, Zecca L, Rosenstiel P, Sievers J, Deuschl G, Lucius R. **Inflammation in Parkinson's diseases and other neurodegenerative diseases: cause and therapeutic implications.** *Curr Pharm Des.* 2007;13(18):1925-8. [PubMed](#)

Contributions

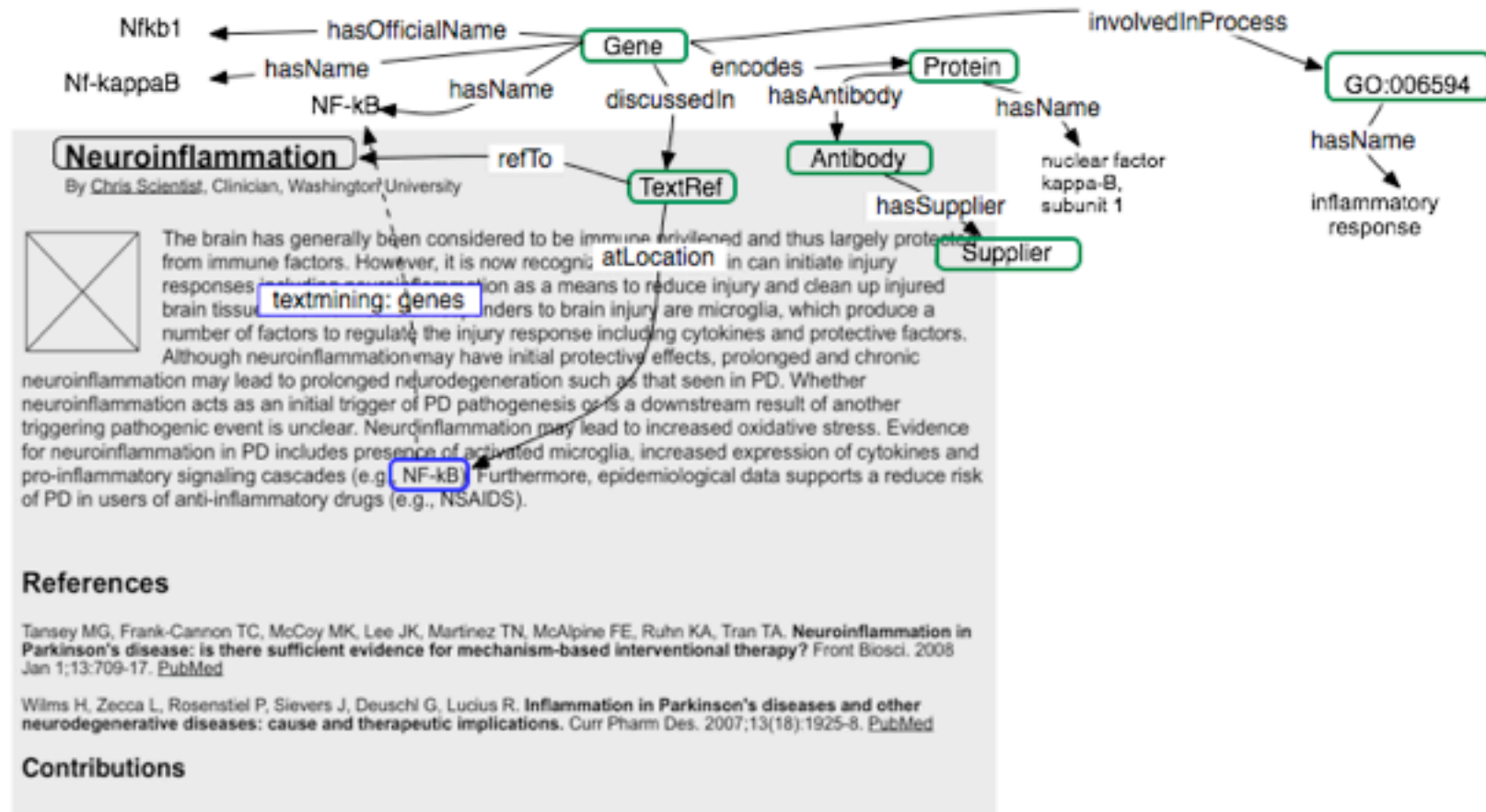
Semi-automatic Annotation



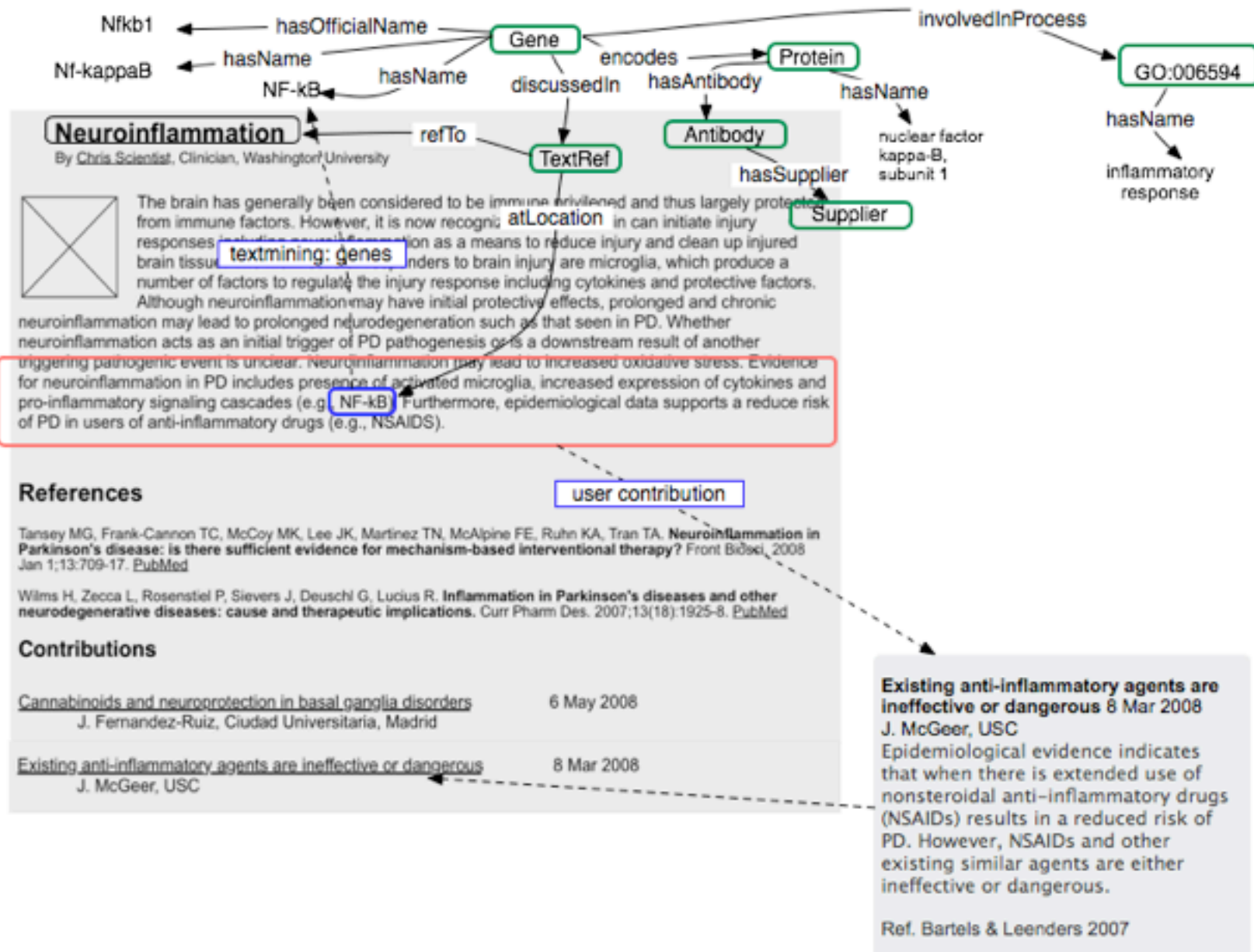
Semi-automatic Annotation



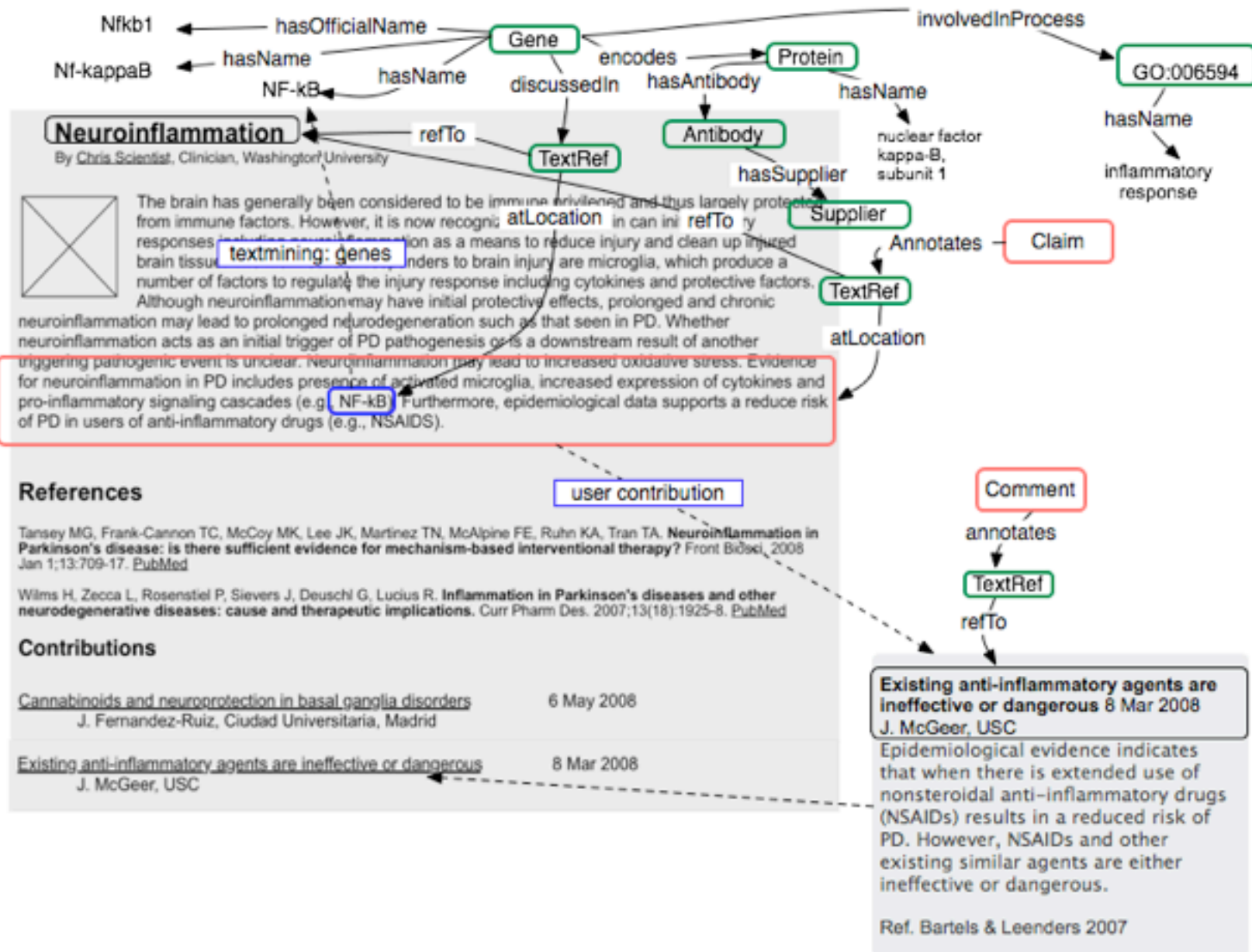
Semi-automatic Annotation



Semi-automatic Annotation

