

# cassi\_2017\_improving\_fitness\_mapping\_research\_priorities\_against\_societal\_needs\_on\_obesity

## Year

2017

## Author(s)

Lorenzo Cassi and Agenor Lahatte and Ismael Rafols and Pierre Sautier and Elisabeth de Turckheim

## Title

Improving fitness: Mapping research priorities against societal needs on obesity

## Venue

Journal of Informetrics

---

## Topic labeling

Fully automated

## Focus

Secondary

## Type of contribution

Established approach

## Underlying technique

Probability-based method (LDavis relevance threshold)

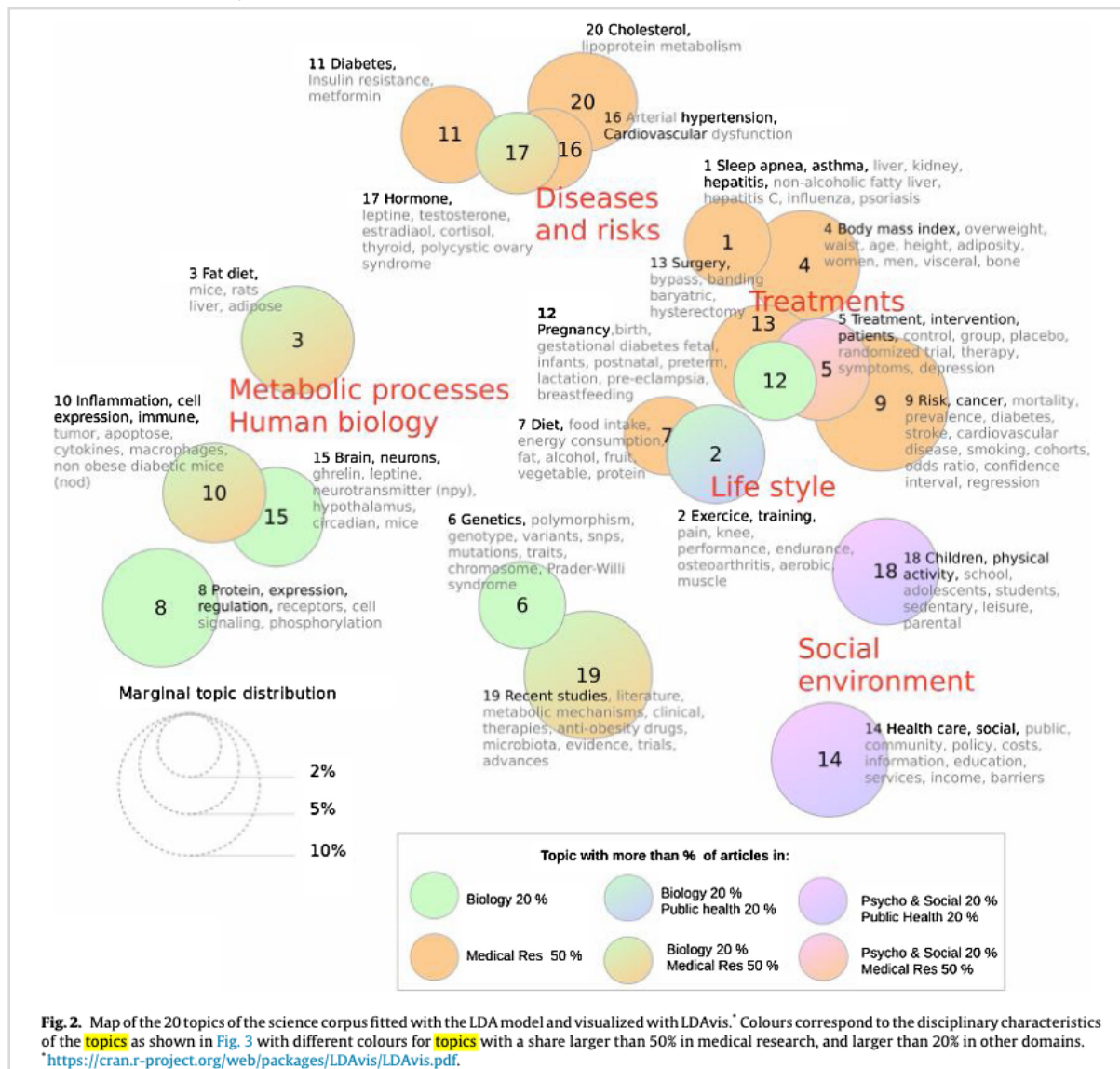
## Topic labeling parameters

Relevance parameter (lambda): 0.6

## Label generation

We analyse and display the results with R scripts and interactive visualisation provided by the LDAvis program. On a LDAvis figure, topics are close if they have similar term distribution.

