World Patent Information 55 (2018) 27–36

Contents lists available at ScienceDirect

World Patent Information

journal homepage: www.elsevier.com/locate/worpatin

# Patenting activity in the food safety sector

Yeong Jae Kim[a,*], Kaye Husbands Fealing[b], Evgeny Klochikhin[c]

[a] Tyndall Centre for Climate Change Research, University of East Anglia, Norwich, NR4 7TJ, UK
[b] Georgia Institute of Technology, 685 Cherry Street, Atlanta, GA, 30332-0345, USA
[c] Parkofon, Alexandria, VA, United States

ARTICLE INFO

*Keywords:*
Food safety
Patent
Machine learning
Research taxonomy

ABSTRACT

Research on science and technology policy has heavily relied on patent data. However, relatively few studies of food safety patent activity appear in scholarly literature. This paper provides a discussion on patents as a measure of new knowledge generation in the food safety sector. In so doing, there are inherent challenges to identifying a research taxonomy for this multidisciplinary area. To overcome these challenges, the paper uses a natural language approach that can be applied to other research areas where boundaries of fields are not well defined.

## 1. Introduction

Food safety is a national priority in the United States and around the world. In a 2010 report [1], the Centers for Disease Control and Prevention stated that one in every six people in the United States gets sick from foodborne illness, 128,000 cases of foodborne illness require medical treatment, and approximately 3000 people die every year in the United States from foodborne illness. Outbreaks of foodborne illnesses occur with surprising frequency and more than $2 billion are spent annually on food-safety research and development (R&D) at the U.S. Department of Agriculture (USDA) [National Institute of Food and Agriculture (NIFA), Agricultural Research Service (ARS) and Economic Research Service (ERS)], and the U.S. Department of Health and Human Services [Food and Drug Administration (FDA) and the National Institutes of Health (NIH)]. Other federal agencies, such as the National Science Foundation (NSF), sponsor research that informs biological solutions and practices in the food-safety sector. Health outcomes are typically the focus of studies on impacts related to investments in research and development (R&D) related to food safety. But preceding those outcomes are outputs, such as human capital produced during training on research projects (e.g., graduate students), papers published on findings from the research, and patents granted to protect the intellectual property embodied in products and processes produced as a result of the research. It is this latter output—patents—that we seek to examine in this paper.

The scope of food-safety research spans from farm-to-fork. Husbands Fealing et al. [2] [1] discuss ways in which the impact of food-safety research is evident throughout the entire supply chain of food production and distribution: agricultural inputs, pre-harvest environmental factors, harvest-related and postharvest factors, manufacturing techniques, storage and transportation conditions, food-processing factors, retail and consumer handling, and surveillance systems. Food-safety research includes all stages of research, including basic, translational, applied, and data acquisition (e.g., environmental and food sampling). Therefore, evaluating the impact of federal funding on food-safety research requires examining the full span of food safety activities (farm-to-fork) and research at all stages of exploration.

One challenge faced when investigating the relationship between funding of food-safety research and outputs of that funding is the development of a taxonomy that defines food safety. A multidisciplinary area, food safety is difficult to define using traditional methods. The existing scientific taxonomy does not provide a comprehensive definition of food safety that includes multiple scientific domains, levels of examination, and industry sectors. Merely looking up food safety in, for example, the North American Industry Classification System codes does not yield a complete list of sectors comprising food safety.

Another challenge is that patents are not the primary currency of food-safety research. Based on the literature review, we did not find a sizable corpus of literature on food safety patents. Food scientists[2] who

* Corresponding author.
*E-mail addresses:* y.kim@uea.ac.uk (Y.J. Kim), khf@gatech.edu (K. Husbands Fealing), contact@parkofon.com (E. Klochikhin).

[1] The scope of food-safety research is a well-illustrated in Figure 2.1(page 13) [1].

[2] A dozen food-safety experts attended the December 2015 workshop sponsored by the research team and funded by the USDA-NIFA. Two participants are also co-authors of chapter 2 of Husbands Fealing et al.: Lee-Ann Jaykus is a William Neal Reynolds Distinguished Professor in the Department of Food, Bioprocessing & Nutrition Sciences at North Carolina State University; and Laurian Unnevehr is Professor Emerita in the Department of Agricultural and Consumer Economics at the University of Illinois.

https://doi.org/10.1016/j.wpi.2018.08.005
Received 8 February 2018; Received in revised form 13 August 2018; Accepted 27 August 2018
0172-2190/ 

participated in a workshop sponsored by the research team acknowledge that outputs of their research are public goods—that is, a product or process that is not necessarily developed for private benefit. Therefore, a focus on patents underestimates the full benefit to society of food-safety research, since it is more important to get a new product or process to market to save lives than it is to delay distribution owing to the patenting process [2](p.145).

Fanfani, Lanini, and Torroni [3] showed that patents related to agriculture and food industries in Italy are a weak indicator of food innovation. They stated that it is important to consider commercialization that is not a result of patents. Therefore, although patent data are widely used as a measure of innovation in some manufacturing sectors [4–8], more recent literature shows that there is not necessarily a strong correlation between patenting and innovation [9]. For this reason, using only patent data to measure food safety innovation can be misleading. A patent is not a perfect measure of innovation, since not all commercialized products or processes are patented especially in food-safety sectors.