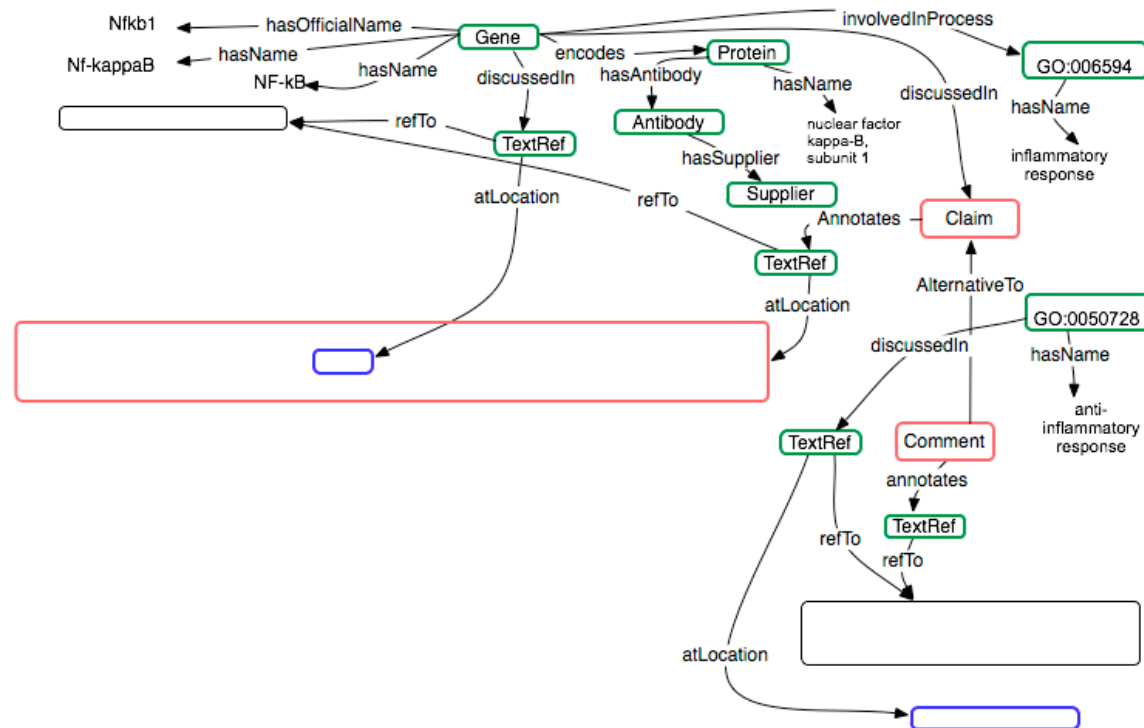


Semi-automatic Annotation

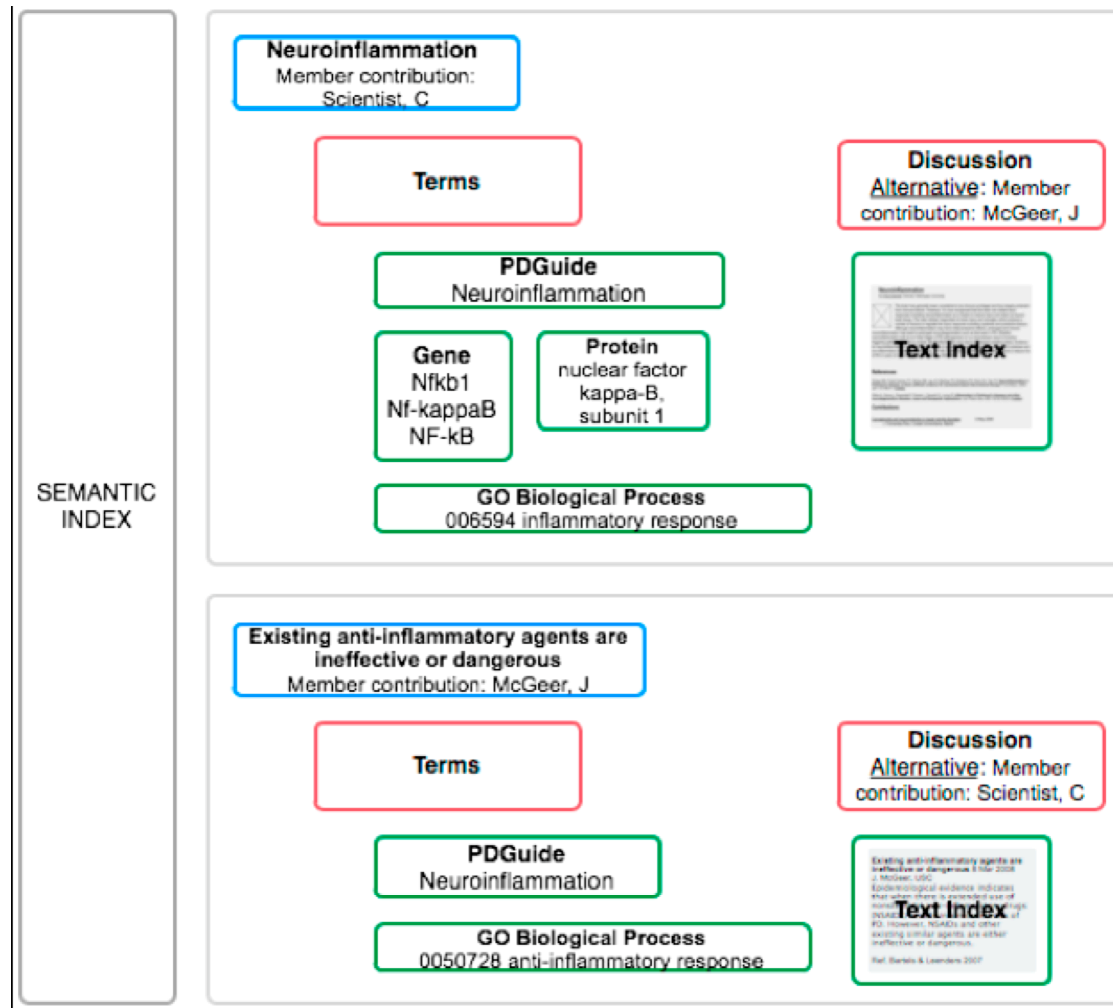


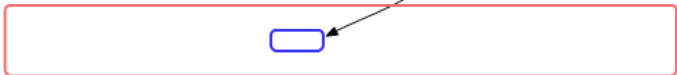
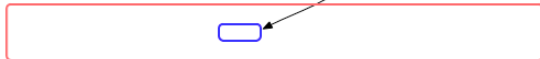
Entity	hasOfficialName	involvedInProcess
Nfkb1	1	





Build Semantic repository





Summary

- Reusable framework for building communities
- Easy to adopt
- Leverages linked data
- Uses shared ontologies/vocabularies
- Interoperable with other knowledge-bases

Team

IIC

Software Engineer

Mark Goetz

Stéphane Corlosquet

Tom Green

Alister Lewis-Bowen

Louis Weitzman

Project Management

Sudeshna Das

Tim Clark

Agaric Design

Ben Melancon

HSCI

Lisa Girard

Brock Reeve

NeuroCommons

Alan Ruttenberg

Jonathan Rees

Alzforum

June Kinoshita

Elizabeth Wu