

CASE-STUDY IN INFORMATION RETRIEVAL

A critical evaluation of search methods for sourcing information on the internet about links between diet and socio-economic status in the UK

Veronica Tuffrey
April 2020

Module INM305 Information Retrieval, MSc Data Science - City, University of London

1. INTRODUCTION

This document reports the findings of an information retrieval (IR) task chosen by the author, and an evaluation of the search methods applied. The search methods and evaluation methods were specified by the university in the assignment instructions.

1.1 Topic, motivation, and task type

The topic chosen for this case-study relates to linkages between dietary intake and socio-economic status in the UK. The author's interest in the topic is both personal and professional (the findings may eventually feed into her MSc dissertation research). Box 1 provides the topic description in the style needed for TREC workshops¹ [1].

- **Title:** UK diet and socio-economic status
- **Description:** Find information on dietary inequalities related to socio-economic status in the UK
- **Narrative:** Relevant documents will provide information about dietary inequalities in the UK. Relevant documents may relate to dietary intake of single or multiple food items, single or multiple food groups, single or multiple nutrients, or dietary "patterns". Documents with general statements that dietary inequalities exist are irrelevant, as are documents that describe inequalities in health, and/or describes dietary patterns without linking them with socio-economic status.

Box 1: Topic for the case-study

This IR task falls into the category of "conscious topical needs" rather than "verificative needs" or "muddled topical needs"² as defined by Ingwersen [2], because the user (this report's author) was aware that the phenomenon of dietary inequality in the UK exists, and required details about it. As described by Belkin et al. [3], the driving force of the task (the information need) is the user's recognition of inadequacy in their state of knowledge. This "anomaly with respect to the problem" led to an "anomalous state of knowledge" or ASK. In this case the ASK is a simple gap, not an internal inconsistency, or conflict with evidence.

1.2 Process of retrieval task

To undertake the search, the user started from identification of the information need (cognitive stage, as described in Section 1.1), and then created a search to be submitted to an information retrieval system (syntactic stage) [4]. The syntactic stage was divided into the linguistic stage (achieved using facet analysis as described in Section 2); the strategic stage (method used to combine facets), and the tactical stage (choice of terms and choice of operators) which are both described in Section 3.

The evaluation of the search engines required a binary classification of all retrieved items as "relevant" or "not relevant". For this exercise, a relevant document was defined as one that provided pertinent information that met the narrative description in Box 1, or such a relevant document was retrievable using one-click from the page returned.

¹ TREC ("Text REtrieval Conference") is workshop series, run by the US National Institute of Science and Technology, that focuses varied information retrieval research areas, or "tracks".

² *Verificative needs* exist when the user wants to verify or locate specific items, and bibliographic data relating to the source are *known* to the user; *Conscious topical needs* exist when the user wants to clarify, review or pursue aspects of a *known subject matter*; while *Muddled topical needs* exist when the user wants to explore some new concepts or concept relations *outside known* subject matter.

1.3 Process and policy of evaluation

The methods of evaluation applied are listed below in Table 1. The first three measures relate to precision, and the next four measures are classed as “diagnostic”, as they provide information on issues that might affect precision, and therefore help in interpretation of the precision measures. Only the top ten ranked documents were examined for each search. Table 2 provides definitions of the measures, and any underlying assumptions.

Measure	Explanation	Metric
Precision measures		
P @ 5	Precision at 5 documents retrieved	Between 0 and 100%
P @ 10	Precision at 10 documents retrieved	Between 0 and 100%
EAP	Estimated Average Precision for the top 10 documents	Between 0 and 100%
Diagnostic measures		
RT – Dup	Rate of Repeated documents = Number of duplicates per search	Between 0 and 10
LB	Link Broken = Number of broken links per search.	Between 0 and 10
NT	Not retrieved = Number of documents not retrieved by that search.	Between 0 and 10
Spam	Spam = Number of Spam documents per search	Between 0 and 10

Table 1: Evaluation methods used

Measure	Definition/process or assumption
P@5	Definition: (Number of relevant documents in first 5 documents retrieved) * 100% / 5
P@10	Definition: (Number of relevant documents in first 10 documents retrieved) * 100% / 10
EAP	Process: For the first 10 documents retrieved, relevance is assessed. For each relevant document, the precision at that rank is calculated using the above definition (replacing 5 or 10 by the document's rank). For other documents precision is 0. Precision values for all ten documents are summed and divided by the number of documents retrieved (here 10). Assumption: The calculation assumes there are at least 10 relevant documents.
RT - Dup	Definition: The webpage appears identical to, and contains the same information as, a webpage retrieved earlier. This is only noted for relevant documents.
LB	Definition: Error message shows attempt to access page was unsuccessful
NT	Definition: 10 – (Number of documents retrieved) Assumption: There are at least 10 relevant documents.
Spam	Definition: Highly irrelevant document returned due to the manipulation of web page content for the purpose of appearing high up in search results for selected keywords [5]

Table 2: Definitions and assumptions underlying the evaluation methods

2. FACET ANALYSIS

To enable structured searching, “facet analysis” was first used to decompose the topic into component elements or “facets” thereby facilitating flexibility and expressive power [6].

2.1 Facets used for search

Table 3 shows how the case-study topic was decomposed into facets using the PMEST approach (see Section 2.2).

Table 3 (left): Facets used for case-study

Facet category	Case-study facets
Personality	Diet
Matter	Food or nutrients
Energy	Socio-economic variation
Space	UK
Time	Present day

2.2 Development of facets

Various potential approaches for facet analysis were considered. A well-accepted approach for facet analysis is PMEST, which is founded on methodology developed by Ranganathan in 1933 for library classification [7], and by which topics are decomposed into five fundamental categories of personality, matter, energy, space, and time, see Table 4 for explanations. Also frameworks such as PICO (Population, Intervention, Comparator, and Outcome) [8] can be used for facet analysis preceding searches relating to health interventions, and ad-hoc schemes using the user's own choice of concepts are an alternative option.

