

UKSG webinar Q&A

‘AI: Empowering Libraries & Making It Real’

Manisha Bolina | Ken Chad | Ben McLeish

Questions that came during the presentation:

“Q: Are there demo sites of Yewno, Dimensions we can play with? I'd be interested to see how they work on humanities subjects, not just science?”

- **Ben:** Yes: app.dimensions.ai is free for researcher use. Publications and datasets are included in this with a limited number of facets. The remainder is part of the licensed version.
- **Manisha:** Yewno is multi-disciplinary here are some examples which are not STEM
[Research into French Enlightenment with Yewno Discover](#)
[Yewno Discover - Literary Research "Uncle Tom's Cabin"](#)
[Yewno Discover - Business/Economics help for returning students](#)

“Q: How do you deal with accessibility when user can't use an interface like Yewno with lines and squares vs. circles etc.?”

- **Manisha:** we have a VPAT document which covers Accessibility, please contact me directly for more information.

“Q: Do you have examples of AI being applied in digital archival collections?”

- **Ken:** I don't. Unfortunately, I missed this webinar last month: Artificial Intelligence and Archives - AUGUST 4, 2020 CLIR blog
<https://www.clir.org/2020/08/artificial-intelligence-and-archives/>

A good contact might be Adrian Stevenson Senior Technical Coordinator at Jisc. Adrian is the technical lead for the UK Archives Hub service and a member of Jisc Artificial Intelligence and Cloud Centres of Excellence teams. Twitter: @adrianstevenson Email: adrian.stevenson.@jisc.ac.uk Tel: +44 (0) 161 413 7561
- **Manisha:** Yes, I do, please contact me directly and I would be happy to go through how Yewno has dealt with content from Digital Archives.

“Q: Ken, could you please point some studies that show that Library use is a good predictor of student success? Would be good for advocacy”

- **Ken:** I quoted the following in [The new role of the library in teaching and learning outcomes](#). By Ken Chad & Helen Anderson. HELIbTech briefing paper No. 3. 20 June 2017
Australasian Survey of Student Engagement [AUSSE] 2012 Institution Report (p. 40-41)
Library Use and Undergraduate Student Outcomes: New Evidence for Students; Retention and Academic Success. Soria, Krista M., Jan Fransen, and Shane Nackerud. Portal: Libraries and the Academy. Vol 13, No. 2 (2013), pp. 147-164. https://www.press.jhu.edu/journals/portal_libraries_and_the_academy/portal_pre_print/current/articles/13.2soria.pdf

Also, of value is ‘Library Impact Data Project’ Graham Stone, Dave Pattern & Bryony Ramsden. Sconul Focus 2012 .
https://www.sconul.ac.uk/sites/default/files/documents/8_0.pdf “The project has successfully demonstrated that there is a statistically significant relationship between student attainment and two of the indicators: e-resources use and book-borrowing. This relationship has been shown to be true across all eight partners in the project that provided data for these indicators”.

The Jisc ‘Library analytics and metrics project’ (LAMP) from 2011 did some interesting work <https://www.jisc.ac.uk/rd/projects/library-analytics-and-metrics-project>. The project has ended and the work is now part of the Jisc Learning Analytics service <https://www.jisc.ac.uk/learning-analytics>

You will find more references on libraries and student success via Google.

“Q: Does the software process all the information resources such as newspaper, articles etc together, so the network graph consist of all the resources together?”

- **Manisha:** Yes, we have literally just included News. The main Knowledge Base in Yewno is academic content. We added news because people wanted to search things like “George Floyd”, but they can’t do that in academic content as no one has published anything about “George Floyd”. If you would like a list of News sources, we are using please contact me directly

“Q: About “Plutchik”: artificial intelligence chatbot for searching NCBI databases. How does it work and which technology uses?”

- **Manisha:** I’m sorry I can’t answer this question.

