

CONTENT INSIGHT

Total: The total number of paragraphs analyzed by Harmonizer.

Unique: The number of paragraphs with no other matching paragraphs.

Exact: The number of paragraphs with one or more exact matches elsewhere in the content.

Close: The number of paragraphs that are close but not exact matches to other paragraphs in the content.

Match: The total number of paragraphs that have one or more matches.

Needed: All matches are divided into Match Groups. When moving to a reusable content format, it is possible to consolidate content to a single, reusable paragraph. "Needed" is the number of paragraphs that must be retained in a reuse model.

Potentially Redundant: The maximum number of paragraphs that could be eliminated.

Matching Sequences: Sequence matches can be useful in identifying areas of your content where a more complex content model might be useful, such as a *Topic* in the DITA content model.

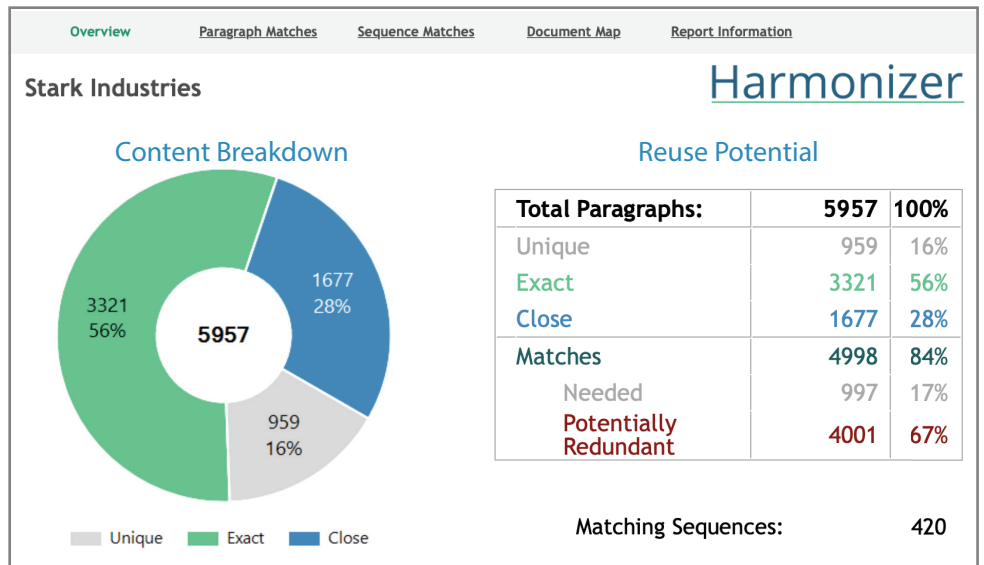
Identify content reuse across multiple document sets and source formats

INTELLIGENT CONTENT REUSE ANALYSIS

Harmonizer is a software application that processes and analyzes thousands of pages using natural language processing to identify redundant content in document collections. Harmonizer examines content across a variety of source formats. HTML and Excel reports from Harmonizer help operationalize information and influence reuse strategy when moving to XML.

SAVE TIME AND MONEY

- ✓ Identify redundant content and variants for clean-up efforts
- ✓ Analyze multiple source formats in one report (HTML, XML, Word, Framemaker, SGML, and more)
- ✓ Create leaner content collections to streamline management, translation, and localization
- ✓ Standardize text and editorial style across large content sets



Contact us today to get started!
+1.800.321.2816

61-18 190th Street
 Suite 205
 Fresh Meadows, NY 11365
 +1 718.357.8700
 info@dclab.com
 www.dclab.com