Facet category	Relates to the...
Personality	..most specific or core point of the topic (referred to as “the subject” below)
Matter	..substance, properties or materials of the subject
Energy	..processes, operations, activities, actions, or problems
Space	..geographical location of the subject
Time	..dates or seasons of the subject

Table 4: Explanation of PMEST approach for facet analysis

For other topics, the author has experienced difficulties with the PMEST approach, as it can be challenging to distinguish between Personality, Matter, and Energy. In those cases, she adopted an ad-hoc scheme to ensure all key elements of the topic were encompassed. But for the case-study topic, she found it straightforward to include all necessary aspects using PMEST, for example the facet “matter” could embrace the various features of “diet”.

Although this study did not pertain to a health intervention, the author did consider using an adapted version of PICO (where Population = UK population; Intervention=Food and Nutrients; Comparator = Socio-economic status and Outcome = Diet). But this approach did not offer any advantage over PMEST, and neither did an ad-hoc scheme. Further, PMEST has an advantage over PICO by including a time aspect. Although initially the time facet was considered useful, ultimately it was not used. (The narrative (Box 1) does not mention time trends, but the author chose to maintain the option of including a historical perspective in her dissertation literature review, where the decision would depend on the documents returned).

3. SEARCH STRATEGY

3.1 Strategy used to generate search queries

The task instructions defined the search strategy to comprise both Boolean and “bag-of-words” (alternatively termed “basic” [9]) searches. The strategy adopted was to fine-tune the Boolean search query to optimise returns from the academic databases, and to fine-tune the bag of words query to optimise returns from web search engines. In other words, the queries were not optimised for individual engines, to enable fair comparison between their outputs.

i. For the *Boolean searches*, the author generally adopted the conventional strategy of applying an OR operator to the terms and synonyms within facets, and then applying the AND operator between facets [4]. Having considered the three well-known strategies for combining facets: building blocks, successive fractions and citation pearl growing [4], the building blocks approach was adopted rather than successive fractions due to the author’s preference to review the volume of documents returned for each facet before the facets were combined. The citation pearl growing approach was used before combining facets to identify search terms additional to those initially identified using logic. Table 6 shows the terms initially compiled in the left column, and the final terms used for the evaluation in the right column. The latter incorporates keywords for documents returned in early Google searches, and also terms used to exclude documents, identified from early searches using the information service Proquest.

ii. For the “*bag of words*” searches, the citation pearl approach was also used to try to fine-tune choice of terms for the evaluation searches. None were changed, see Table 7.

Facet	Initial Search	Final Search
Personality (S1)	diet*	
Matter (S2)	food OR nutrient*	diet* OR food OR *nutrient* OR eat* OR intake
Energy (S3)	(Socio-economic AND variation) OR inequalit*	(soci* OR income OR education) AND (inequalit* OR diff* OR varia*)
Space (S4)	UK OR United Kingdom OR England OR Wales OR Scotland OR North* Ireland	UK OR United Kingdom OR Wales OR Welsh OR Engl* OR Scotland OR Scottish OR North*Ireland OR North*Irish OR GB or Brit*
For exclusion (S5)		NOT(society OR impact OR intervention OR evaluation OR advert* OR media OR environment)
Combination	S1 AND S2 AND S3 AND S4	S2 AND S3 AND S4 AND NOT S5

Table 6: Queries used for Boolean searches

Initial Search: UK diet socio-economic inequality	Final Search: UK diet socio-economic inequality
--	--

Table 7: Queries used for initial and final “bag of words” searches

3.2 Tactics used to generate search queries

i. Boolean searches: In order to reduce recall and increase precision, the decision was quickly made to search for keywords only in document titles. Early Boolean searches using the Proquest database with four facets (left columns of Table 6) revealed the best approach was to reduce the number of facets to specify the documents required (from five to four), and to add an extra search using keywords to exclude documents on certain topics. This was because initially many documents were returned describing studies of: ways socio-economic inequalities in diet could be addressed using public health interventions; the influence of advertising; the influence of social media; sustainable diets, and findings reported by academic Societies. Thus, precision was increased by using keywords to exclude documents.

The wild-card operator was helpful to ensure inclusion of multiple variants of keywords. For example, diet* would also return dietary and diets; *nutrient* would return macro-nutrient, micro-nutrient and nutrients, and soci* would return documents with the terms social class, socio-economic, and socioeconomic in them. It was necessary to include both the keywords Scotland and Scottish, because Scot* returned documents about Nova Scotia.

ii. Bag of words searches: Trials using Google showed the exact forms of keywords included did not influence the documents returned (e.g. socio-economic vs socio economic) and also that for maximum precision it was preferable to include as few terms as possible (e.g. to just include the term diet, and not also food, fruit, vegetables, fat, sugar and/or salt).

4. EVALUATION OF SEARCHES

4.1 Findings from searches

The process and policy for searches was described in Section 1.3. Table 1 gives a summary of the results, while details of search engines, queries and items returned are in the Appendices.