There is anecdotal evidence that the food safety innovation was largely driven by both private and public sector funding on Hazard Analysis and Critical Control Points (HACCP) systems to control pathogens for the U.S. meat industry [10]. On one hand, private companies play an important role in inducing agricultural biotechnology innovation [11]. On the other hand, agricultural biotechnology patenting heavily relies on public research funding [12]. However, the impact of public funding may be realized for some time in the food safety sector similar to the low-carbon technology sector [13].

Although patent data are not a perfect measure of food safety innovation, there are several research papers that use patents as a proxy of the subfield of agriculture. For example, one study found that innovators are getting clustered in the agriculture, water, food, and bioenergy innovation ecosystem in Colorado using patent data [14]. King and Schimmelpfennig [15] also relied on patents from the USDA-ERS and the Agricultural Biotechnology Intellectual Property Database to investigate the quantity, quality, and composition of agricultural biotechnology intellectual property rights of the major agricultural biotechnology firms and their subsidiaries: Dow, DuPont, Monsanto, BASF, Bayer, and Syngenta. While this is the most comprehensive report on agricultural biotechnology innovation in general, their paper does not specifically focus on food safety patent activity.

There is also literature on seed industry and intellectual property rights owing to tremendous industry consolidation in the agricultural sector [16,17] and evolving roles of intellectual property protection rights in the agricultural biosciences [18–20]. Salay, Caswell, and Roberts conducted a survey for case studies of food safety innovation, but their taxonomy of food safety was not fully specified [21].

This paper, therefore, contributes to the literature by showing how machine learning techniques can be used to develop a taxonomy on food safety and to identify food safety patents. Those identified food-safety patents are further examined to address three questions: (1) How are food-safety patents classified? (2) Which firms are actively participating in food safety patenting? (3) What are the geographical and sectoral distributions of food safety patenting? The paper is organized as follows. First, we discuss the methodological background. Second, this paper describes new data and methods used to define food-safety research, which can be further applied to other multidisciplinary sectors. Third, we validate our results. Fourth, we analyse results and then conclude.

## 2. Methodological background

In this paper, we have two methodological contributions. The first methodological contribution is the application of text analysis techniques, using Wikilabeling to establish the taxonomy, which we then used to discover food-safety patents [22]. This technique is described in chapter four of the Husbands Fealing et al. [2]. Information retrieval and identification using Wikilabeling determines a group of topics based on words in documents. This process generated a list of topics within a corpus. Similarities between individual documents, such as government awards and Wikipedia webpages, were matched using the following method:

1) determine if a standalone Wikipedia article exists within the list of significant n-grams from within the corpus and an existing taxonomy;
2) evaluate the similarities between individual documents and Wikipedia webpages; and
3) identify keywords and phrases that represent the food safety sector.

The model was trained on a database of grant abstracts from NIH, NSF and USDA. The primary advantage of applying Wikilabeling is that it allowed us to derive a list of potential labels from the corpus that reflected the existing taxonomy, for example, NSF's Survey of R&D Expenditures at University and Colleges. Therefore, Wikilabeling enabled us to update and extend the existing research taxonomy.

The second key methodological contribution is the use of the U.S. Patent and Trademark Office's (USPTO) PatentsView database. This database is used to identify food safety patents and to retrieve additional data about patent assignees, inventors, their locations, and patent classifications. The most significant advantage of using the PatentsView database is accuracy of the disambiguated assignee, inventor, their locations, and patent classifications [23]. PatentsView uses a patent assignee disambiguation technique,[3] the Jaro-Winkler approach, to cluster entities. Of course, a certain amount of manual check is inevitable. Additionally, the same John M. Smith might apply for two patents with and without the middle initial. If one were looking at exact matches, then these two inventors would be considered different individuals while in fact, they reside in the same city, the patent is in the same technology area, they work for the same company, and so on. The new inventor disambiguation algorithm, authored by the research team from the University of Massachusetts at Amherst and integrated into PatentsView in 2016, uses discriminative hierarchical co-reference as a new approach to increase the quality of inventor disambiguation [24,25]. For locations—city/state/country text as it appears in source files—area algorithmically matched against a master geocode file from Google and MaxMind open source files.

## 3. Methods

We applied the keywords used in searching food-safety research based on the search string approach referenced in Husbands Fealing et al. [2] (p. 170). A three-stage process was used to extract the final search strings needed to identify food safety patents. Fig. 1 summarized this approach graphically.

(1) Combine two advanced techniques—search string approach and Wikilabeling—to identify possible food-safety research.
(2) Validate the initial sets through expert curation. Using this finalized food safety search strings (shown in the appendix) and patent classifications, retrieve the relevant food safety patents.
(3) Validate the results, using query-side and retrieval-side methods.

Patent documents are more complex than award abstracts owing to the legal language characteristics that do not necessarily show the nature of patent content in lay terms [26]. Therefore, we used a combination of both text analysis and patent technology classifications to identify food safety patents. Additionally, we manually validated food safety patents to reduce Type I (false positive) and Type II (false

[3] https://www.uspto.gov/about-us/organizational-offices/office-policy-and-international-affairs/patentsview-inventor.