Food Standards Agency
food.gov.uk

Publication Analysis Tool: Interim demo

24/02/2020

Agenda

1. Recap: Problem and objective
2. Features of the proposed tool
3. The microbiome test case and further use cases
4. Demo
5. Next steps

Problem and objective

Question: what are the trends in research and media coverage with respect to the human microbiome - and what do these suggest about how this issue will evolve over the next 5-10 years?

The aim is to build a tool to summarise these trends, to inform the new horizon scanning process.

The microbiome is being used here as a test case. It is hoped that the tool could be extended to other topics of interest in the future - and other use cases.

Features of the proposed tool

1. Fetch and store metadata about relevant articles

- Find and ingest a research backlog - what has been happening in this area over the last 5-10 years?
    - Populate metadata from manually retrieved list
    - Carry out keyword/title searches (PloS, Crossref)
- Ongoing monitoring and automated updating - regular updates of new research in this area
    - Retrieve articles from alert/ToC emails
    - Perform scheduled searches

Food Standards Agency
food.gov.uk

Features of the proposed tool

2. Create useful metrics

- Volume of research - is this a growing/fading issue, prominence of sub-categories
- Who is producing this research and what themes/topics can be identified
- Presence of key entities – companies, places, researchers, products
- Distinguishing between medical and food related research

Features of the proposed tool

3. Deliver results to the user

- User-friendly dashboard that can be understood by users across the agency
- Visualizes the metrics with interactive filtering options
- Returns list of references relevant to identified trends
- Timeline showing when trends change or emerge to help project into the future

Food Standards Agency
food.gov.uk

The microbiome test case and future use cases

- Manually retrieved references (from Web of Science), time horizon searches of two databases (PLoS and Crossref), and references scraped from email alerts (*Google Scholar*, *Journal of Agricultural and Food Chemistry*, *Microbiome*, *Cell*, *Gut* and *Scientific Reports*)

- Future use cases
    - setting up further test cases (and monitoring them)
    - uploading a list of DOIs and using the tool to visualise your literature
    - 'instant' search of key databases and visualisation of results

Demo

The Shiny App is now deloyed online

Food Standards Agency
food.gov.uk

Features of the proposed tool - progress

1. Fetch and store metadata about relevant articles

- Find and ingest a research backlog - what has been happening in this area over the last 5-10 years?
    - Populate metadata from manually retrieved list - *YES*
    - Carry out keyword/title searches (PloS, Crossref) - *YES but want to expand to PubMed and others*
- Ongoing monitoring and automated updating - regular updates of new research in this area
    - Retrieve articles from alert/ToC emails - *YES*
    - Perform scheduled searches - *NOT YET!*

Food Standards Agency
food.gov.uk

Features of the proposed tool

2. Create useful metrics

- Volume of research - is this a growing/fading issue, prominence of sub-categories - *YES for whole database NOT YET for sub-topics*
- Who is producing this research and what themes/topics can be identified - *YES although better text cleaning would improve this*
- Presence of key entities – companies, places, researchers, products - *NOT YET*
- Distinguishing between medical and food related research - *PARTIALLY*

Food Standards Agency
food.gov.uk

Next steps

- Remainder of this phase
  - storage and automation
  - text cleaning
- Next stage

We want to hear from you!

Do you think you would use this or a similar tool?

datascience@food.gov.uk

https://github.com/FoodStandardsAgency/hs-test-case-dev