



**The Power of the Pen, the Press, and the Algorithm:  
Evaluating the Framing of Google News Regulatory  
Efforts in the European Union and Australia**

by

**Shannon Ors**

**Thesis Advisor: Professor Lynda Oswald**

A thesis submitted in fulfillment of the requirements of the  
Michigan Ross Senior Thesis Seminar (BA 480), 22 April 2021.

## Acknowledgements

I would like to take a moment to thank all those who helped this thesis become a reality. In particular, thank you Professor LaFontaine and Professor Tasoluk for guiding me through the research process and always being so willing to share your wisdom with the class. It was an honor to learn from such accomplished researchers and I truly appreciate all the energy you invested in my success. I also owe my deepest appreciation to Professor Oswald who so willingly agreed to serve as my thesis advisor back in September. Thank you for always responding to my emails, offering incredibly detailed feedback on my writing, and introducing me to intellectual property law. Without your guidance, this thesis would not have been possible. Thank you to Rohit Singla at Ross Research Computing who encouraged me to learn Python for this thesis and fixed many bugs in my code. A special thank you is also deserved to our PhD student advisor, Paul Organ, who calmed my nerves about approaching potential faculty advisors and encouraged me to go with my gut and pursue this topic. Additionally, thank you to my peers in BA 480 who kept me sane and made even the stressful moments a little more fun. Huge shoutout to Student Interest Group #1 — Maansi Dalmia and Riya Gupta — I don't know what I would have done without your support, especially amidst the loneliness of remote school. And of course, thank you to Anna Zidarescu for being the best housemate a girl could ask for. You are truly an amazing friend for listening to me worry about this thesis for the past eight months. Lastly, thank you to my parents and my sister for being my biggest cheerleaders — love you all so much.

## **Abstract**

This thesis aims to compare and contrast stakeholder positions to proposed news aggregator regulations in the European Union (EU) and Australia (AU). The EU and AU have both sought to apply regulatory pressure to force Google News to remunerate press publishers for aggregating their news content. The research setting of the EU and AU was selected because the two governments have differing conceptions of copyright stemming from their respective civil law and common law legal traditions. Thus, this thesis hypothesizes that the EU approach of protecting the creator itself versus the AU approach of defending the entire public domain of news content will be reflected in consultation submissions. In other words, this thesis will identify the frames or “specific aspect[s] of a policy proposal that [are] emphasized in a policy debate by a specific actor” in the EU versus AU consultation submissions (Entman, 1991). While there is existing literature that empirically demonstrates the benefits and threats of news aggregators as well as a body of literature advocating for a re-framing of the news aggregator debate, no studies have systematically extracted the frames put forth by stakeholders. Thus, by performing a quantitative text analysis of EU and AU stakeholder consultation submissions (n = 243), this thesis will offer a systematic mapping of current frames and draw connections to both jurisdictions’ legal conceptions of copyright. The results generated by this analysis will aim to inform future stakeholder framing already being considered in other news markets around the globe, including the United States.

<b>Contents</b>	
<b>Introduction</b>	5
<b>Statement and Justification of Problem</b>	7
<b>Background</b>	10
What is a News Aggregator?	10
History of Google News	10
Early Google News Litigation and Regulation: 2005 - 2014	12
EU Directive on Copyright in the Digital Single Market	13
ACCC Draft Mandatory Bargaining Code for News Media and Digital Platforms	14
<b>Literature Review</b>	15
Do News Aggregators Compete with Press Publishers?	15
Is Google News a Substitute or Complement to News Outlets?	16
Does Google News Benefit or Harm Press Publishers?	17
Is There a Better Way Forward for Media Businesses?	18
<b>Theoretical Framework</b>	19
Framing Theory	19
Conception of Copyright: Civil Law versus Common Law	20
<b>Methodology</b>	23
Data Collection	23
Data Preparation	25
Question 1: Measuring Textual Similarity	26
Question 2: Topic Modeling	27
Question 3: Manual Text Analysis	28
Limitations	28
<b>Results</b>	30
Summary Statistics	30
Question 1: Measuring Textual Similarity	34
Question 2: Topic Modeling	35
Question 3: Manual Text Analysis	36
<b>Discussion</b>	46
<b>Appendices</b>	51
<b>References</b>	59

## Introduction

In the early aughts, Larry Page and Sergey Brin found themselves in Rupert Murdoch's inner circle. It was an unlikely pairing — the co-founders of Google and the Executive Chairman of News Corp — as the two businesses not only represented competing ways to access information, but also markedly different management styles. Yet, Murdoch was curious. While hosting Page and Brin at his ranch in California, Murdoch bombarded the two brains behind the search engine with questions. Murdoch's third wife, Wendi Deng Murdoch, paraphrased one question in particular that fittingly captured a problem Murdoch would grapple with for the rest of his career, "why you no read newspapers? (Wolff & Wolff, 2008). The public record doesn't confirm how Page and Brin responded, but it didn't take long for Murdoch to identify his own answer to that question — Google News. By 2009, Murdoch would pivot to launch an aggressive crusade against Google that would span across the globe, most notably in the three major markets of News Corp's newspaper portfolio — the United States, the United Kingdom, and Australia.

Murdoch's argument centered on the perceived threat of news aggregators to the business model of news organizations. As he explained to a group of U.S. Federal Trade Commission regulators, "our friends online...think they have a right to take our news content and use it for their own purposes without contributing a penny to its production...to be impolite, it's theft" (Murdoch, 2009). Yet, Google offers a different argument, emphasizing the consumer benefit of news aggregation and "virtual newsagent" role that drives traffic to online news sites (Tryhorn, 2009). Eleven years later, the stakeholders involved the news aggregator policy debate have expanded well beyond Murdoch probing Page and Brin on his California ranch. Yet, these two competing arguments have guided research across disciplines to understand the legality, business implications, and economic incentives attached with requiring Google News to pay press publishers for their news content. At the crux of this research, is the question succinctly distilled by one research team

— new rights or new business models? (Senftleben et al., 2017). In other words, should the government intervene to provide press publishers with the right to receive remuneration from aggregators who use their content, or should it be the responsibility of press publishers to innovate with business models that will thrive in the digital world?

Subsequently, this thesis is concerned with understanding how the competing perspectives put forth by news aggregators, press publishers, and other peripheral stakeholders interact within the context of the news aggregator policy debate in the European Union (EU) and Australia (AU). This will be accomplished via a text analysis of the consultations administered by the two governments during the policy formation stages of the debate. Embedded within this analysis is the distinction between the civil law and common law understanding of copyright that ground my hypothesis of how stakeholder responses between the two countries will differ. This analysis is informed by framing theory which captures the idea that there is meaning in what a stakeholder chooses to emphasize and ignore in their consultation submission. Framing theory will be utilized to organize and group the outputs of the text analysis. Thus, this thesis is divided into three parts. First, the text of the submissions will be converted into a vector space to calculate the similarity of consultation submissions. Second, topic modeling will be performed to extract frames present in the EU and AU consultations. A frame is defined by Entman (1991) as a “specific aspect of a policy proposal that is emphasized in a policy debate by a specific actor.” Third, the results of the topic modeling will be verified via manual text analysis to understand whether the identified frames are unique or duplicated across the civil law legal system of the EU and common law legal system of AU. These results will be coupled with a qualitative analysis of the policy aftermath to gauge the success of different frames and better understand which frames may prove to be most useful as other countries pursue news aggregator regulation in the future.

## Statement and Justification of Problem

There has been substantial global interest in regulating news aggregators. The digital disruption to the newspaper industry has motivated governments to pursue an aggressive regulatory agenda to address the perceived market power imbalance between press publishers and news aggregators. While proposed and enacted policies all share the core goal of making news aggregators pay for the use of news content, the scope of these efforts vary significantly across governments. Subsequently, this thesis aims to explore how news industry stakeholders inform different policy approaches within the research setting of the EU and AU. Since the EU has a civil law legal system and the AU has a common law legal system, the two countries have different conceptions of copyright and thus offer an ideal framework for contrasting differing regulatory approaches. The scope of this thesis will be further narrowed to one news aggregator — Google News. This decision to focus the stakeholder narrative on Google News is prompted by Google News being the most visited news aggregator in the world (Newman, 2020). Moreover, Google News has been the explicit target of both EU and AU regulators as elaborated upon later.