	Search Service	Query Type	P@5	P@10	EAP	RT-Dup	LB	NT	SPAM
1	Google web search	Boolean	20%	10%	5.0%	0	0	0	0
		Bag of Words	60%	40%	29.3%	4	0	0	0
2	Bing Web search	Boolean	0	0	0	0	1	0	0
		Bag of Words	60%	30%	26.0%	1	0	0	0
3	Google images	Bag of Words	0	0	0	0	0	0	0
4	Bing images	Bag of Words	20%	10%	10.0%	2	0	0	0
5	YouTube	Bag of Words	0	0	0	0	0	0	0
6	Bing Video	Bag of Words	0	0	0	0	0	0	0
7	ProQuest Dialog	Boolean	60%	50%	39.7%	4	0	0	0
8	DuckDuckGo	Boolean	0	0	0	0	0	0	0
		Bag of Words	40%	20%	15.0%	1	0	0	0
9	Social Searcher	Bag of Words	40%	20%	20.0%	0	0	0	0
10	Startpage	Boolean	20%	10%	2.5%	0	0	0	0
		Bag of Words	60%	40%	27.7%	4	0	0	0
11	Web of Science	Boolean	40%	70%	47.3%	0	0	0	0

Table 8: Summary of results using the varied search methods

4.2 Evaluation of the search engines: Comparison of search engines

i. Best search service: Table 8 shows that several services had the same good result when one just considers the first five documents returned (P@5 of 60%) - Proquest Dialog, Google, Bing and Startpage. Web of Science had the best result when one considers the first ten (P@10 of 70% and EAP of 47%). Web of Science provides access to multiple databases, some of which are academic. Since the search task required documents containing information of sufficient reliability for inclusion in the literature review for an academic assignment (Section 1.2), this finding that an academic search engine was the most productive was expected.

ii. Comparison of online search with web search: Table 8 shows online bag of words searches were as effective as the Boolean searches using academic engines when considering only the first five documents returned (see i). Compared to P@5, the lower values for P@10 (40% for Google and Startpage and 30% for Bing) were in large part due to the high values for RT-Dup (four repeated documents for Google and Startpage). This finding that bag of words web search led to precision values similar to online search was unexpected. The exercise also clearly showed the value of undertaking both online and web searches, as although there was some overlap in items returned, there were also several different ones.

iii. Comparison of 'bag of words' search with 'Boolean' search: For the web search engines, "Bag of Words" searches gave more precise outputs than Boolean. However, as mentioned in Section 3.1, the Boolean search query used was optimised for use with academic databases, and it is likely that the precision of Boolean searches using the web engines would be increased by simplifying the query (in fact the query had already been simplified somewhat by removing the keywords used with the AND NOT requirement, because the engines had appeared to use these terms to return items, rather than exclude them). It seems that complex Boolean searches are impractical outside the context of specialised information services.

iv. Comparison of web search with meta search: Table 8 shows that the values for precision from bag of words searches using the DuckDuckGo meta-search engine were lower than those using Google, Startpage and Bing, even though the number of repeated documents returned by DuckDuckGo was lower (only one). The precision of the Boolean search using DuckDuckGo was zero, like the Web search engines Bing. However, one difference was noted between web search and meta-search - the outputs from searches using the latter were quite unstable. The exact documents returned differed each time the search was run using DuckDuckGo (the precision stayed at zero).

v. Use of search engines for images and videos: Searches for images using Google and Bing, and for videos using YouTube and Bing, did not help to achieve the retrieval task. The precision values were almost uniformly 0%. The images and videos returned were irrelevant because they related to health and economic inequalities rather than dietary inequalities. Interestingly, the first (and only) relevant image returned (it was through Bing) was of the first page of the most useful document returned via the text searches.

vi. Use of Social Search: The engine Social Searcher returns items from social networks like Facebook and Twitter. While the first two documents returned were relevant, they had already been found via the traditional services. Given the task's academic character the finding that Social Search did not swell the volume of relevant documents was expected.

vii. Comparison of private engines with the main search engines: The evaluation results using the privacy-oriented search engines StartPage and DuckDuckGo did not differ in a consistent way from those using Google and Bing. Just like the main search engines, for this task the privacy-oriented engines had poor results using the Boolean search query submitted, and quite good results using the bag of words search query.

4.3 Evaluation of the search engines: Impact of user interfaces on the results

i. For the text-based web searches the user interfaces for were all very similar and so it is unlikely they contributed to the differences between the engines seen in Table 8. In accordance with the task instructions, the tool 2Dsearch (<https://www.2dsearch.com/>) was used to construct the Boolean query for submission to Google and Bing. While this tool can aid formulation of Boolean queries for web searches, for the current task the Boolean query had already been formulated for use with Proquest. So, the author's challenge was simply learning to use the interface. The providers of the webpage extol the benefits of using a "two-dimensional canvas" to combine objects rather than a "one-dimensional search box", but in the author's view, it is only for complex searches that the tool would provide a notable benefit.

ii. For the image and video searches, the Google images interface looks like the text interface, and the Bing images interface has a background photo (which is pleasant but would not have impacted on the search). In contrast, the Bing and YouTube video interfaces show screenshots of suggested videos, which were successful in distracting the author and thereby impacted the task at hand by delaying relevant activity.

iii. For the online search engines, even though the author had existing experience using academic databases such as Medline, she found the ProQuest Dialog interface was not particularly intuitive. The Web of Science interface was easier to navigate as the font was larger and the items more spaced out on the screen. It is likely that the author found Web of Science easier because she had already gained experience using ProQuest.

4.4 Evaluation of the search engines: *Quality of material retrieved*

Appendix D shows the information sources returned by the text searches³ were of multiple types - the majority were academic journal articles, and news stories about such articles (noted as duplicates if the original article had been judged as relevant). Additionally, reports and blogs originating from academic, governmental, non-governmental and United Nations organisations were sourced. Only DuckDuckGo returned essays sourced from essay-writing mills.

With respect to the documents returned that were judged as containing relevant subject matter (see Section 1.2), the author evaluated quality as good in most cases. Most were academic articles in peer-reviewed journal so had been independently assessed by others working in the same field. While peer-review provides no guarantee of the reliability of the information⁴, the system confers a higher degree of legitimacy to content that has passed through it, compared to other sources. This report's author has familiarity with the process as both author and reviewer and is aware through colleagues of the vital contribution that journal editors also make. So, she is likely to be biased towards judging the quality of information from peer-reviewed sources as high. She also trusts information from government sources, and again this view is biased by having undertaken work supported by government agencies.