“Q: As far as I know, algorithm like topic modelling (be it unsupervised or supervised) is conducted on particular resources like journal articles and for the once we have some idea about the collection. So, does the knowledge graph compile topics from all the resources together or focused on particular resource?”

- **Manisha:** Yewno has created a its own topic model which consist of 33 topics, 800 subtopics and 5million concepts. The topic model is used to cluster project concepts/documents together. This is how we then understand the semantic relationship between documents.

“Q: How well can researchers' ambiguation be done using AI and are researchers' names standardized? This is for us one of the major issues performing research analytics.”

- **Ben:** Author disambiguation remains a challenge across all known platforms. AI certainly helps, but it's always going to be a mix of hand-curated data, ORCID and similar methods for researcher identification, and AI.
- **Manisha:** I agree with Ben, I would defer to ORCID too.

“Q: How does an AI team look like in a library? Librarians, IT? What are their profiles?”

- **Ben:** This varies strongly across institutions. Presently the presence of a data scientist (or team of them) is usually the telltale sign of AI use. Most libraries do not have a working AI team.
- **Manisha:** my answer is the same as Ben's. I deal with anyone from the Library Director to the Systems people to Collection management to those that sit in the Research team or the department.

“Q: Is there any info which is excluded in Yewno/Dimensions searches, and if so who decides what that is?”

- **Ben:** Dimensions - If results sit below a threshold of confidence for relevance for a search they are excluded or rather they are dropped down the list. This threshold relies on the AI. It is similar to the level of confidence Google uses to include or exclude results. It is not driven by human decisions.
- **Manisha:** We deal with academic content, so you will not find Fiction content in Yewno.

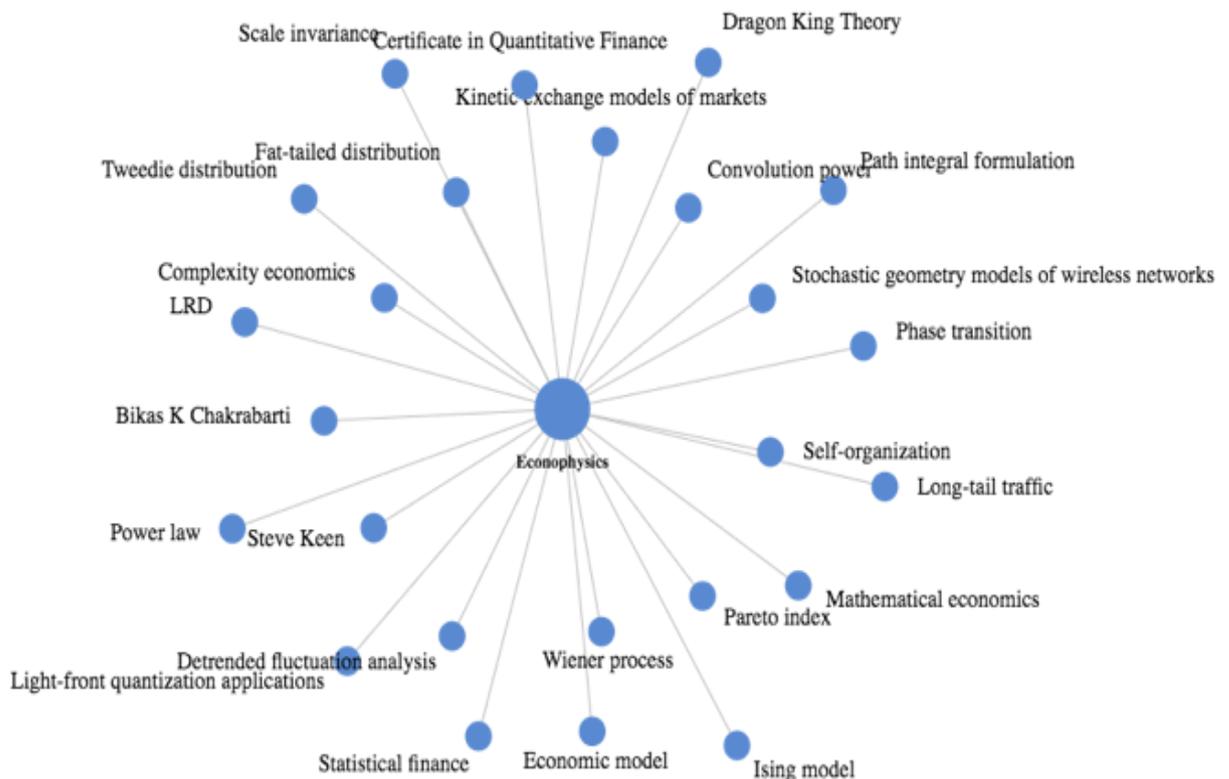
“Q: Is it possible to literature review automatically through AI”

