![CCC Copyright Clearance Center | RightsDirect]

How it works

Using XML for mining

Identify full-text articles.

Automatically retrieve and index full-text XML articles.

Query the index using Linguamatics I2E.

Integrating RightFind XML for Mining and Linguamatics I2E

Get instant access to full-text articles — Automatically index in I2E

While text mining data from scientific article abstracts provides some value, there are limits to what data can be found. The ability to mine the full text of the article ensures that you don't miss vital data, discoveries and assertions that are only published in the full-text. For detailed descriptions of methods and protocols and the complete information on all study results, full-text is an essential resource.

However, obtaining full-text articles for text mining is often a struggle. When you do get access to the full-text, you must contend with multiple formats and inconsistent license terms — all of which can inhibit your text mining efforts.

To address these issues, Copyright Clearance Center (CCC), in partnership with Linguamatics, developed an integrated solution to make it simple for I2E end users to obtain and index full-text XML articles from a wide range of scientific publishers.

I2E, Linguamatics' powerful, market-leading Natural Language Processing (NLP) text analytics platform, enables you to answer business-critical questions by rapidly extracting relevant facts and relationships from large document collections. CCC's XML for Mining solution makes it easy to quickly identify and automatically index full-text article collections from multiple publishers for querying in I2E.

Combining XML for Mining with Linguamatics I2E gives you instant access to full-text content, enabling you to make discoveries and connections that can only be found in full-text articles. The solution provides faster speedto-actionable-insight, reduces costs and mitigates your organization's copyright infringement risk.

The integration of XML for Mining with Linguamatics 12E enables you to:

- Automatically index full-text content in XML format in I2E from CCC's database of millions of articles from major STM publishers

- Discover and purchase XML-formatted articles outside of your company's subscriptions for mining, and index them in I2E

- Store, manage and analyze XML-formatted article content for all your literature-based text mining projects

Benefits

- **Don't miss vital data and assertions.** The combination of XML for Mining with Linguamatics I2E enables you to make discoveries and connections that can only be found in full-text. Now you can obtain and mine XML-formatted content from publications you subscribe to and discover articles that fall outside of company subscriptions, giving you the most complete article collection for mining.

- **Build the most comprehensive collection of full-text articles in XML format for mining.** Get the most comprehensive coverage of full-text article content and text mining rights available on the market today.

- **Save time and money.** Spend more time focusing on analysis and discovery and less time wrangling data sources. XML for Mining provides aggregated full-text article content and normalized metadata from a wide range of publishers in a single collection. Integrating XML for Mining with I2E reduces the time and costs associated with article conversions, content management, formatting, I2E index creation and individual negotiations with publishers.

- **Manage copyright compliance.** Because all of the content available in XML for Mining is pre-authorized for commercial text mining, you get the peace of mind that your text mining projects comply with copyright, minimizing your organization's infringement risk.

Linguamatics

Linguamatics, an IQVIA company, delivers market-leading NLP-based AI solutions for high-value knowledge discovery and decision support from text. It empowers customers to speed up drug development and improve patient outcomes by breaking down data silos, boosting innovation, enhancing quality, and reducing risk and complexity. Linguamatics NLP platform is used by top commercial, academic and government organizations, including 19 of the top 20 global pharmaceutical companies, the US Food and Drug Administration (FDA) and leading US healthcare organizations. Linguamatics NLP platform can be used to mine a wide variety of text resources, such as scientific literature, patents, Electronic Health Records (EHRs), clinical trials data, news feeds, social media and proprietary content. Linguamatics is an integration and referral partner for CCC's RightFind XML for Mining solution.

CCC | RightsDirect

Get more Information

Learn how combining Linguamatics I2E and XML for Mining can improve the results of your text and data mining queries, reduce costs and mitigate your infringement risk.

U.S. organizations:
- ⊕ copyright.com/LinguamaticsI2E
- ✉ solutions@copyright.com

Outside U.S. organizations:
- ⊕ rightsdirect.com/LinguamaticsI2E
- ✉ solutions@rightsdirect.com

About CCC
A pioneer in voluntary collective licensing, CCC (Copyright Clearance Center) helps organizations integrate, access, and share information through licensing, content, software, and professional services. With expertise in copyright and information management, CCC and its subsidiary RightsDirect collaborate with stakeholders to design and deliver innovative information solutions that power decision-making by helping people integrate and navigate data sources and content assets.