The objective of this thesis is to compare and contrast stakeholder responses to proposed EU and AU news aggregator regulations — Article 15 of the “EU Directive on Copyright in the Digital Single Market” and “ACCC Draft Mandatory Bargaining Code for News Media and Digital Platforms” respectively. While both policies aim to require news aggregators to compensate press publishers for aggregating their news content, the EU policy confers a “press publishers’ right” whereas the AU policy pursues a competition law-based approach via bargaining rights. This thesis proposes that the civil law conception of copyright in the EU and the common law conception in the AU will translate to how the news industry stakeholder narrative is framed in the EU and AU. In broad terms, civil law jurisdictions aim to protect “creators and their claims to ensure the authenticity of their works” whereas common law jurisdictions focus on “the audience and its hopes



for an expansive public domain” (Baldwin, 2014). Thus, I hypothesize that the EU stakeholder narrative will emphasize creators (writers, journalists, authors) whereas the AU stakeholder narrative will emphasize access to the news content itself. To evaluate this hypothesis, I will perform a quantitative text analysis of consultations conducted by the EU and AU during the policy formation stages of their news aggregator regulations. These consultations collect responses from relevant stakeholders including press publishers, journalists, advocacy groups, collective management organizations, and Google. The methodology will be guided by three research questions:

- 1) How textually similar are the EU and AU consultation responses?
- 2) What frames do EU and AU consultation respondents employ in their submissions?
- 3) How do the frames identified in Question 2 appear in the consultations and inform the contrasting regulatory approaches of the EU and AU?

The first question will assess the textual similarity between each consultation submission. This metric captures how similar the vocabulary usage is across a set of documents by creating a vector space model, a method for representing the contents of a document numerically by counting the repetition of each word. The output of this method will be coordinates assigned to each consultation submission to plot each submission on a cluster map. In other words, this first question aims to uncover whether the EU and AU consultation responses share a similar vocabulary by evaluating the distance between each submission on a cluster map. Question 2 will further explore the cluster map by extracting the most repeated keywords from each cluster via topic modeling. These keywords will offer a starting point to uncover the frames employed by consultation respondents. Yet, recognizing a keyword does not necessarily equate to a frame, this method will be supplemented by manual text analysis of the submissions to verify and contextualize the results found in Question 2. And to answer the third question, the frames identified in Question 2 will be compared between the EU and AU. This analysis will be supplemented by qualitative research

regarding the policy debate aftermath to evaluate the success of each frame in civil law versus common law contexts.

The motivation for this thesis rests upon the fact that while the EU and AU were the first governments to pursue sweeping policy reform to address news aggregators, they will certainly not be the last. This is evidenced by similar regulatory efforts already being considered from India to Canada. While there has been significant research dedicated to the economic implications and legality of news aggregators, a gap in the literature exists that evaluates how the EU and AU news industry stakeholders frame their arguments. The answers to these three research questions can inform what argument framing resonates in common law versus civil law jurisdictions within the context of the news aggregator debate. Moving forward, these insights will be valuable to stakeholders in other markets who are currently evaluating the best framing approach for engaging in the news aggregator debate in their own country.

## **Background**

In order to establish the boundaries of this thesis, it is important to provide information regarding the definition of a news aggregator, the operating ethos of Google News, a brief history of prior attempts to regulate Google News, and descriptions of the EU and AU policymaking processes. Because regulation targeting Google News has existed from the onset of the aggregator itself, it is important to understand the context that fueled this increasingly divisive policy debate.

### ***What is a News Aggregator?***

This thesis will use Kimberley Isbell's (2010) definition of a news aggregator — “a website that takes information from multiple sources and displays it in a single place.” While the term news aggregator is often used interchangeably with news reader, feed reader, or RSS reader, the vocabulary of this thesis will be limited to news aggregator. Isbell groups news aggregators into four categories: feed aggregators, specialty aggregators, user-curated aggregators, and blog aggregators. As mentioned earlier, the scope of this thesis will be narrowed to Google News, which Isbell categorizes as a feed aggregator. Generally, the traditional conception of a news aggregator is a feed aggregator which indexes material from multiple websites and organizes content by source, topic, or story (Isbell, 2010). Outside the scope of this thesis are the remaining three categories of news aggregators identified by Isbell. Since Isbell's categorization of news aggregators, social media sites have launched features that perform similar functions to traditional feed aggregators such as Facebook News. Moreover, it could even be argued that social media platforms could be holistically categorized as user-generated aggregators. However, because this thesis focuses solely on Google News, these distinctions will not be explored further.

### ***History of Google News***

The following section details the aggregator's origins, algorithmic architecture, business model, and relationship with press publishers to color the landscape of consultation submissions.

The impetus for Google News occurred in the hours after 9/11. Head of Google Search, Amit Singhal, realized a search for “New York Twin Towers” yielded no relevant results because Google was indexing content from a month earlier. So, for the first time ever, Google placed links to *CNN* and *The Washington Post* reporting on its homepage with the message, “please visit these sites to get the news of the day, because our search is failing you” (*The Evolution of Search*, 2011). The beta version of Google News launched a year later on September 22, 2006. (Bharart, 2006). Today, Google News operates 70 country-specific editions and indexes articles across 80,000 news publications. While press publishers and even other aggregators such as Apple News rely on human editors to select and package articles, Google News is entirely automated (Sullivan, 2009). The aggregator runs on an algorithm separate from the search engine that has been overhauled repeatedly since 2002. Most notably in 2018, Google News began leveraging artificial intelligence and machine learning to deliver personalized news content to users — reinforcing the initial vision of Google’s co-founders to make editorial decisions without human editors (Upstill, 2018). Yet, despite these changes, the basic interface remains strikingly similar to the initial beta version displaying the headline, first few sentences of an article, and link to the original publication.

While the interface of Google News appears seemingly straightforward, there is a veil of mystery surrounding its algorithm and business model. A particular point of contention among publishers is the lack of specificity offered by Google regarding how the algorithm prioritizes which content it displays. Additionally, because the aggregator does not run any advertisements, the value Google derives from funneling readers to the search engine where searches do yield advertising revenue is unclear. The only time Google News has been publicly valued was in 2008 at Fortune’s Brainstorm Tech Conference where Google Vice President, Marissa Mayer, offhandedly quoted a \$100 million estimation (Fortt, 2008). This number has been the source of intense scrutiny as press publishers sought to blame Google News for declining subscription rates and the shuttering of local

newsrooms. While Google has responded by launching its Digital News Initiative in Europe, scrapping its controversial First Click Free policy that circumvented news paywalls, and offering publishers more user data to target subscribers — these concessions have seemingly not quelled press publisher’s complaints.

### ***Early Google News Litigation and Regulation: 2005 – 2014***

The early lawsuits involving Google News stemmed from press publishers suing Google for making their news articles available on the aggregator without their permission. The plaintiffs argued Google must obtain permission from press publishers (the copyright owners) prior to aggregating their news articles. Google responded by pointing out (1) any press publisher can “opt-out” from their content being aggregated, (2) the snippets of content displayed qualify as fair use, and (3) there is a public good benefit to expanding access to news content. The following section will survey the outcomes of early litigation against Google News.

In 2005, Agence-France-Presse (AFP), a newsgathering agency based in Paris, filed the first copyright infringement lawsuit against Google News (Cozens, 2005). Before the AFP case was resolved, Copipresse, a collective management organization (CMO) representing French and German-language newspapers in Belgium, filed a similar lawsuit against Google News (Crampton, 2007). The outcomes of both cases were not outright victories for press publishers. AFP settled with Google on undisclosed terms and Copipresse content was removed from the search engine index entirely after the court ordered Google to pay royalty payment to Copipresse (Auchard, 2007; Sterling, 2011). These early lawsuits proved a larger, more consolidated response would be needed if press publishers wanted to hold Google News accountable.

The shift from individual lawsuits to regulatory attempts can be attributed to Spain and Germany, who were the first to enact a press publisher right which “gives press publishers the exclusive right to authorize the reproduction and making available to the public of their press

publications for online uses” (Mullooly, 2020). In 2013, the German Parliament passed Section 87f-h of the “German Copyright Act 9” which introduced a press publisher right. The Spanish Competition Authority (CNMC) soon after amended the “Ley de Propiedad Intelectual,” introducing a similar press publisher right over news content. Yet, there was one important distinction — in Germany, the right was waivable, allowing press publishers to continue to be indexed by news aggregators, whereas the right was not waivable in Spain. Google News responded accordingly, switching to an opt-in policy in Germany and shutting down operations in Spain entirely (Román, 2014).

