Curating Medical Writing Training Content *with Crawl Technology*

SYNTEREX
By Jeanette Towles, MA, RAC-Drugs
https://synterex.com

Table 1: Discovery & Curation of Crawl Results

	Number of Topics
Initial suggested topics (first 2 days)[a]	34
Rejected topics	**3**
Inactive client	3
Total accepted topics	**31**
Number of published topics (AI-suggested topics with curation needed)[a]	2
Disease-specific information	1
Regulation- or guidance-specific information	0
Client-specific information[b]	0
Other[c]	1
Number of confirmed topics (AI-suggested topics with no curation needed)[a]	29
Disease-specific information	10
Regulation- or guidance-specific information	3
Client-specific information[b]	16
Other	0
Subsequent suggested topics (next 29 days)[a]	**286**
Number of rejected topics	**0**
Total accepted topics	**292**
Number of published topics (AI-suggested topics with curation needed)[a]	286
Disease-specific information	48
Regulation- or guidance-specific information	9
Client-specific information[b]	192
Other[d]	37
Number of confirmed topics (AI-suggested topics with no curation needed)[a]	6
Disease-specific information	2
Regulation- or guidance-specific information	2
Client-specific information[b]	1
Other[e]	1
Additional topics discovered (year to date)[f]	**246**
Total topics discovered	*566*

a. Collected November through December 2021.
b. Includes client name, project code, or document name.
c. IT software (Workspace ONE).
d. IT software (e.g., SalesForce, PleaseReview, Teams), n=17; medical writing document types (e.g., clinical study report, protocol), n=17; process (style guide, SOP), n=2; community organization, n=1.
e. Process (quality control).
f. Suggested January 2021 through August 2022.

Background

An issue that managers often face when onboarding and coordinating resources is having an abundance of data and information and not knowing how to collate or transfer it in an effective way. A "crawl" is the process of using a script or program to collect all content from one or more web pages into a dataset. We will outline the journey of implementing and optimizing crawl technology that leverages artificial intelligence (AI) to curate content and enhance the digital employee experience.

Methods

In Q4 2021, our staff was engaged in discussions about streamlining the employee onboarding experience. We defined what types of information represent the most common inquiries from new employees, in particular entry-level medical writers. The top categories were disease-specific information, regulation- or document-specific information, and client-specific information.

Around the same time, we became aware of the Viva Topics software through our regularly scheduled review of new technologies from the annual Microsoft Ignite conference in 2021. Viva Topics is a Microsoft technology (Redmond, WA) that leverages AI, Office 365 search capabilities, Microsoft's application programming interface, and various core services (e.g., SharePoint, OneDrive) to put data together and enhance the discovery experience. We selected Viva Topics for these abilities in particular:

- Documenting who is the best resource for a specific therapeutic area
- Having a concise glimpse into a client and their portfolio
- Identifying subject matter experts (SMEs) across technologies, applications, and clients or departments
- Learning more about what topics matter to users

In addition, we liked that in Viva Topics the SME could further update the information in the case of inaccurate results and maintain the information over time, as well as the relative cost-effectiveness of the tool (around $50 USD per user per year). A topic is defined as any concept that is considered organizationally important. The breadth of the crawl is customizable by the software administrators; when deploying Viva Topics, the administrators selected for it to crawl all of the company's SharePoint sites.

Results

Topic Discovery and Curation

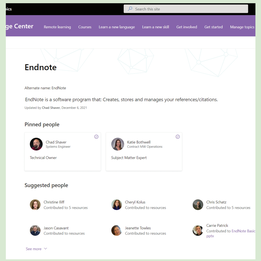
We assigned a "Knowledge Manager" on our Medical Writing Operations team and assigned them permissions to work with the Information Technology (IT) lead and other SMEs to further curate topics, including the following tasks:

1. Confirm AI-generated topics ("suggested topics") as valid
2. Reject topics to prevent the content from being viewed as a topic
3. Create topics that were not discovered by AI or identify topics that could be edited to be more helpful or accurate.

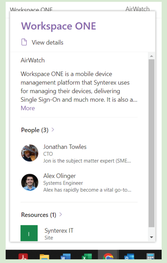
The initial topics AI discovered were mostly client- and disease-specific information. The majority of these terms were confirmed to be topics of interest for the company; only 3 terms were rejected **(Table 1)**.

Once these initial terms were confirmed, the AI continued to discover new topics over the course of the month. None of these topics was rejected, and most of them required curation (manual edits to improve the quality of the result, identify the proper SMEs, add detailed descriptions, or link additional files or sites).

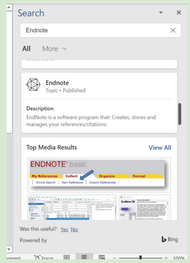
Figure 1: User Interaction with Topics



Topics in Search Results



Topics Highlighted on SharePoint Pages



Searching Inside of Office Apps

User Interaction with Topics

When a topic is created, there are 3 main ways in which the user can interact with the topic **(Figure 1)**.

From these 3 points of origin, users can select the topic to learn more about it through the topic details and can even connect with the SMEs directly via chat. Content permissions on sites are set based on the user group in which the user is listed in Microsoft Office, so while the user may be able to see a term, they may have to obtain access to it to interact with it further on a SharePoint site or inside of an Office application.

Conclusions

Viva Topics is a relatively inexpensive product that makes crawl technology that leverages AI accessible to a wider variety of users than has been possible previously. To maximize the investment in crawl technology, plan to have a Knowledge Manager who has broad knowledge across the organization, to understand how the topics interrelate and curate the topics. Collated topics may need minor maintenance over time to ensure they stay up to date with current regulations and staff; the silver lining is that the initial exercise of refining the topics will be time well spent, as AI will continue to provide suggestions to improve topics as changes occur in your environment. Users can also interact with the information and note whether the description was useful, and AI uses this information to report to administrators and refine descriptions and topic highlighting.

Resources

Microsoft Viva Topics: https://docs.microsoft.com/en-us/viva/topics/topic-experiences-overview

Acknowledgment

Thank you to Jonathan Towles for technical peer review of this poster.