4.5 Reflection on the process and result of relevance assessment

Relevance is defined as a "*measure of information conveyed by a document relative to a query*" [10]. As described in Section 1.2, for this case-study, an item was judged relevant if it provided (or was linked by no more than one click to another item that provided) information of sufficient pertinence with respect to the narrative provided in Box 1. Unlike other evaluations like TREC [1], there was no chance of variations in, or wrong interpretation of, the narrative, since only one person was doing the evaluation, and this same person wrote the narrative.

To assess relevance, a tick-box exercise was used requiring 100% positive responses to the questions based on the narrative in Box 1. It was impossible to assess relevance if an item was not accessible in the public domain, and then the item was deemed irrelevant. Initially there was a question about whether the document reported new findings or simply provided overviews of others' research, but later this question was discarded because overviews were re-evaluated as useful. Also, initially documents published earlier than 2010 were rejected, but later this criteria was dropped in order to gain a historical perspective to the topic.

The author may ultimately use those items deemed relevant to inform part of her dissertation introduction. This goal provided motivation to apply the exercise carefully. Fortunately, sufficient items were returned, there is no need for further information retrieval. In future she would experiment with using more fields than just the title when running database searches.

5. SUMMARY

The case-study reported was based on the 'ASK hypothesis' and the 'best-match' principle, that is, given a representation of a request for information ('query'), the best response will be the text whose representation most closely matches it [3]. Several findings have practical importance for researchers such as the author, and so the evaluation proved to be a valuable as well as interesting exercise for her. Examples are the worth of undertaking searches using both online and web engines; the need to fine-tune queries for individual search engines to best harness their power, and the benefits of reviewing more than the first ten items returned.

³ The videos and images returned were judged irrelevant (except for one image) so are not considered.

⁴ One eminent scientist commented "*We have little or no evidence that peer review 'works,' but we have lots of evidence of its downside*" [10].

APPENDICES

A REFERENCES

- [1] Chowdhury GG. *Searching and retrieval*. 3rd edn. London: Facet Publishing; 2010.
- [2] Ingwersen P. *Information Retrieval Interaction*. London: Taylor Graham; 1992.
- [3] Belkin NJ, Oddy RN, Brooks HM. ASK for Information Retrieval: Part I. Background and Theory. *Journal of Documentation* 1982; 38:61-71.
- [4] MacFarlane A, Russell-Rose T. *Search Strategy Formulation: A Framework For Learning*. In *Proceedings of the 4th Spanish Conference on Information Retrieval*. 2016: 1-8.
- [5] Manning CD, Raghavan P, Schütze H. *Introduction to information retrieval*. Cambridge UK: Cambridge University Press; 2008.
- [6] Tunkelang D. *Faceted Search*. In *Synthesis Lectures on Information Concepts, Retrieval, and Services*. Edited by Marchionini G. Chapel Hill, US: University of North Carolina; 2009
- [7] Ranganathan SR. *Colon Classification*. Madras, India: India: Madras Library Association; 1933.
- [8] NICE. *Chapter 4 Developing review questions and planning the evidence review*. In *Developing NICE guidelines: the manual*. London: NICE (National Institute for Health and Care Excellence); 2018
- [9] Bradley P. *Expert internet searching*. London: Facet Publishing; 2017.
- [10] Smith R. Classical peer review: an empty gun. *Breast cancer research* 2010; 12:S13.

B TABLE OF SEARCH ENGINES AND THEIR WEB ADDRESSES

	Search Service	Query Type, and Web address
1	Google Web Search	Boolean: https://app.2dsearch.com
		Bag of Words: http://www.google.co.uk
2	Bing Web Search	Boolean: https://app.2dsearch.com
		Bag of Words: http://www.bing.com/?cc=gb
3	Google images	Bag of Words: http://www.google.co.uk/imghp?hl=en&tab=wi
4	Bing Images	Bag of Words: http://www.bing.com/?scope=images&nr=1&FORM=NOFORM
5	YouTube	Bag of Words: http://www.youtube.com/?gl=GB&hl=en-GB
6	Bing Video	Bag of Words: http://www.bing.com/videos/browse
7	ProQuest Dialog	Boolean: https://trials.proquest.com/access?token=QFAOSBYXQ1HNSIUF6UXU
8	DuckDuckGo	Boolean: https://duckduckgo.com/
		Bag of Words: https://duckduckgo.com/
9	Social Searcher	Bag of Words: https://www.social-searcher.com/
10	Startpage	Boolean: https://startpage.com/uk/advanced-search.html?hmb=1
		Bag of Words: https://startpage.com/
11	Web of Science	Boolean: accessed via http://libguides.city.ac.uk/az.php

C TABLES OF SEARCH QUERIES

Boolean web search	
1A, 2A, 8A, 10A	(diet OR eat OR food OR intake OR nutrient) (education OR income OR socio-economic) (difference OR inequality OR variation) (England OR "Northern Ireland" OR Scotland OR UK OR "United Kingdom" OR Wales OR GB OR Britain)
"Bag of Words" web search	
1B, 3,4,5,6,8B, 10B	UK diet socio-economic inequality

	Boolean web search using Proquest Dialog	Field	Number of documents returned
Set 1	ti(diet* OR food OR *nutrient* OR eat* OR intake)	Document title	10,969,026
Set 2	ti((soci* OR income OR education) AND (inequalit* OR diff* OR varia*))	Document title	212,416
Set 3	ti(UK OR United Kingdom OR Wales OR Welsh OR Engl* OR Scotland OR Scottish OR North*Ireland OR North*Irish OR GB OR Brit*)	Document title	9,247,132
Set 4	ti(society OR impact OR intervention OR evaluation OR advert* OR media OR environment)	Document title	20,619,934
	S1 AND S2 AND S3 AND NOT S4		44