### ***EU Directive on Copyright in the Digital Single Market***

Following the German and Spanish regulatory efforts, the EU commission adopted the Digital Single Market (DSM) strategy in 2015 to address how digital technologies are changing how copyrighted content is produced, accessed, and distributed (European Commission, 2015). While the DSM signaled the EU embrace of a complete overhaul of the copyright laws governing the 28-member bloc, there was a particular focus on news aggregators, particularly Google News. This focus on Google was no secret. As Andrej Savin, an internet government professor at the Copenhagen Business School explained, “it’s a fact of life that most innovative companies come from the West Coast...the proposals coming from Brussels definitely have the American companies in mind” (Scott, 2016). Accordingly, the “Directive on Copyright in the Single Market” (CDSMD) was passed by the Council of the European Union on April 15, 2019 with a 19-6 vote (European Council, 2019). Of interest to this thesis is Article 15 (formerly Article 11) which grants a press publisher right in order “to strengthen [press publisher’s] bargaining position and improve their remuneration when they negotiate the use of their content by online platforms” (European Commission, 2018). The ambiguity of Article 15 was left to individual member states to resolve the

implementation issues of CDSMD. At this point, it is still unclear how the practical application of Article 15 will impact Google News.

### ***ACCC Draft Mandatory Bargaining Code for News Media and Digital Platforms***

Following the passage of the CDSMD in the EU, the Australian Competition and Consumer Commission (ACCC) began negotiating the “Draft Mandatory Bargaining Code for News Media and Digital Platforms” (DMBC). Currently, the proposed version of the DMBC, released on July 31, 2020, would only apply to Facebook and Google because both platforms are perceived to hold a significant bargaining power imbalance with Australian news media businesses (ACCC, 2020).

Unlike the press publisher right provided in the EU Directive, the DMBC takes a competition law approach by requiring Google to share revenue from news links with eligible media companies determined by the Australian Communication and Media Authority (ACMA). If Google cannot reach an agreement with press publishers on payment within three months of negotiation, it can move to arbitration overseen by the ACMA. The DMBC also empowers the ACCC to issue copyright infringement notices and fines for breach of the code. The DMBC was passed by the Australian Parliament on February 25, 2021, an effort led by federal treasurer Josh Frydenberg.

## Literature Review

Given news aggregators emerged barely two decades ago, academic research covering the topic is fairly nascent. However, literature does exist discussing the economic effects of news aggregators, including competitive dynamics and substitutability. There is also a body of literature concerned with evaluating the narrative of stakeholders within the media landscape and identifying opportunities to re-frame policy. Subsequently, this literature review will sample research related to these topics to help locate the contributions of this thesis.

### *Do News Aggregators Compete with Press Publishers?*

The question of whether news aggregators help or hurt press publishers is not settled. A handful of studies have administered surveys to address the question of whether a competitive relationship exists between press publishers and news aggregators. Before discussing these studies, it is important to be aware that there are two spheres of competition within media: inter-media and intra-media competition. Inter-media competition refers to competition between different forms of media such as online publications competing with traditional print media. Intra-media competition captures competition within an industry such as online news and is often characterized as the more intense form of competition because of high product substitutability (Picard, 1998; Chyi and Sylvie, 1998). The advent of news aggregators introduced new dynamics to both inter and intra-media competition, however most research concentrates on the intra-media competition between news aggregators and online news outlets.

Yang and Chyi (2011) analyzed reader survey data collected from 27 U.S. daily newspapers and found evidence of a non-competitive relationship between online newspaper sites and other news sites such as local and regional news, information sites, national news sites, and portal news sites. This initial study was built upon by Huang (2013) which extended Yang and Chyi's examination of intra-media competition to news aggregators in the online news market in Taiwan.



The study found online news sites do not exhibit a competitive relationship with one another. However, a statistically significant competitive relationship did exist between Yahoo! News (a news aggregator) and Apple Daily (a tabloid newspaper whose content was not available on the aggregator). The mixed results of these studies prompted Lee and Chyi (2015) to further investigate the relationship between news aggregators and other news media, controlling for key demographic and psychological factors. Their survey results found Google News, Yahoo News, and Huffington Post all had a negative correlation with other media outlets. This is the most convincing evidence so far that a non-competitive relationship exists between news aggregators and traditional news websites.

### ***Is Google News a Substitute or Complement to News Outlets?***

As governments began regulating Google News, a body of literature emerged using these markets as natural experiments to examine whether news aggregators are substitutes or complements to press publishers. News aggregators would be considered substitutes for traditional news consumption if they reduce the incentive for consumers to click article links and read the entire article on the press publisher's own website. In other words, if consumers are satisfied with the snippets of news content provided by a news aggregator, they would substitute that aggregator instead of visiting each press publisher's individual website. News aggregators would be considered complements if they drive traffic to press publisher's online news sites. The "press publisher right" passed in Germany and Spain described earlier offers the ideal research setting for a natural experiment because the regulations prompted Google News to shut down in Spain. Thus, scholars have been able to evaluate news consumption before and after the shutdown in Spain. Calzada and Gill (2020) studied these scenarios and found the Google News shutdown reduced traffic to Spanish news outlets between 8% and 14%. The study revealed that the degree to which the shutdown affected an outlet's traffic depended on its value proposition — outlets that focused on sports,

regional news, or were lower-ranked tended to be the most harmed by the shutdown. Another closely related paper published by Athey et al. (2017) found the shutdown reduced overall news consumption of users in Spain by approximately 20% and reduced news outlet page views by 10%. Similar to Calzada and Gill, Athey et al. found the decline was most significant for smaller publishers. Both these papers reference Chiou and Tucker (2016), which was one of the first papers to pursue a natural experiment related to news aggregators by analyzing news consumption following the Associated Press's (AP) content dispute with Google News that resulted in the removal of AP content for the aggregator. Chiou and Tucker found traffic to other news outlets from Google News declined by 28% following the AP removal and contributed to a decline of 110 million visits each month to news outlets. All of these studies point to the increasing evidence that Google News is a complement to traditional press publisher's websites given their website traffic suffered when Google News shutdown in certain instances.

### ***Does Google News Benefit or Harm Press Publishers?***

In order to make sense of the conclusions drawn above that news aggregators do not compete and are not substitutes with news outlets; the question remains — why then have press publishers been so aggressive in their lobbying efforts to regulate Google News? The answer to this question becomes clearer after unpacking research that describes how news aggregators have changed the landscape of news consumption habits. From a press publishers' perspective, Jeon and Nasr (2016) found that news aggregators increase the quality of news content, thereby increasing consumer surplus and welfare. Yet a separate study conducted by Jeon (2018) also found evidence that while news aggregators increase traffic to individual news articles, they reduce traffic to newspaper homepages. And even more concerning, Dellarocos et al. (2016) found that the more information provided by the news aggregator in terms of snippet length, the less likely users are to click to the original article. This sampling of research indicates that while there is mounting empirical

evidence that news aggregators drive consumption, there is competing evidence that this type of news consumption may not be advantageous to press publishers and more importantly, may not offer a clear pathway for monetization.

### ***Is There a Better Way Forward for Media Businesses?***

The last portion of this literature review will concentrate on studies that have examined the policy narratives put forth by media business stakeholders. This also ties to the motivation for this thesis because these papers often advocate for a re-framing of the current discussion guiding policy ideas such as “a press publisher right.” For example, Searle (2020) evaluated responses to UK government consultations on copyright to uncover an overall defensive tone among media stakeholders that their business models need protecting. Yet, Searle contextualizes these results to also suggest that the business model narrative may be unproductive and a shift to an innovation focus may be needed. A different narrative is put forth by Frosio (2020) who reviews the failures of the CDSMD to suggest the EU should shift away from author’s rights and instead advocate for a “welfare and cultural theory.” And while perhaps a bit more distant from the media narrative, Khan (2017) illuminated that the consumer welfare argument put forth by online platforms are not equipped to capture the market power imbalance and a different approach is needed when justifying the application of competition law.

