

## Abstract

# Mapping Gene Patents: Text Mining Enabled Legal Analysis of Existing Property Rights in Natural & Synthetic Genes<sup>1</sup>

---

Kimani .F. Goddard, Maastricht University Faculty of Law, PhD Fellow  
Pedro Hernandez Serrano, Maastricht University Institute of Data Science

The core proposition of this paper is that the increasing complexity of patent thickets in genetics, coupled with rapid technological advancement in this field, requires that patent decisions should be grounded in precise empirical evidence. However, the sheer volume of patent data presents practical difficulties for precise legal examination of gene patent applications and challenges to existing patents. We propose that text similarity algorithms and Natural Language Processing (NLP) provide the possibility for precision, as a means to create tools that allow patterns and relationships within patent disclosure data to be identified, in order to produce insights which could be useful for enhancing the standard of review which decides outcomes of patent examinations and disputes.

Application of NLP analytical methods in a legal context is crucial, as the scope of patents issued for natural and synthetic gene sequences continues to be a source of uncertainty and ethical unease. This is fuelled by the fact that gene editing and isolation of gene sequences, have rendered human manipulation of biological products and processes more efficient and more ambitious, within the last 10 years. In 2018 alone, the EPO recorded a 12.1% increase in biotechnology patent applications, with an additional 2.1% increase recorded in 2019, the following year. Despite this, in May 2020 the unease with this expansion of biotechnology property rights, manifested itself through the EBA opinion on G 0003/19, in which it expanded the boundaries of exclusions from patentability, based on EPC Art 53(b) which renders “plant or animal varieties or essentially biological processes for the production of plants or animals...” patent in-eligible. The EBA tailored its decision to integrate developments which had transpired since the pronouncements in G 2/12 and G 2/13, setting aside previous interpretations of Art. 53(b), it confirmed that Rule 28(2) is not incompatible with the Article. Further, that the exclusionary provision should be interpreted dynamically, so as to encompass later emerging “aspects or developments which were unknown at the time the decision was issued, or irrelevant to the case, or otherwise not taken into consideration.”<sup>2</sup>

---

<sup>1</sup> This paper is part of a research collaboration between Maastricht University Faculty of Law and the Maastricht University Institute of Data Science.

<sup>2</sup> Opinion EPO - EBA: G 0003/19 Pepper (pursuant to Tomatoes II & Broccoli II)