	Boolean web search using Web of Science	Field	Number of documents returned
Set 1	TI=(diet OR food OR *nutrient* OR eat* OR intake)	Document title	538,842
Set 2	TI = ((soci* OR income OR education) AND (inequalit* OR diff* OR varia*))	Document title	36,386
Set 3	TI = (UK OR United Kingdom OR Wales OR Welsh OR Engl* OR Scotland OR Scottish OR North*Ireland OR North*Irish OR GB or Brit*)	Document title	482,908
Set 5	#1 AND #2 AND #3		27

D TABLES OF OUTPUTS

	Short-code for cross-reference	Topic of item	Website address	Rel evant?	Reason for non-relevance.
1A Google Boolean search via 2Dsearch					
1	Gov	Poverty and income inequality statistics	https://www.gov.scot/collections/poverty-and-income-inequality-statistics/	No	Poverty and income
2	Ji	Socioeconomic inequality in salt intake in Britain 10 years after a national salt reduction programme	https://bmjopen.bmj.com/content/4/8/e005683	Yes	
3	Barton	Time to address the double inequality of differences in dietary intake between Scotland and England	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6088540/	No	Geographical comparison, not Soc-Ec
4	Barton	Time to address the double inequality of differences in dietary intake between Scotland and England	https://www.ncbi.nlm.nih.gov/pubmed/29947325	No	(repeat of irrelevant doc)
5	F.Times	Income Inequality increases in UK	https://www.ft.com/content/06c0933c-5ed8-11ea-b0ab-339c2307bcd4	No	Income inequality
6	Guardian	Large rise in takeaway shops highlights dominance of fast food in deprived areas	https://www.theguardian.com/inequality/2017/jul/25/large-rise-takeaway-shops-highlights-dominance-fast-food-deprived-areas-england	No	No data
7	Jay Rayner	Diet, health, inequality: why Britain's food supply system doesn't work.	Guardian Newspaper website, article on March 22, 2020	No	No details
8	Oxfam	Food Poverty	https://policy-practice.oxfam.org.uk/our-work/inequality/food-poverty	No	No data
9	NHS	Food Poverty	http://www.healthscotland.scot/media/2222/food-poverty.pdf	No	No data
10	UCL	Diet makes a difference to learning	http://www.educationalneuroscience.org.uk/resources/neuromyth-or-neurofact/diet-makes-a-difference-to-learning/	No	Psychology

1B Google web search using "bag of words"					
1	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	Yes	
2	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	No	REPEAT
3	Roberts	Social and economic inequalities in diet and physical activity	http://www.oneeastmidlands.org.uk/sites/default/files/library/vid_19253_Social_and_economic_inequalities_in_diet_and_physical_activity_04.11.13.pdf	Yes	
4	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.cambridge.org/core/journals/british-journal-of-nutrition/article/socioeconomic-dietary-inequalities-in-uk-adults-an-updated-picture-of-key-food-groups-and-nutrients-from-national-surveillance-data/9F814CB09FCC78ED8C4F1E1E1D80737B	No	REPEAT
5	Tiffin	Inequalities in diet and nutrition	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	Yes	
6	Yau	Time trends in adherence to UK dietary recommendations and associated sociodemographic inequalities, 1986-2012: a repeated cross-sectional analysis	https://www.nature.com/articles/s41430-018-0347-z	yes	
7	Roberts	Social and economic inequalities in diet and physical activity	http://www.oneeastmidlands.org.uk/sites/default/files/library/vid_19253_Social_and_economic_inequalities_in_diet_and_physical_activity_04.11.13.pdf	No	REPEAT
8	Alkerwi	Demographic and socioeconomic disparity in nutrition: application of a novel Correlated Component Regression approach	https://bmjopen.bmj.com/content/5/5/e006814	No	About Luxembourg
9	Dowler1	Food and health inequalities: the challenge for sustaining just consumption	https://www.tandfonline.com/doi/abs/10.1080/13549830802478736?mobileUi=0&journalCode=cloe20	No	No data
10	PHE	New briefing paper explores inequalities in diet and physical activity	https://www.gov.uk/government/news/new-briefing-paper-explores-inequalities-in-diet-and-physical-activity	No	REPEAT

2A Bing Boolean search					
1	D'Odorico	Food Inequality, Injustice, and Rights	https://academic.oup.com/bioscience/article/69/3/180/5320252	No	About human rights
2	Wikipedia	Diet (nutrition)	https://en.wikipedia.org/wiki/Diet_(nutrition)	No	Just about diet
3	Tech	Diet variation, quantification, and misreporting	https://dapa-toolkit.mrc.ac.uk/diet/introduction/diet-variation-quantification-misreporting	LINK BROKEN	No data
4	?	Food inequity	https://food-inequality.weebly.com/definition.html	No	Global
5	UCL	Diet makes a difference to learning	http://www.educationalneuroscience.org.uk/resources/neuromyth-or-neurofact/diet-makes-a-difference-to-learning/	No	Psychology
6	FAO	The state of food security and nutrition in the world	http://www.fao.org/state-of-food-security-nutrition/en/	No	Global
7	BNF	A healthy balanced diet	https://www.nutrition.org.uk/healthylivin	No	Just diet

			g/healthydiet/healthybalanceddiet.html		
8	BBC	What is a balanced diet? - BBC Bitesize	https://www.bbc.co.uk/bitesize/topics/zrf/fr82/articles/zppvv4j	No	Just diet
9	NHS	Top diets review	https://www.bbc.co.uk/bitesize/guides/z9gk87h/revision/2	No	Just diet
10	Gov	Variation	https://www.bbc.co.uk/bitesize/guides/z9gk87h/revision/2	No	Biology