As noted in this sampling of papers, there is hardly a consensus of how the future policy debate regarding the regulation of online platforms and their relationship with media enterprises should be framed. Subsequently, this thesis will contribute to the groundwork by providing a systematic overview of current frames being pursued by stakeholders affected by news aggregator policies — information that is necessary to better understand the context for proposed frames detailed above.

## **Theoretical Framework**

This thesis is grounded in the interaction of two theoretical ideas: framing theory and differing legal conceptions of copyright. The relationship between the two theories is captured in the hypothesis which leverages framing theory to suggest that EU and AU stakeholders will tailor their arguments according to the civil and common law conceptions of copyright. In other words, framing theory offers the structure which will express how EU and AU stakeholders position their arguments in two different legal systems.

### ***Framing Theory***

The methodology of this thesis will be informed by framing theory. The most widely cited definition of framing is, “to select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation and/or treatment recommendation for the item described” (Entman, 1993). This definition is widely accepted by interest group scholars across political science, psychology, sociology, and communications to evaluate the arguments interest groups promote and how this affects their policy influence.

The framing of specific policy issues that have been analyzed using this theory include the death penalty, the Gulf War, the Kosovo crisis, pornography, and the Monica Lewinsky scandal (Kluver and Mahoney, 2015). All of these studies performed systematic text analysis to extract frames from a set of documents related to the policy issue of interest. As referenced earlier, a frame is “a specific aspect of a policy proposal that is emphasized in a policy debate by a specific actor” (Entman, 1991) Thus, by reviewing public documents such as parliamentary debates, committee deliberations, presidential speeches, news articles, or consultations, a researcher is able to simplify a complex issue and extract “words and images that make up a frame” that are used to “promote perceptions and interpretations that benefit one side while hindering the other” (Entman, 2010).

Thus, issues best suited for the application of framing theory include those that are “highly salient, controversial, and partisan” where different groups of stakeholders are emphasizing different aspects of a policy’s implications (Kluver and Mahoney, 2015).

To build upon the *Literature Review*, research that has utilized framing theory to evaluate texts related to public policy issues was further examined. For example, Schonhardt-Bailey (2005) performed quantitative text analysis on 20 speeches related to national security given by John Kerry and President George W. Bush during the 2004 Presidential Election. The results identified multiple frames, including a “gratitude” and “fear” frame employed by Kerry and Bush respectively. According to Schonhardt-Bailey, Kerry “invoked the image of shared military experience and sacrifice” whereas Bush “painted the world as a fearful and evil place” (Schonhardt-Bailey, 2005).

Inspired by Schonhardt-Bailey, Kluver and Mahoney (2015) sought to extract frames from two EU public policy debates — (1) regulating CO<sub>2</sub> car emissions and (2) international rail passenger rights. The textual data used in this study was consultation responses separately collected by the European Commission on both policy proposals. Through quantitative text analysis, Kluver and Mahoney identified the presence of a “press,” “industry,” and “environment” frame in the CO<sub>2</sub> debate and an “accessibility,” “voluntary,” and “passengers” frame in the rail passenger debate which she contextualized in the policy aftermath to discuss which stakeholders prevailed, concluding it was the automobile industry and rail operators. In this thesis, a similar methodology will be applied to the news aggregator consultation responses in order to compare the extracted frames to the hypothesized frames informed by the EU and AU conceptions of copyright, discussed next.

### ***Conceptions of Copyright: Civil Law versus Common Law***

The first copyright act in the world was the “Copyright Act of Queen Anne,” passed in 1710. It is often considered “the source of Anglo-American copyright law” (Bracha, 2010). The idea for the statute originated with the Stationers’ Company of London, a printing guild who had been

petitioning the Crown for nearly a decade for increased protection for printed books. However, after multiple bills proposed by the Stationers' Company failed, the guild changed strategy in 1707 by "eliminating all references to censorship and shifting the gravity center of the argument to the protection of authors and the encouragement of learning" (Bracha, 2010). By positioning authors as the "prime beneficiaries" of the statute, it was enacted three years later. Interestingly, the history of the "Copyright Act of Queen Anne" not only offers an example of the value of re-framing a policy debate, but it also serves as a starting point for understanding how the conception of copyright has evolved across time and place in the following centuries.

There is no single common-law or civil law intellectual property (IP) legal system. As the World Intellectual Property Organization (WIPO) points out, "even within the two general legal traditions, common and civil, laws differ from jurisdiction to jurisdiction; and every individual system of law has its own history and development" (Harms & World Intellectual Property Organization, 2018). A unilateral conception of copyright within the civil law or common law tradition does not exist. Yet, there is still a rich theoretical discourse surrounding the differences of the two legal traditions that will be drawn upon in this thesis. It is also important to note that the differences of how copyright is conceived will be grounded in theory, rather than procedural differences. WIPO defines the IP differences in common law versus civil law countries as "relatively small" with only two major procedural differences: the role of the judge and use of precedents (ibid). Subsequently, the theoretical approach of this thesis is captured by the phrase "conception of copyright" to denote the focus on comparing the nature of copyright within the two legal traditions.

The differences in how civil versus common law conceive of copyright begins with the name of the right itself. The civil law tradition uses "author's right" whereas the common law tradition uses "copyright" (Harms & World Intellectual Property Organization, 2018). This distinction serves as a starting point to understand how the two traditions view their role in protecting creative works.

As mentioned earlier, civil law targets, “creators and their claims to ensure the authenticity of their works” (Baldwin, 2014). This idea is said to romanticize the role of the individual artist and “protect the creator’s vision from commercialization and exploitation” (ibid). Consequently, the idea of author’s rights is often criticized with restraining, “distribution, inhibiting experimentation, and public exposure” (ibid). On the other hand, the common law perspective uses the term copyright to, “focus on the audience and its hopes for an expansive public domain” (ibid). By regarding the creator as an entrepreneur and the work as a product, copyright aims to encourage innovation and promote dissemination. It is this idea that gives creators, “a limited economic monopoly over their work to stimulate their creativity, eventually enrich the public domain, and thereby serve the public interest” (ibid). Criticisms of copyright often associate the idea as “philistine and commercial, treating noble creation as a mere commodity” (ibid). This idea of the civil law conception favoring the creator and the common law conception favoring the consumer’s access to content serves as the theoretical framework for my hypothesis.

## **Methodology**

This thesis will be guided by the three research questions defined in the *Statement and Justification of Problem*. The first question examines the textual similarity of EU and AU consultation responses by generating a vector space model to visually represent each submission on a cluster map. The second question will build upon this cluster map and perform topic modeling to extract relevant keywords from submissions and begin to identify the frames employed in the EU and AU submissions. Lastly, the third question will perform manual text analysis to verify the frames identified in the second question. These frames will also be contextualized in the aftermath of the final policy proposals to draw conclusions about which frames prevail in a common law versus civil law legal system. Before expanding upon this methodology, the consultation datasets that serve as the building blocks of this thesis will be described.

### ***Data Collection***

This thesis will use two public consultations, one conducted by the European Commission and the other by the Australian Competition and Consumer Commission (ACCC). The practice of administering consultations occurs early in the policymaking process to gather public opinions from stakeholders on a particular issue. In the case of the news aggregator debate, both the EU and AU used consultations to scope stakeholder opinions on news aggregators to determine whether proposed regulations are an appropriate path forward. The raw submission files were downloaded from the EU and AU consultation websites for analysis. Submissions explicitly referenced in this thesis were compiled into two PDFs titled “EU Selected Submissions” and “AU Selected Submissions.” Both files were publicly uploaded to Dropbox and the links to these files are included in *References*.

The EU consultation, “Public Consultation on the Role of Publishers in the Copyright Value Chain and Panorama Exception,” was administered from March 23, 2016 to June 15, 2016. As the



title infers, the consultation combined two separate public policy issues related to copyright — press publications and architecture or sculpture in public places (the “panorama exception”). The complete dataset includes 6,203 responses, counting both the press publication and the “panorama exception” sections. However, only the press publication section is relevant to this thesis which received 3,957 replies. These responses were split into five categories by the European Commission: registered organizations, non-registered organizations, individuals, anonymous, and the #fixcopyright coalition. Given the AU consultation did not include replies from individuals in their personal capacities, this analysis of the EU consultation was limited to submissions from registered and non-registered organizations. The distinction of “registered” versus “non-registered” relates to an organizations status in the European Commission’s transparency register, a database that tracks lobbying efforts of “registered” organizations. It should be noted that although the individual, anonymous, and #fixcopyright coalition categories represented over 80% of the total responses to the consultation, the majority of these submissions shared a standardized form response.