- **Manisha:** Yes! Here are a couple of videos Yewno Discover for Literature Reviews:
[Yewno Discover Literature Review: Charles Dickens' Great Expectations](#)
[Yewno Discover - Literature Review of Moby Dick](#)

“Q: What is econophysics? (founders background)”

- **Manisha:** Econophysics is an interdisciplinary research field, applying theories and methods originally developed by physicists in order to solve problems in economics, usually those including uncertainty or stochastic processes and nonlinear dynamics. Some of its application to the study of financial markets has also been termed [statistical finance](#) referring to its roots in [statistical physics](#). Econophysics is closely related to [social physics](#).

Here's Yewno's Knowledge Graph showing you the related concepts:



“Q: I have lot of OER (open educational resources) for the village farmer, we would like to implement AI on OER? How that could be possible & helpful for illiterate underprivileged farmer?”

- **Manisha:** I'm sorry you would have to be literate to use Yewno.

“Q: Google now use the same AI technologies in its question answering when it highlights in yellow web content that will be relevant in a page?”

- **Manisha:** Sorry I don't know the answer to this question.

“Q: Have there been any comparative studies published on the systems results against traditional academic library systems (Ex Libris Alma etc) and conventional outcomes/assessments of search/retrieval and content?”

- **Ben:** Probably not, or not that I'm aware. Alma and other library systems are using old technology based on Solr indexing. This is extremely far removed from being able to classify items live according to disciplines.
- **Manisha:** I'm of the same mind as Ben on this. Yewno is reading the full text of a document and not just the metadata so any comparison would be moot.

“Q: What is the cost of implementation and what exactly is involved?”

- **Manisha:** For Yewno, please contact me directly for a price for your institution.

“Q: What tools would you recommend, Manisha, for creating and managing Knowledge Graphs?”

- **Manisha:** What do you want to use it for? If it's for research, I recommend Yewno Discover!

“Q: Even though AI checks the content and 'understands' the concept, the results seem still to be based on some selective (only out of 500m academic results) popularity, popularity nevertheless, and pushes more popular concepts. Would this not potentially hinder the approach to the original question?”

- **Ben:** In the case of Dimensions popularity is not used (this would imply using Altmetric or citations to measure popularity I presume. This is not how the service works.) The dominance of the results in certain fields for any search is based on the categories that those items describe themselves as being in. You can order results by online attention or citations etc, but conceptual classification is simply matching the category systems with the content.
- **Manisha:** Popularity is not something we use, again like Ben says, it's not how the algorithm works. Results are based on the strength of the semantic connection. Yewno doesn't discriminate between a commercial publisher's

document and a small society's document. We do not collection "popularity" information.

"Q: Multidimensional semantic vectorspace. Inference engine. What does vector and inference mean in this context?"

- **Manisha:** Multidimensional semantic vectorspace: a place where we cluster semantically similar concepts together.
Inference engine: The technology is able to extract concepts through inference (the "keyword" itself does not have to be present for us to find a concept as it is understanding the context and meaning).