2B Bing web search using "bag of words"					
1	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	Yes	
2	Roberts	Social and economic inequalities in diet and physical activity	http://www.oneeastmidlands.org.uk/sites/default/files/library/vid_19253_Social_and_economic_inequalities_in_diet_and_physical_activity_04.11.13.pdf	Yes	
3	Eureport	Health Inequalities: dietary and physical activity-related determinants	https://ec.europa.eu/jrc/en/health-knowledge-gateway/societal-impacts/inequalities	No	About Europe, no data
4	Roberts	New briefing paper explores inequalities in diet and physical activity	https://www.gov.uk/government/news/new-briefing-paper-explores-inequalities-in-diet-and-physical-activity	No	REPEAT
5	Tiffin	Inequalities in diet and nutrition	http://centaur.reading.ac.uk/24681/1/offprint.pdf	Yes	
6	blog	Tackling Socio-Economic Inequality In The UK Is An Urgent Human Rights Issue	https://eachother.org.uk/tackling-socio-economic-inequality-in-the-uk-is-an-urgent-human-rights-issue/	No	Economic inequalities
7	Essay	Socio Economic Inequalities: Health	https://www.ukessays.com/essays/health-and-social-care/socio-economic-inequalities-health-and-social-care-essay.php	No	Health inequalities
8	blog	The Scale of Economic Inequality in the UK	https://www.equalitytrust.org.uk/scale-economic-inequality-uk	No	Economic inequalities
9	pechey	Socioeconomic inequalities in the healthiness of food choices: Exploring the contributions of food expenditures	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4910945/	No	About shopping
10	gov	Chapter 5: inequality in health	https://www.gov.uk/government/publications/health-profile-for-england/chapter-5-inequality-in-health	No	Health inequalities

3A Google images using "bag of words"					
1	Gov	Chapter 5: inequality in health	https://www.gov.uk/government/publications/health-profile-for-england/chapter-5-inequality-in-health	No	Health inequalities
2	Gov	Understanding health inequalities in England PHE	https://publichealthmatters.blog.gov.uk/2017/07/13/understanding-health-inequalities-in-england/	No	Health inequalities
3	Gov	Chapter 5: inequality in health	https://www.gov.uk/government/publications/health-profile-for-england/chapter-5-inequality-in-health	No	(repeat of irrelevant doc)
4	Gov	Understanding health inequalities in England PHE	https://publichealthmatters.blog.gov.uk/2017/07/13/understanding-health-inequalities-in-england/	No	(repeat of irrelevant doc)
5	Book	Inequalities in health (e.g. by region, ethnicity, socio-economic position or gender) and in access to health care, including their causes	https://www.healthknowledge.org.uk/public-health-textbook/medical-sociology-policy-economics/4c-equality-equity-policy/inequalities-distribution	No	Health inequalities

6	Gov	Chapter 5: inequality in health	https://www.gov.uk/government/publications/health-profile-for-england/chapter-5-inequality-in-health	No	(repeat of irrelevant doc)
7	Gov	Chapter 5: inequality in health	https://www.gov.uk/government/publications/health-profile-for-england/chapter-5-inequality-in-health	No	(repeat of irrelevant doc)
8	Gov	Chapter 5: inequality in health	https://www.gov.uk/government/publications/health-profile-for-england/chapter-5-inequality-in-health	No	(repeat of irrelevant doc)
9	Gov	2 Health inequalities - extent, causes, and policies to tackle them	https://publications.parliament.uk/pa/cm200809/cmselect/cmhealth/286/28605.htm	No	Health inequalities
10	poverty and health	Poverty and health	https://www.health.org.uk/infographic/poverty-and-health	No	Health inequalities

4 Bing images using "bag of words"					
1	Maguire	Maguire first page		yes	
2	Maguire	Maguire table		no	REPEAT
3	Maguire	maguire table		no	REPEAT
4	Cookson	Socio-economic inequality in healthcare		no	Health inequalities
5		Medical advancements and the growing inequality in healthcare		no	Health inequalities
6	Sayed	Socio-economic inequalities in childhood obesity		no	Body size not diet
7	PHE	Factors that affect health outcomes		no	Health inequalities
8		Stunting rates in India		no	Body size not diet
9	SocHealth	Mortality by SES		no	Health inequalities
10		Sustainable development goals		no	Global

5 YouTube using "bag of words"					
1		Inequality in the UK		no	Economics
2		Income inequality in the UK		no	About income
3		UK's income inequality crisis		no	About income
4		How economic inequality harms societies		no	Economics
5		Social inequality, Khan academy		no	Sociology
6		Social inequalities in health		no	Health inequalities
7		Economic Inequality, TED talk		no	Economics
8		Social issues in the UK (inequality)		no	Sociology
9		40 years after Thatcher: inequality in uk		no	Sociology
10		How economic inequality harms societies		no	Economics

6 Bing Videos using "bag of words"					
1		Income inequality in the UK	High Pay Centre	no	Health inequalities
2		Wealth inequality in the UK	Inequality Briefing	no	Wealth inequalities
3		income and wealth inequality	Crash course	no	Economics
4		health inequalities	NIHRtv	no	Health inequalities
5		health inequalities	Michael Marmot	no	Health inequalities
6		Tax, inequality and human rights	CESRvideo	no	Economics
7		Report highlights race inequality in the UK	5 News	no	Racial inequalities
8		Structural inequalities	Namibian broadcast	no	Economics
9		Social gradient of health and income inequality	Tomo	no	Health inequalities
10		Focus on health inequalities	ARC NWC	no	Health inequalities