Amongst the registered and non-registered organization submissions, relevant stakeholder groups were identified to align with the AU consultation, resulting in some submissions to be eliminated from the dataset that represented non-relevant interests such as performing arts organizations. Moreover, while the consultation accepted responses written in English, French, and German, this thesis evaluated only responses submitted in English. After cleaning the dataset accordingly, the final EU consultation dataset included 132 submissions. Further information about stakeholder representation and geographic distribution of these submissions is found in the *Sample Statistics* section.

The AU consultation, “Digital Platforms Inquiry,” was administered in two phases. The first phase, lasting from February 26, 2018 to April 3, 2018 sought responses to an “Issues Paper” with the goal of understanding, “the impact of digital platforms on the supply of news and journalistic

content” (ACCC, 2018). Responses from the second phase, open from December 10, 2018 to February 15, 2019, were prompted by a “Preliminary Report” which outlined 11 recommendations and eight areas for further analysis by the ACCC. The first phase received 77 responses and the second phase received 124 responses, for a total of 201 responses. All responses are in English. For the purposes of this thesis, the responses from both phases will be considered in aggregate. If a stakeholder responded to both phases, the two responses were aggregated into one observation for analysis. After combining necessary submissions and removing submissions that represented non-relevant stakeholder groups, the final AU consultation dataset included 111 submissions.

### ***Data Preparation***

Merging the EU and AU datasets yielded a complete dataset of 243 submissions — considered the “corpus” of this thesis. Before performing quantitative text analysis, the corpus was pre-processed. Both the pre-processing and later text analysis were completed using the Natural Language Toolkit (NLTK) module within Python. In order to manipulate the corpus with NLTK, each submission was converted from a PDF to a TXT file. As the EU consultation included a multiple-choice section, the conversion of the EU submissions required additional formatting to include only the short answer responses in the TXT file. Once all 243 submissions were converted to TXT files and saved in a directory, a DataFrame was built in Python to organize the corpus. The 5 x 243 DataFrame dedicated a row to each submission and columns to organize important identifier information associated with each submission including organization name, headquarter location (denoted by “ctry”), and stakeholder type (Table 1).

**Table 1. DataFrame Sample**

Index	Organization Name	Cons.	Ctry	Stakeholder Type	Submission Text
0	University of Vienna: Media Governance...	EU	AT	Education or Research Institution	“The controversial outcomes in Spain and in Germany of introducing...”
1	Allied for Startups	EU	BE	Startup	“Today “writers” are everywhere blogging and commenting on...”

2	AMEC / FIBEP	EU	BE	Trade Association (Media)	“For services where contents need to be communicated to the clients...”
3	Association of Belgian Publishers	EU	BE	Book Publisher	“When signing a contract with their authors publishers receive the...”
...	...	...	...	...	...
242	StartupAUS	AU	AU	Technology Company	“StartupAUS does not support Preliminary Recommendation 7...”

After organizing the corpus, a series of NLTK modules were applied to the “Consultation Text” column to prepare the submissions for text analysis. This included converting all letters to lower case and removing stop words, punctuation, and special characters. Stop words are common words in language that do not convey relevance when identifying meaning such as “the.” The final step of the pre-processing stage was word stemming, a NLTK module that converts words to their base forms such as from “reporting” to “report.”

***Question 1: Measuring Textual Similarity***

After the dataset was prepared, the methodology to achieve the results needed for Question 1 began. To calculate the textual similarity of the submissions, the “Document Clustering with Python” guide was referenced (Rose, 2020). First, the corpus was transformed into a vector space using term frequency-inverse document frequency (TFIDF). TFIDF represents text numerically by counting the frequencies of individual words in each submission. In comparison to other weighting schemes, TFIDF assigns smaller word weights to commonly used words (Bochkay et al., 2020). This attribute is especially helpful given all submissions are responding to one of two consultation prompts and often repeat vocabulary featured in these prompts. The underlying assumption of TFIDF is that words repeated less often contain more meaning in relation to the document. This prevents repetition of less-meaningful words from distorting the final output. After representing the corpus in the TFIDF matrix, the cosine similarity between each submission and the other submissions in the corpus were calculated. Cosine similarity ranges between zero and one, where

zero indicates complete dissimilarity and one indicates complete agreement. The advantage of cosine similarity is it naturally controls for the length of each submission, ensuring a longer submission can be compared to a shorter submission. The cosine similarity of each submission was then subtracted from one to achieve the cosine distance.

Once cosine distance was calculated, it can be inputted into a k-means clustering algorithm to visually represent the similarity between each submission. K-means clustering is an iterative process to determine the optimal number of clusters. Each submission is assigned to a cluster and the mean of the cluster is calculated to find the centroid. Then, the submissions are reassigned to clusters and the centroids are recalculated until the algorithm reaches convergence. As a safeguard, an elbow plot was used to find the optimal number of clusters, which was confirmed to be eleven clusters (Appendix 1). The final step before mapping the output of the k-means clusters is to apply multidimensional scaling which translates the pairwise distances captured in the cosine distance metric into a two-dimensional array with x and y coordinates. The result is each submission can be plotted on an array, belonging to one of the eleven clusters.

### ***Question 2: Topic Modeling***

To understand how the submissions are framing their responses, with particular emphasis on uncovering if a distinction exists between the EU and AU framing as articulated in the *Hypothesis*, topic modeling was pursued. At its core, topic modeling is an unsupervised machine learning technique that scans a corpus, detecting word and phrase patterns to extract topics that best characterize the text. In contrast to supervised machine learning techniques, topic modeling does not require a predefined list of topics. The type of topic modeling utilized in this thesis built upon the methodology of the k-means clustering used for Question 1. By coding the clusters to yield the top six most frequent words of each cluster, a basic version of topic modeling was achieved. The benefit of this approach was it leveraged the TFIDF weighting scheme described earlier.

### ***Question 3: Manual Text Analysis***

Once the output of the topic modeling was retrieved, manual text analysis of the individual submissions was performed to verify the results of the topics identified. In other words, the submissions were reviewed, and relevant topics were categorized and compared with the results of the topic modeling. This qualitative analysis is referenced in more detail in the *Results* section, but ultimately sought to contextualize the machine learning topics identified in the original context of the submissions themselves.

Informed by the results of the topic modeling and manual text analysis, Question 3 also includes a qualitative review of the policy outcomes and current discussion related to news aggregator policy globally. This was completed via online research of the events following the consultations and regulatory activity in other countries. Acknowledging this research method is largely anecdotal, it will serve as a compliment to the systematic frame extraction completed for Question 2.

### ***Limitations***

The limitations of this methodology stem from the reality of comparing two consultations conducted by two different governments. First, the EU and AU consultations inherently had different prompts, formatting, and collection methods that could distort the ability to compare responses and reach generalizable conclusions. Second, while both consultations accepted responses from any member of the general public, it is unclear if distribution methods such as targeting specific types of organizations for responses were employed. This could bias the pool of stakeholder responses received and contribute to an unrepresentative sample. Third, non-English language responses and responses from certain stakeholder groups such as performing arts organizations were removed from the sample for data cleaning purposes that could impact the results. Lastly, the data cleaning methodology employed did not account for differences in Australian versus British versus

American English, which posed issues when performing topic modeling across submissions and caused some terms with the same meaning to appear twice in the output such as neighboring and neighbouring.

