



TDM: How publishers are supporting researchers

Text and Data Mining

PERCEPTIONS AND OPPORTUNITIES



AWARENESS OF TEXT MINING IS STILL LOW

PERCENTAGE OF RESPONDENTS THAT USED TDM:



INTEREST IN TEXT MINING IS GROWING

2/3 KEEN TO USE IN THEIR OWN RESEARCH

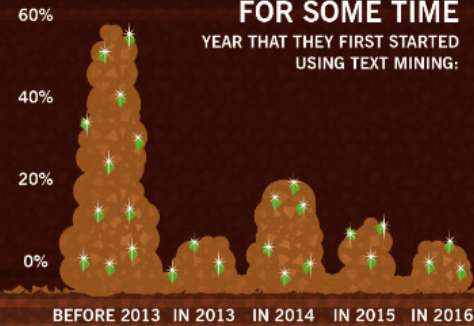
BIGGEST TEXT MINING BENEFITS

AUTOMATING SYSTEMATIC LITERATURE REVIEW
EXTRACTING INFORMATION, CONCEPTS AND NEW FACTS



EXPERIENCED MINERS HAVE BEEN USING TEXT MINING FOR SOME TIME

YEAR THAT THEY FIRST STARTED USING TEXT MINING:



INFRASTRUCTURE SUPPORTS CURRENT LEVELS OF TDM

RATINGS ON RESOURCES AND SUPPORT AVAILABLE FOR TEXT MINING



CHALLENGES FOR MINERS!



POTENTIAL FOR INCREASED TDM IN THE FUTURE



‘How to’ videos

- What is text mining?

<https://www.youtube.com/watch?v=l3cjbB38Z4A>

- How does text mining work?

<https://www.youtube.com/watch?v=xxqrIZyKKuk&feature=youtu.be>

How do publishers help?

- Raise awareness of the potential of TDM
- Provide access
- Optimize published content for text mining
- Invest in APIs to deliver optimal experiences to both humans and machines
- Provide technical support
- Collaborate with other publishers and academic institutions
- Support the development of specialized tools for STM content
- Employ NLP specialists

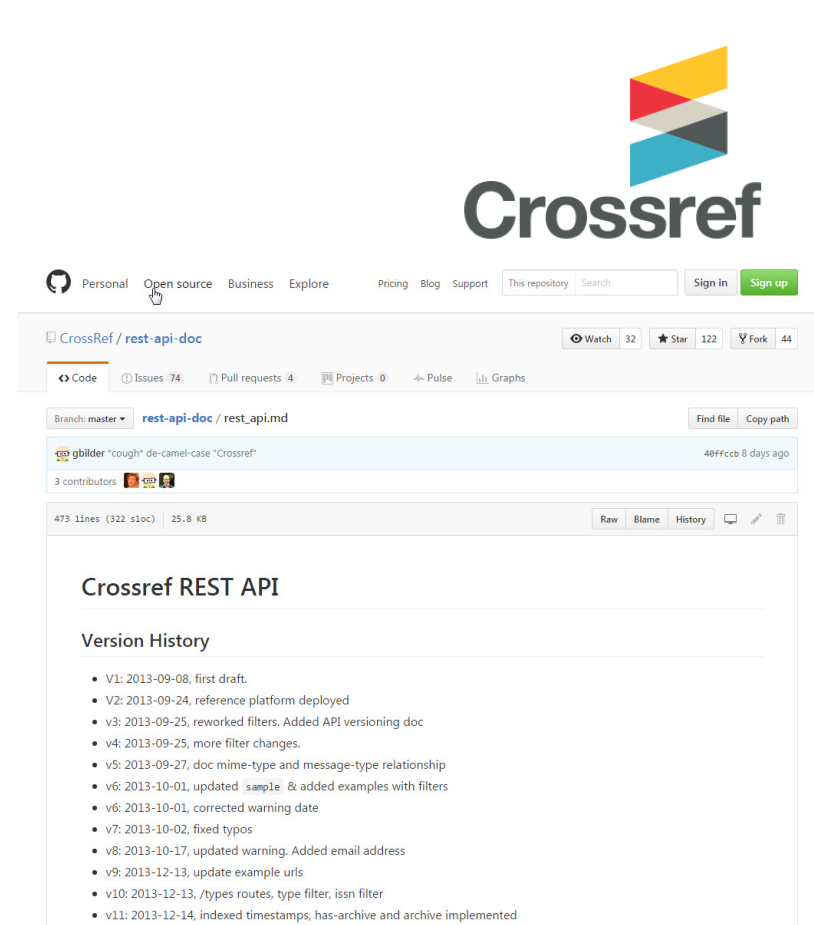


Supporting cross-publisher TDM

Providing one point of access for researchers to TDM all published STM content

- Free public service
- Common API (protocol) for requesting machine readable full text from many different publishers
- ‘One stop shop’

<http://tdmsupport.crossref.org/>



The screenshot shows the GitHub repository page for Crossref's REST API. The repository is named 'rest-api-doc' and is located under the 'CrossRef' organization. The page displays the repository's name, a search bar, and navigation links for 'Personal', 'Open source', 'Business', 'Explore', 'Pricing', 'Blog', and 'Support'. The repository has 32 watchers, 122 stars, and 44 forks. The current branch is 'master', and the file path is 'rest-api-doc / rest_api.md'. The commit message is 'gbuilder "cough" de-camel-case "Crossref"'. The file size is 25.8 KB. The page content includes the title 'Crossref REST API' and a 'Version History' section with a list of updates from 2013-09-08 to 2013-12-14.

Crossref REST API

Version History

- V1: 2013-09-08, first draft.
- V2: 2013-09-24, reference platform deployed
- v3: 2013-09-25, reworked filters. Added API versioning doc
- v4: 2013-09-25, more filter changes.
- v5: 2013-09-27, doc mime-type and message-type relationship
- v6: 2013-10-01, updated `sample` & added examples with filters
- v6: 2013-10-01, corrected warning date
- v7: 2013-10-02, fixed typos
- v8: 2013-10-17, updated warning. Added email address
- v9: 2013-12-13, update example urls
- v10: 2013-12-13, /types routes, type filter, issn filter
- v11: 2013-12-14, indexed timestamps, has-archive and archive implemented

Thank you!

Gemma Hersh

Policy and Communications Director, Elsevier

g.hersh@elsevier.com

Twitter: gemmahersh