7 Proquest Diagog using Boolean Search					
1	Barton	Trends in socio-economic inequalities in the Scottish diet: 2001-2009	Public Health Nutrition journal	yes	
2	McNeill	Socio-economic differences in diet, physical activity and leisure-time screen use among Scottish children	Public Health Nutrition	yes	

		in 2006 and 2010: are we closing the gap?	journal		
3	News	News article about McNeill study		No	REPEAT
4	Bolton-Smith	Nutrient intakes of different social-class groups - results from the Scottish-heart-health-study (SHHS)	British Journal of Nutrition	Yes	
5	Reynolds	Healthy and sustainable diets that meet greenhouse gas emission reduction targets and are affordable for different income groups in the UK	Public Health Nutrition journal	No	About environment
6	Thane	Intake and sources of phylloquinone (vitamin K-1): variation with socio-demographic and lifestyle factors in a national sample of British elderly people	British Journal of Nutrition	Yes	
7	News	News article about McNeill study		No	REPEAT
8	News	News article about McNeill study		No	REPEAT
9	Schofield	The diets of pregnant and post-pregnant women in different social groups in London England and Edinburgh Scotland UK - calcium iron retinol ascorbic acid and folic acid	British Journal of Nutrition	yes	
10	News	News article about McNeill study		No	REPEAT

8A DuckDuckGo using Boolean Search					
1	Jay Rayner	Diet, health, inequality: why Britain's food supply system doesn't work.	Guardian Newspaper website, Article on March 22, 2020	No	No details
2	Mayo clinic	Diabetes diet: Create your healthy-eating plan	https://www.mayoclinic.org/diseases-conditions/diabetes/in-depth/diabetes-diet/art-20044295	No	Dietary advice
3	Mayo clinic	DASH diet: Healthy eating to lower your blood pressure	https://www.mayoclinic.org/healthy-lifestyle/nutrition-and-healthy-eating/in-depth/dash-diet/art-20048456	No	Dietary advice
4	Data	Food Supply	https://ourworldindata.org/food-supply	No	Global
5	D'Odoric o	Food Inequality, Injustice, and Rights	https://academic.oup.com/bioscience/article/69/3/180/5320252	No	About rights
6	Howley	Vegan vs. Vegetarian	https://health.usnews.com/wellness/compare-diets/articles/vegan-vs-vegetarian	No	About diets
7	?	Food inequity	https://food-inequality.weebly.com/definition.html	No	Global
8	Diet Doctor	Low carb and keto made simple	https://www.dietdoctor.com/	No	Dietary advice
9	Medical News Today	What is the difference between veganism and vegetarianism?	https://www.medicalnewstoday.com/articles/325478	No	Dietary advice
10	National Institutes of Health	Nutrient Recommendations: Dietary Reference Intakes (DRI)	https://ods.od.nih.gov/Health_Information/Dietary_Reference_Intakes.aspx	No	Nutritional principles

8B DuckDuckGo using Bag of Words					
1	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	Yes	
2	blog	Tackling Socio-Economic Inequality In The UK Is An Urgent Human Rights Issue	https://eachother.org.uk/tackling-socio-economic-inequality-in-the-uk-is-an-urgent-human-rights-issue/	No	Economics
3	blog	The Scale of Economic Inequality in the UK	https://www.equalitytrust.org.uk/scale-economic-inequality-uk	No	Economics
4	Tiffin	Inequalities in diet and nutrition	http://centaur.reading.ac.uk/24681/1/0ffprint.pdf	Yes	
5	Essay	Income Inequalities and Socio-Economic Development	https://www.ukessays.com/essays/economics/income-inequality-and-economic-and-social-developments-economics-essay.php	No	Income inequalities
6	Vinke	Socio-economic disparities in	https://www.thelancet.com/journals/ec	No	About

		the association of diet quality and type 2 diabetes incidence in the Dutch Lifelines cohort	linm/article/PIIS2589-5370(19)30261-5/fulltext		Netherlands
7	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	No	REPEAT
8	Essay	Socio Economic Inequalities: Health	https://www.ukessays.com/essays/health-and-social-care/socio-economic-inequalities-health-health-and-social-care-essay.php	No	Health inequalities
9	Ryden	Diet cost, diet quality and socio-economic position: how are they related and what contributes to differences in diet costs?	https://www.cambridge.org/core/journals/public-health-nutrition/article/diet-cost-diet-quality-and-socioeconomic-position-how-are-they-related-and-what-contributes-to-differences-in-diet-costs/8261829C0654575691C1C2DFAF31CB46	No	About Sweden
10	Cookson	Socio-Economic Inequalities in Health Care in England	https://onlinelibrary.wiley.com/doi/full/10.1111/j.1475-5890.2016.12109	No	Health inequalities

9 Social Searcher using Bag of Words					
1	Tiffin	Inequalities in diet and nutrition	https://core.ac.uk/download/pdf/107964.pdf	Yes	
2	Roberts	Social and economic inequalities in diet and physical activity	https://researchportal.bath.ac.uk/en/publications/social-and-economic-inequalities-in-diet-and-physical-activity	Yes	
3	Dowler1	Food and health inequalities: the challenge for sustaining just consumption	https://www.tandfonline.com/doi/abs/10.1080/13549830802478736?mobileUi=0&journalCode=cloe20	No	No data
4	Dowler1	Food and health inequalities: the challenge for sustaining just consumption	https://www.tandfonline.com/doi/abs/10.1080/13549830802478736?mobileUi=0&journalCode=cloe20	No	No data
5	Alkerwi	Demographic and socioeconomic disparity in nutrition: application of a novel Correlated Component Regression approach	https://bmjopen.bmj.com/content/5/5/e006814	No	About Luxembourg
6	Dowler2	Inequalities in food and nutrition: challenging 'lifestyles	https://www.researchgate.net/publication/254432822_Inequalities_in_food_and_nutrition_challenging_	No	No data
7	Mutch	PUBH7620 Social Perspectives in Public Health - Week #4 :	https://www.facebook.com/notes/peter-nisbet/pubh7620-social-perspectives-in-public-health-week-4-/10154248569895588/	No	Health inequalities
8		Healthy and environmentally sustainable diets; how easy are they to define and to achieve?	https://z-m-www.facebook.com/events/1181797338669671/	No	Diet and sustainability
9		The food we eat	https://www.facebook.com/thehealthfoundation/photos/a.589197687811716/2210686555662813/?type=1&theater	No	No data
10	Cookson	Socio-Economic Inequalities in Health Care in England	https://onlinelibrary.wiley.com/doi/full/10.1111/j.1475-5890.2016.12109	No	About diet