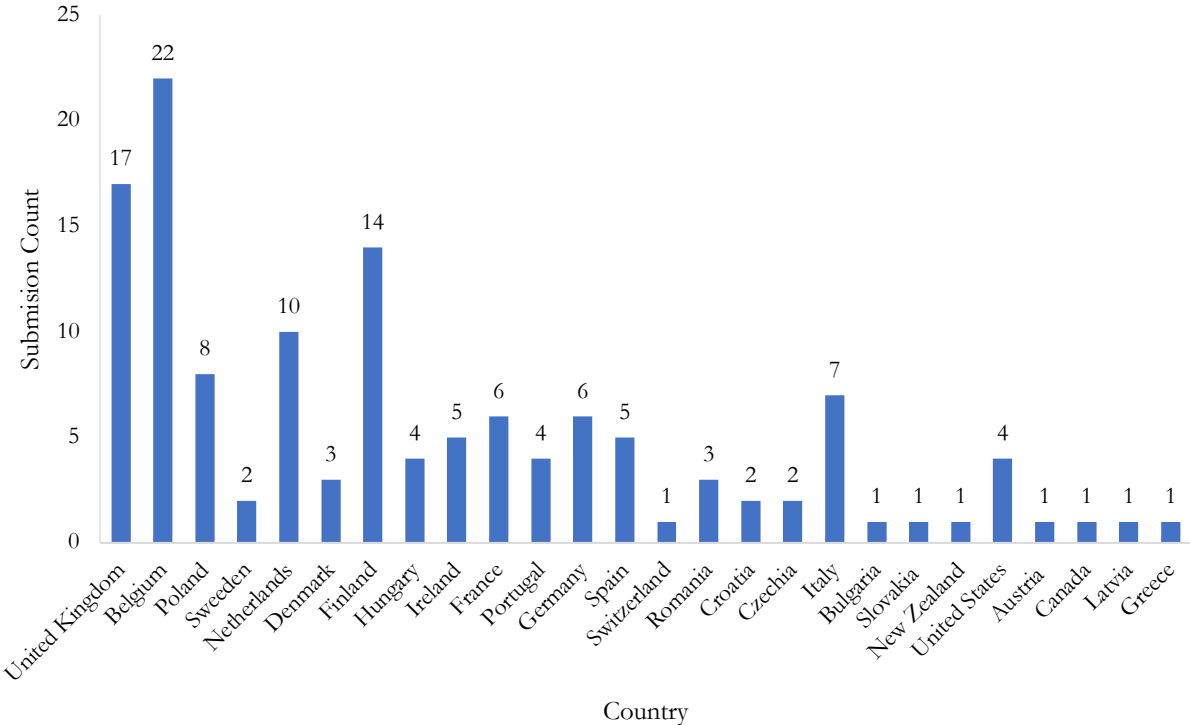
**Results**

*Summary Statistics*

As described in the *Data Collection* section, the dataset includes 243 submissions with 54% representation from the EU consultation and 46% representation from the AU consultation. For both consultations, submissions were received from organizations across the globe, demonstrating that the impact of these policy proposals was felt beyond EU and AU borders.

Figure 1 illustrates the geographic distribution of the EU submissions. The country assignment for submissions with a global presence was determined by headquarter location. For example, Getty Images is considered to be located in the United States with headquarters in Seattle, despite operating in multiple EU member countries. The concentration of submissions from

**Figure 1. Count of EU Submissions by Country**

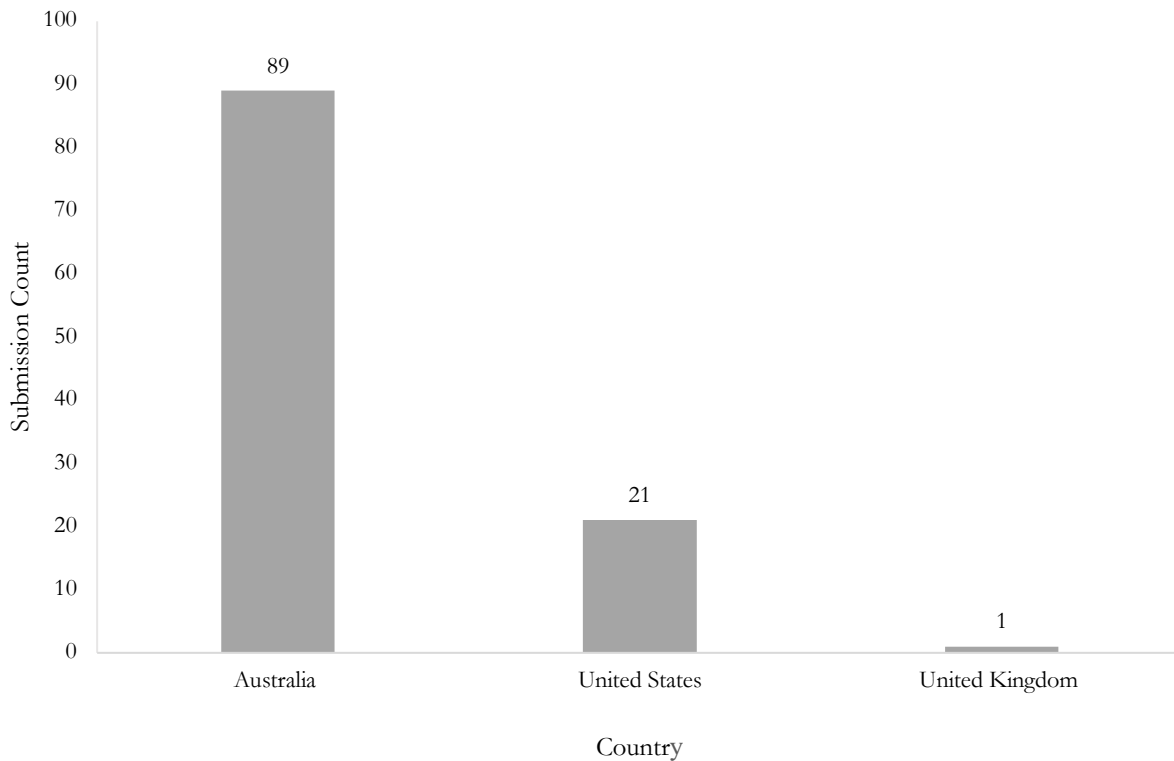


Belgium is driven by a number of principal EU government stakeholders located in Brussels. As expected, the EU consultation submissions have greater geographic diversity in comparison to AU,

given the government operates as a 28-member bloc (before Brexit). And as referenced earlier, despite only English language submissions included in the dataset, 82% of EU member countries are represented. The six EU member countries not represented are Cyprus, Estonia, Lithuania, Luxembourg, Malta, and Slovenia.

In a similar manner, the geographic distribution of the AU submissions is represented in Figure 2. The majority of the submissions were received from organizations in AU, however approximately 20% of the submissions are from organizations headquartered in the United States. This is an important distinction, partially explicable by a large share of stakeholders representing technology and innovation interests headquartered in the United States.

