P-08-09

The *Meat Science* Universe: Meta- and topic analytical text analysis of a scientific journal (after 40 years) (#135)

Elisa Oertel[1], Stefan Baur[2], Daniel Moerlein[1]

[1] Georg-August-University, Department of Animal Sciences, Göttingen, Germany; [2] Gesellschaft für wissenschaftliche Datenverarbeitung (GWDG), eScience, Göttingen, Germany

Introduction

Scientific journals provide empirically established findings to researchers of the respective fields and it is up to each individual to keep track. To do this, scientists use online subscription-based databases such as *Web of Science* or free search engines like *PubMed*. However, the latter lacks a graphical platform that can answer meta-analytical questions at a glance (such as experts in a particular sub-area, countries with the largest research share, evolving topics and the like). These tools would enable both, the public and the science community to easily get started in the abundance of scientific publications. In this research, we have been evaluating computer assisted text analysis to analyse all publicly available abstract and meta information of the journal *Meat Science* as an example.

Methods

All article abstracts ever published in *Meat Science* (ranging from the oldest from 1977 to the last dated 21.02.2019) were downloaded and stored locally *via* the *PubMed* backend interface *EDirect* [1]. On the resulting XML-Files, we carried out an exploratory analysis in **R** [2], working out information like publication strength for all participating countries and their cooperations. Subsequently, we split up the data regarding investigated species and filtered for the top authors as shown in Figure 2. To delve deeper into the data, we constructed a model of the topic distribution covered by this journal using the Latent Dirichlet Allocation algorithm [3]. Afterwards we analyzed topic development per investigated species over time to reveal evolving and decreasing methods/topics in these sectors.

Results

We were able to show that the top countries publishing in *Meat Science* are the United States and Spain, with more than 1000 publications overall, followed by Australia, China, Italy, France and the United Kingdom with about 200-400 publications. Although many countries have already worked in cooperation (70 out of 85 countries), the total number of joint publications is very low (609 out of 7399 publications so far). Also, more than one third of the journals editors have a high publication score, locating them between the top 100 authors. Additionally they tend to show up in more than one publication group (here: cattle&beef, pig&pork, sheep&lamb and poultry&chicken), which can be seen in the expert groups per investigated species in Figure 2. Finally, the topic analysis of, for example, the entire *Meat Science* journal as shown in Figure 3 revealed a growing interest in topics 1 and 3 (comprised of words like: stunning, rigor, boar, proteins, models and prediction) and a decreasing interest in topics 5 and 6 (comprised of words like: consumer, fat, diet, vitamin).

Conclusion

Our approach allowed us to map a nearly complete meta-level image of the *Meat Science* journal: Starting with important countries for the contributions of the journal up to the top authors in the field (implemented for first and last authors as well as for the entire authorship), we conducted a predictive analysis of ongoing topics in the research field (such as protein-analysis and electrical stimulation), by developing a Text Mining and Topic Modeling framework.

Literature

[1] Kans J. Entrez Direct: E-utilities on the UNIX Command Line. 2013 Apr 23 [Updated 2019 Feb 11]. In: Entrez Programming Utilities Help [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2010-.

[2] R Development Core Team. "R: A language and environment for statistical computing." R Foundation for Statistical Computing.Vienna, 2008.

[3] Wang, Chong, and David M. Blei. "Collaborative topic modeling for recommending scientific articles." Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011.

Notes

Figure 2: Experts in Meat Science Grouped barplots showing the top 25 authors per species (author position ignored), illustrating the bandwidth of Meat Science and the respective experts in the given field. The publications are shown in absolute count and are not necessarily disjoint between groups.

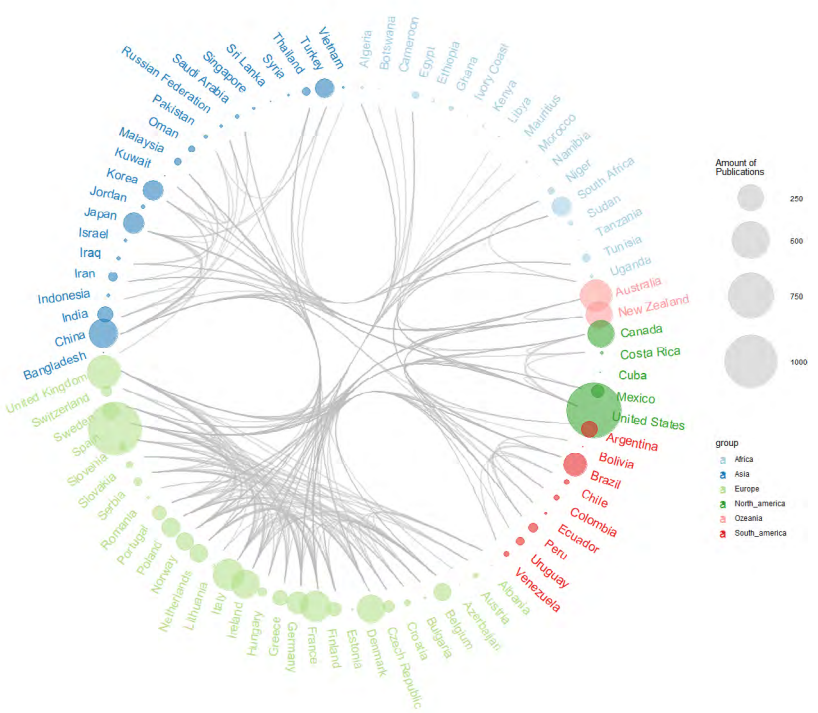

Figure 1: Explorative analysis of Meat Science A circelized dendrogram of all countries contributing publications - Countries are sorted by continent, where bubble sizes indicate the publication strength and grey lines refer to cooperation between countries.

Notes

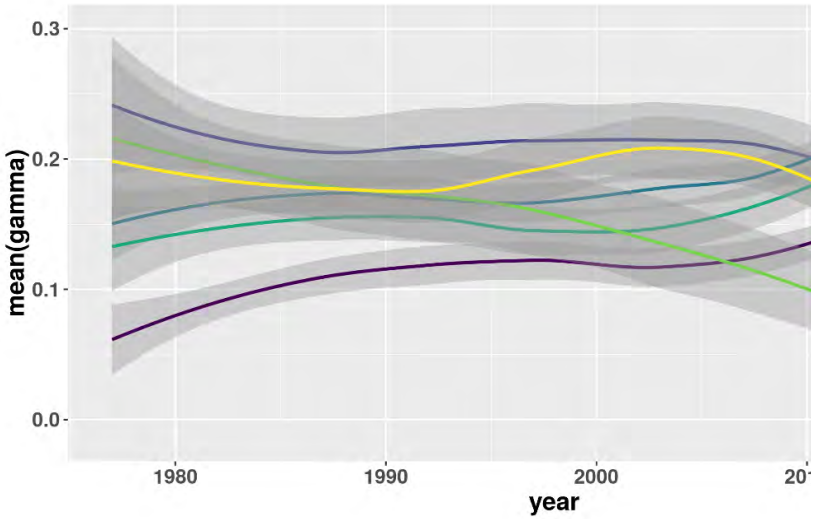

Figure 3: Topic Development in whole Meat Science The distribution of six calculated topics over scientific publications in Meat Science pictured as a timeline. The x-axis displays the year of publication. Gamma denotes the most probable distribution of topics over all documents (it was averaged to facilitate interpretation). The top three terms per topic are shown next to the respective calculated topic in the legend.

Notes