10A StartPage using Boolean Search					
1	Gov	Poverty and income inequality statistics	https://www.gov.scot/collections/poverty-and-income-inequality-statistics/	no	Income inequalities
2	Barton	Time to address the double inequality of differences in dietary intake between Scotland and England.	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6088540/	no	Geographical comparison

3	Barton	Time to address the double inequality of differences in dietary intake between Scotland and England.	https://www.ncbi.nlm.nih.gov/pubmed/29947325	no	(repeat of irrelevant doc)
4	Ji	Socioeconomic inequality in salt intake in Britain 10 years after a national salt reduction programme	https://bmjopen.bmj.com/content/4/8/e005683	Yes	
5	F.Times	Income Inequality increases in UK	https://www.ft.com/content/06c0933c-5ed8-11ea-b0ab-339c2307bcd4	No	Income inequalities
6	NHS	Food Poverty	http://www.healthscotland.scot/media/2222/food-poverty.pdf	No	No data
7	Guardian	Large rise in takeaway shops highlights dominance of fast food in deprived areas	https://www.theguardian.com/inequality/2017/jul/25/large-rise-takeaway-shops-highlights-dominance-fast-food-deprived-areas-england	No	No data
8	Barton	Time to address the double inequality of differences in dietary intake between Scotland and England.	https://rke.abertay.ac.uk/ws/portalfiles/portal/15022448/Barton_TimeToAddressTheDoubleInequalityOfDifferencesInDietaryIntake_Published_2018.pdf	No	(repeat of irrelevant doc)
9	UCL	Diet makes a difference to learning	http://www.educationalneuroscience.org.uk/resources/neuromyth-or-neurofact/diet-makes-a-difference-to-learning/	No	Psychology
10	Oxfam	Food Poverty	https://policy-practice.oxfam.org.uk/our-work/inequality/food-poverty	No	No data

10B StartPage using Bag of Words					
1	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pubmed/25399952	Yes	
2	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351901/	no	REPEAT
3	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	https://www.cambridge.org/core/journals/british-journal-of-nutrition/article/socioeconomic-dietary-inequalities-in-uk-adults-an-updated-picture-of-key-food-groups-and-nutrients-from-national-surveillance-data/9F814CB09FCC78ED8C4F1E1E1D80737B	no	REPEAT
4	Roberts	Social and economic inequalities in diet and physical activity	http://www.oneeastmidlands.org.uk/sites/default/files/library/vid_19253_Social_and_economic_inequalities_in_diet_and_physical_activity_04.11.13.pdf	Yes	
5	Tiffin	Inequalities in diet and nutrition	https://core.ac.uk/download/pdf/107964.pdf	yes	
6	Yau	Time trends in adherence to UK dietary recommendations and associated sociodemographic inequalities, 1986-2012: a repeated cross-sectional analysis	https://www.nature.com/articles/s41430-018-0347-z	yes	
7	Alkerwi	Demographic and socioeconomic disparity in nutrition: application of a novel Correlated Component Regression approach	https://bmjopen.bmj.com/content/5/5/e006814	No	About Luxembourg
8	PHE	New briefing paper explores	https://www.gov.uk/government/news/	No	REPEAT (of

		inequalities in diet and physical activity	new-briefing-paper-explores-inequalities-in-diet-and-physical-activity		Roberts above)
9	Dowler1	Food and health inequalities: the challenge for sustaining just consumption	https://www.tandfonline.com/doi/abs/10.1080/13549830802478736?mobileUi=0&journalCode=cloe20	No	No data
10	Roberts	Social and economic inequalities in diet and physical activity	https://researchportal.bath.ac.uk/en/publications/social-and-economic-inequalities-in-diet-and-physical-activity	No	REPEAT

11 Web of Science using Boolean Search					
1	Berger	Changes in the sugar content of food purchases and socio-economic inequalities: a longitudinal study of British households, 2014-2017	Journal of Epidemiology and Community Health	Yes	
2	Reynolds	Healthy and sustainable diets that meet greenhouse gas emission reduction targets and are affordable for different income groups in the UK	Public Health Nutrition	No	About sustainability
3	Cornelsen	Changes in the sugar content of food purchases and socio-economic inequalities: a longitudinal study of British households, 2014-2017	Journal of Epidemiology and Community Health	Yes	
4	Tong	Dietary cost associated with adherence to the Mediterranean diet, and its variation by socio-economic factors in the UK Fenland Study	British Journal of Nutrition	No	About cost of diet
5	Maguire2	Does exposure to the food environment differ by socioeconomic position? Comparing area-based and person-centred metrics in the Fenland Study, UK	International Journal of Health Geographics	No	About retail food environments
6	Yau	Age, sex and socioeconomic inequalities in fruit and vegetable intake in UK adults, 1986-2012	Journal of Epidemiology and Community Health	Yes	
7	McNeill	Socio-economic differences in diet, physical activity and leisure-time screen use among Scottish children in 2006 and 2010: are we closing the gap?	Public Health Nutrition	Yes	
8	Albar	Does food portion size differ by level of household income? A cross-sectional study using the UK National Diet and Nutrition Survey 2008-11	Lancet	Yes	
9	Barton	Trends in socio-economic inequalities in the Scottish diet: 2001-2009	Public Health Nutrition	Yes	
10	Maguire	Socio-economic dietary inequalities in UK adults: an updated picture of key food groups and nutrients from national surveillance data	British Journal of Nutrition	Yes	