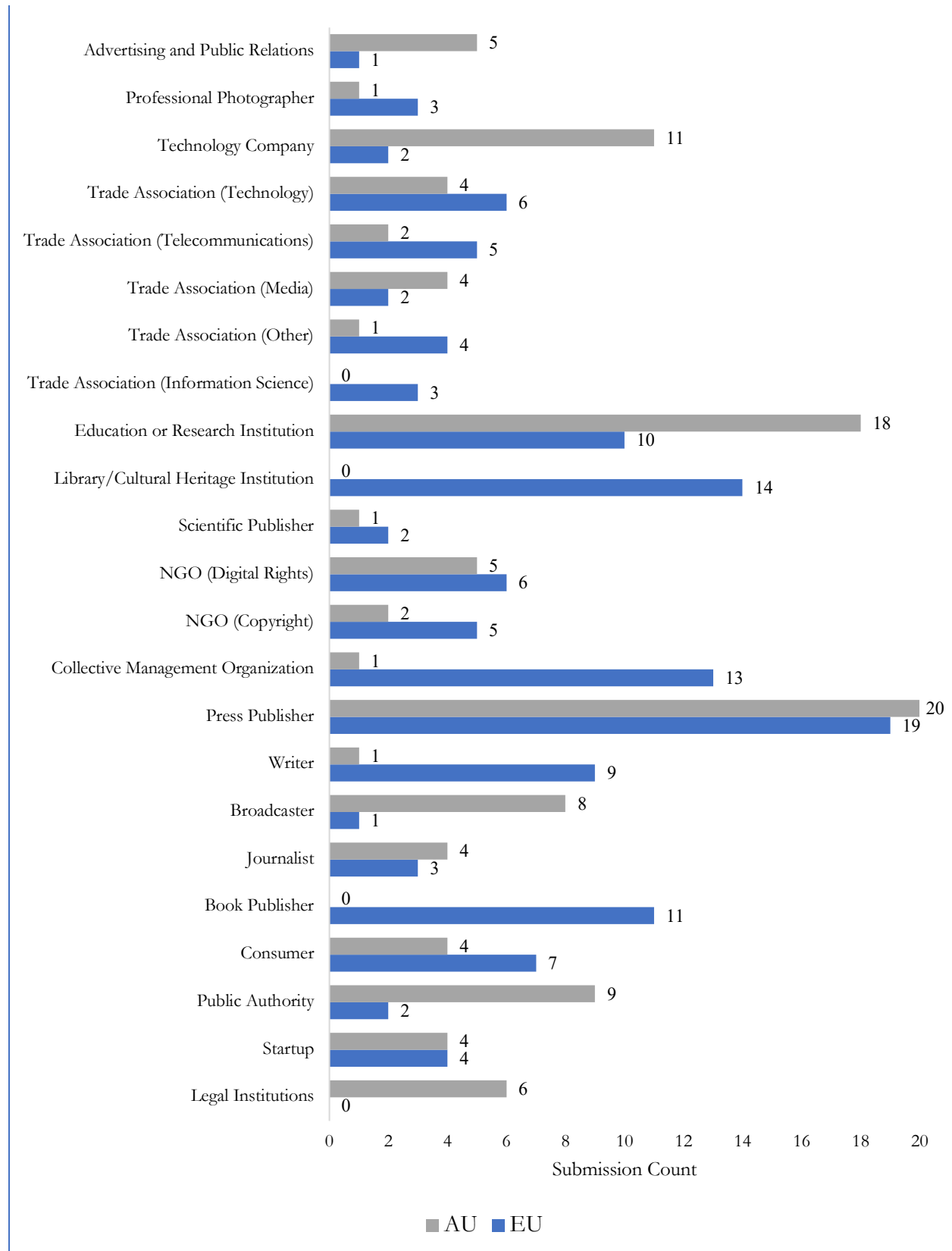
**Figure 2. Count of AU Submissions by Country**



To better understand the identity of submission respondents, beyond geographic location, each submission was labeled by stakeholder type (Figure 3). There were 24 stakeholder categories



**Figure 3. Count of Submissions by Stakeholder Type**



identified, most including representation from both the EU and AU. However, there are several exceptions, most notably there is an EU emphasis on cultural history and information science with 18 EU submissions in contrast to 0 AU submissions in the “Library/Cultural Heritage Institution” category. Another distinction in stakeholder representation is the AU consultation attracted 140% more academic scholars and legal practitioners in comparison to the EU consultation, captured by the combined stakeholder count totals of the “Education or Research Institution” and “Legal Institution” categories.

When considering stakeholder representation in the context of geographic distribution, additional insights emerge, including 75% of stakeholders in the “technology companies” category being located in the U.S. This is unsurprising given U.S. exports of technology products and services totaled \$338 billion in 2019, indicating the global reach of U.S. technology companies and spurring the engagement of U.S. interests in the EU and AU consultations (CompTIA, 2020). This is certainly the case for Google, headquartered in Mountain View, California, with an obvious stake in the news aggregator debate, as established in the Background section. Google response to both consultations, submitting over 100 pages of expert testimony, economic analysis, and stakeholder opinion for consideration. Moreover, even U.S. technology companies without news aggregator products submitted responses to the consultations, reinforcing the mounting threat of the policy proposals and fear among technology companies that the aggressive regulatory agenda will expand to social media, software, and video streaming. The EU consultation received two U.S. technology company responses — Google and Mozilla — while the AU consultation received 9 responses including Oracle, Facebook, Microsoft, Twitter, and Netflix. And it was not solely the companies themselves looking to protect U.S. interests in technology and innovation. The AU consultation attracted responses from public authorities such as the U.S. Chamber of Commerce and technology interest groups such as the BSA Software Alliance and Innovation Defense Foundation — all

headquartered in Washington, D.C. The fact that the AU consultation attracted more responses from technology interests in comparison the EU consultation points to the emphasis on the platforms themselves in the AU regulation. Whereas the EU consultation has a roster of stakeholders representing creative and artistic interests including more submissions in the “professional photographers,” “writers,” and “library/cultural heritage institutions” categories.

**Question 1: Textual Similarity**

Figure 4 plots each consultation response on a two-dimensional plane to represent the textual similarity (use of similar vocabulary) of each submission. Each datapoint represents an individual submission, labeled with a number that corresponds to the organization name, indexed in Appendix 2. The dissimilarity of the EU and AU submissions is strikingly demonstrated with the EU submissions, coded in blue, clustering on the left-side of the plot and AU submissions, coded in grey, clustering on the right-side of the plot.

**Figure 4. Textual Similarity (coded by consultation)**

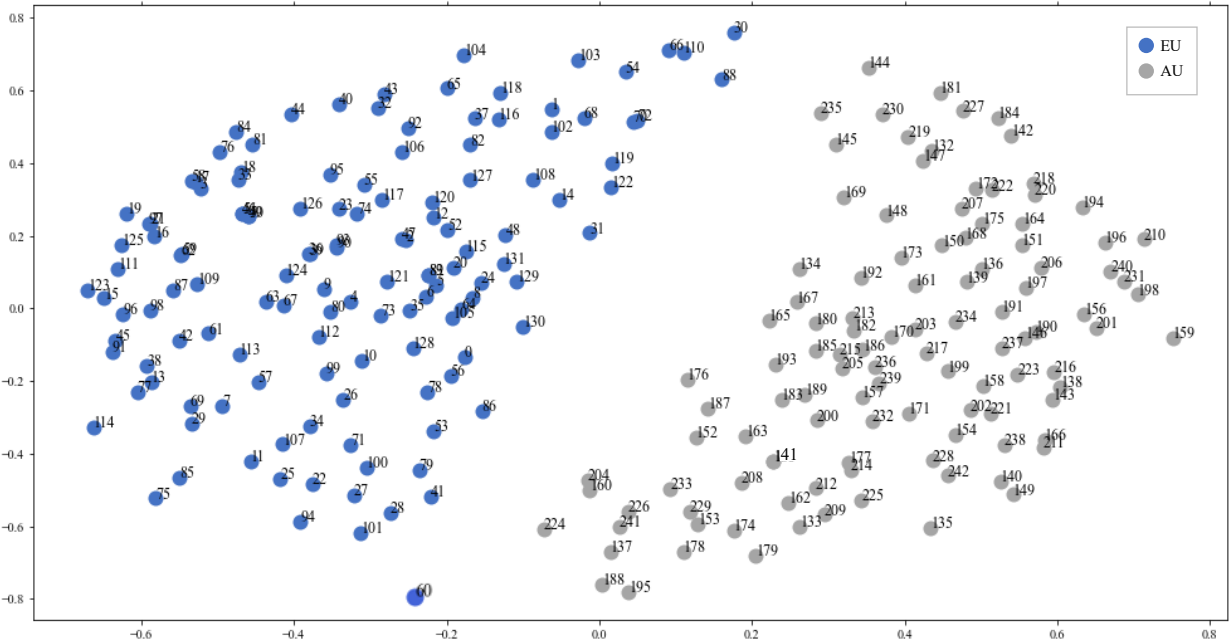
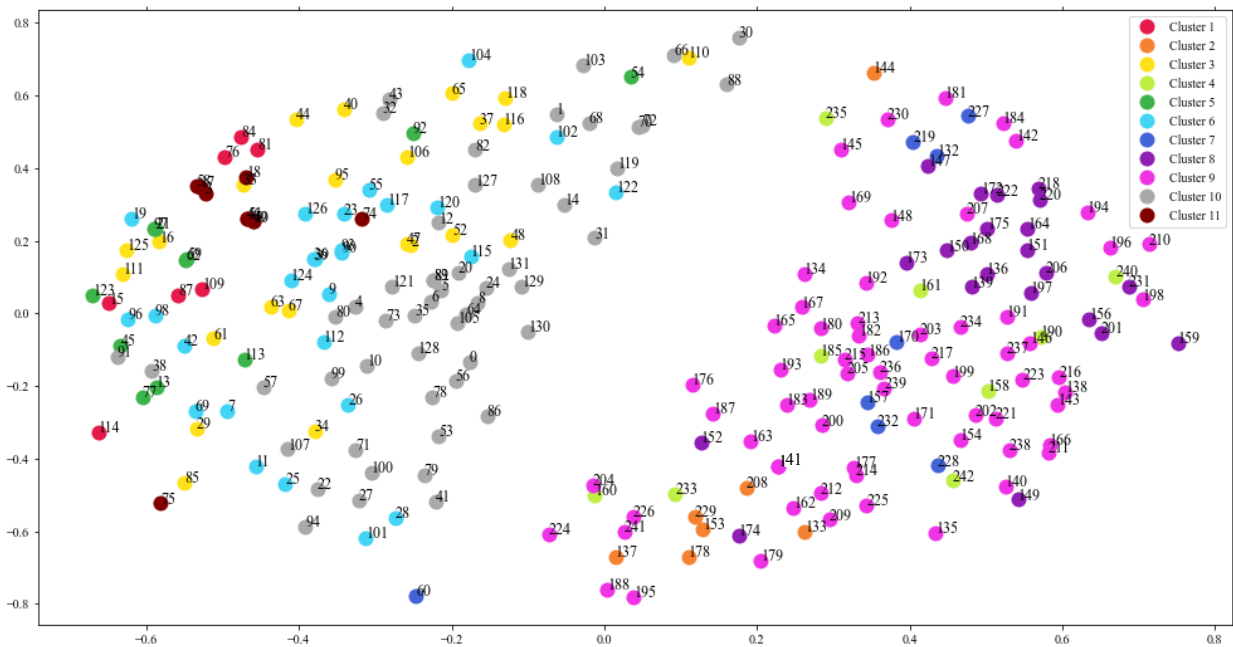