Labels are obtained by selecting a few meaningful terms among the 20 most frequent (or relevant) terms displayed by LDAvis



## Motivation

\

## Topic modeling

LDA

## Topic modeling parameters

Nr of topics (k): {20, 200}

## Nr. of topics

20

---

## Label

One or more words belonging to the topic (selected using an LDAvis relevance threshold)

## Label selection

\

## Label quality evaluation

\

## Assessors

\

---

## Domain

Paper: Health (obesity)

Dataset: Medical domain

## Problem statement

In this paper, we present an exploratory investigation of science supply and societal needs on the grand challenge of obesity.

We illustrate a potential approach that uses topic modelling to explore:

- (a) how scientific publications can be used to describe existing priorities in science production;
- (b) how policy records (in this case here questions posed in the European parliament) can be used as an instance of mapping discourse of social needs;
- (c) how the comparison between the two may show (mis)alignments between societal concerns and scientific outputs.

## Corpus

### Science supply data

Origin: PubMed, Web of Science database

Nr. of documents: 147,322

Details:

- Publications matching the search obes\* in MEDLINE/PubMed during the 2000-2013 period matched with records indexed in the ISI Web of Science (WoS) database
- 'article' or 'review' as document types
- The same search is then directly on the WoS Core Collection and the corpus is defined as the union of these two sets

### Societal demands data

Origin: European Parliament

Nr. of documents: 222

Details:

- questions from Members of European Parliament (MEPs) to the European Commission
- We retrieved all the questions (oral or written) asked during the Seventh Parliament term (2009-2014) and the beginning of the current term (until February 15, 2015) which contain the word obes\* in the text

## Document

- 'article' or 'review' from published in PubMed and present in the Web of Science database
- Questions as short texts (30-400 words), pointing to either the general issue of obesity or to particular causes or consequences of obesity.

## Pre-processing

We use a classification system to identify clusters of publications related to obesity.

We enriched the original corpus with the whole clusters having at least a percentage of publications tagged as obesity by the WoS query.

Documents without an abstract were removed from these sets. This step reduces the original corpus of 147,322 documents to a corpus denoted O of 133,731 documents.

Removal of standard stop words and terms appearing in less than 100 documents

---

```
@article{cassi_2017_improving_fitness_mapping_research_priorities_against_societal_needs_on_obesity,
```

```
  abstract = {Science policy is increasingly shifting towards an emphasis in societal problems or grand challenges. As a result, new evaluative tools are needed to help assess not only the knowledge production side of research programmes or organisations, but also the articulation of research agendas with societal needs. In this paper, we present an exploratory investigation of science supply and societal needs on the grand challenge of obesity -- an emerging health problem with enormous social costs. We illustrate a potential approach that uses topic modelling to explore: (a) how scientific publications can be used to describe existing priorities in science production; (b) how policy records (in this case here questions posed in the European parliament) can be used as an instance of mapping discourse of social needs; (c) how the comparison between the two may show (mis)alignments between societal concerns and scientific outputs. While this is a technical exercise, we propose that this type of mapping methods can be useful to domain experts for informing strategic planning and evaluation in funding agencies.},
```

```
  author = {Lorenzo Cassi and Agnor Lahatte and Ismael Rafols and Pierre Sautier and Elisabeth {de Turckheim}},
```

```
  date-added = {2023-03-16 19:41:56 +0100},
```

```
  date-modified = {2023-03-16 19:41:56 +0100},
```

```
  doi = {https://doi.org/10.1016/j.joi.2017.09.010},
```

```
  issn = {1751-1577},
```

```
  journal = {Journal of Informetrics},
```

```
  keywords = {Research agenda, Science mapping, Societal needs, Obesity, Topic modeling},
```

```
  number = {4},
```

```
  pages = {1095-1113},
```

```
  title = {Improving fitness: Mapping research priorities against societal
```

```
needs on obesity},
```

```
url = {https://www.sciencedirect.com/science/article/pii/S1751157717301542},
```

```
volume = {11},
```

```
year = {2017}}
```

#Thesis/Papers/Initial