Figure 5 is the same plot, but coded by cluster assignment, grouping the 243 submissions into 11 clusters. There is limited crossover of the EU and AU submissions within each cluster. EU submissions comprise Cluster 1, 3, 5, 6, 10, and 11 and AU submissions comprise Cluster 2, 4, 8, and 9. Cluster 7 is the one cluster that exhibited a mixture of EU and AU submissions with K-Monitor (#60), a digital rights NGO based in Hungary, joining a slew of consumer-focused stakeholders based in Australia. The overall separate grouping of the EU and AU submissions prompts further questions related to what differences exist in the vocabulary and content of the EU and AU submissions that account for these results, answered in the results to Question 2.

**Figure 5. Textual Similarity (coded by cluster assignment)**



**Question 2: Topic Modeling**

To begin to uncover the framing of the EU and AU submissions, topic modeling was performed on the clusters to derive the top six words for each cluster (Table 2). After evaluating the output, a narrative began to emerge with the vocabulary of the EU submissions emphasizing the individual creator. The most frequent words appearing in the EU clusters were “publisher,”

“authors,” and “rights.” This suggests an inward-looking response, as stakeholders considered the regulation through its impacts on the current ecosystem of news with particular focus on authors versus publishers. This is a sharp contrast to the AU clusters which much more frequently call out the perceived threat of platforms such as Google News. This is evidenced by the most frequent words appearing in the AU clusters being “platforms” and words describing news content such as “media,” “news,” and “information.” The results in Table 2 echo the results established in Question 1 with hardly any of the frequently used words identified overlapping between the EU and AU clusters.

**Table 2. K-Means Topic Modeling**

Cluster	Cons.	Output
Cluster 1	EU	authors, rights, publishers, remuneration, visual, neighbouring
Cluster 3	EU	rights, publishers, neighbouring, neighbouring, new, authors
Cluster 5	EU	rights, publishers, journalists, authors, works, neighboring
Cluster 6	EU	publishers, rights, authors, content, press, copyright
Cluster 10	EU	publishers, rights, new, new, copyright, content
Cluster 11	EU	publishers, book, authors, book, rights, remuneration
Cluster 2	AU	journalism, media, fund, news, abc, publications
Cluster 4	AU	platforms, digital, Australian, digital, infringements, content
Cluster 8	AU	privacy, data, consumers, recommend, platforms, information
Cluster 9	AU	platforms, digital, news, advertisers, digital, media
Cluster 7	EU/AU	children, media, platforms, digital, young, regulatory

***Question 3: Manual Text Analysis***

The “creator” frame manifests itself across the EU submissions. Yet, it is important to note that the emphasis on the individual creator does not mean all of the submissions hold the same

position towards the proposed regulation. Rather, as noted in the *Theoretical Approach* section, a frame is emphasizing one aspect of an issue over another. Thus, in this case, while the EU consultation largely emphasizes the impact of the proposed regulation on creators such as journalists, writers, and authors — all the submissions do not necessarily hold the same opinion towards whether that impact is positive or negative.

The EU submissions also share a tendency to barely mention digital platforms such as Google News and instead focuses on the power dynamics of creators versus publishers and questions related to the ownership of the news content itself. This tactic raises questions, specifically if the focus on the traditional players in the news media industry and subsequent infighting hindered the larger goal of regulating Google News. To better understand these dynamics, Table 3 maps the general positioning of each EU cluster towards the proposed regulation and evidence of how the “creator” frame is expressed in each cluster.

**Table 3. EU Clusters | Creator Frame**

Cluster	Cluster Tag	Framing Examples (representative of cluster)	
1	creator exploitation	84: The Authors Guild of the Netherlands	“[publishers] would get even more... while authors get even less. The underlying problem is the lack of a level playing field between publishers and authors and fair contracts” (EU Selected Submissions, 211).
		87: Norwegian Visual Artists Copyright Society	“It is most likely that publishers will seek ways to circumvent authors and visual creators in order to ensure that such remunerations are paid entirely to themselves” (EU Selected Submissions, 191).
3	severe economic implications	85: Vrijschrift	“It would unjustly benefit the biggest publishers...as well as introduce unnecessary friction in the marketplace through transaction costs” (176).
		95: Impresa	“A new Publisher’s Right would allow the Publisher to not depend on any acquisition of third-party rights (e.g., Author) in order to enforce its rights. This would make enforcement simpler, quicker, cheaper and with less parties involved” (EU Selected Submissions, 162).
5	hurt creator welfare	13: European Federation of Journalists	“the neighboring right could conflict with journalists’ own reproduction and making available right over the content on which they still own the rights (if not exclusively licensed)” (EU Selected Submissions, 151).
		54: Martin Kroll	“the more rights you grant to nonoriginal stakeholders in any whatsoever artistic or creative field, the more it could backfire to the original artist” (EU Selected Submissions, 134).
6	enhance creator welfare	23: The Union of Publishers of Bulgaria	“The acknowledgement of publishers in all sectors as right holders would have a positive impact on authors, as publisher’s revenues will

			rise and a fair share of the increased income would certainly benefit authors and allow more editions to be profitable” (EU Selected Submissions, 123).
		95: The Polish Chamber of Press Publishers	“A publisher’s right would be without prejudice to authors’ rights and will have no impact on the contractual relationship between publishers and journalists” (EU Selected Submissions, 105).
		126: The Publishers Association	“Authors would see the technical environment around their content better protected and publishers’ capacity to invest in new content and technology strengthened” (EU Selected Submissions, 86).
10	damage all stakeholders	22: Wikimedia	“The results we can observe in the countries that attempted this are negative through the board. Both publishers and journalists are losing traffic and exposure of their content, leading to a drop in revenues” (EU Selected Submissions, 71).
		86: Scoop Foundation for Public Interest Journalism	“For the same reasons that it would be a disaster for online service providers, internet, regulators and the rule of law it would be a disaster for the public - it would break the internet (EU Selected Submissions, 62)
		103: Association of Spanish Startups	“The idea that everything published in writing on the internet should get an additional, new set of copyright rules is scary. It promises a new wave of legal uncertainty, complexity and red-tape for all the businesses, large and small” (EU Selected Submissions, 46).
		131: Mozilla	“The harm that would come from imposition of such an extraneous and unnecessary mechanism would stifle innovation and creativity and would lead to less interesting content available to be published - above and beyond the proximate and specific problems that publishers face - creating long-term damage to the entire ecosystem” (EU Selected Submissions, 21).
11	neutral	50: Hachette Livre	“the impact on publishers of the creation of a new neighboring right in EU law needs to be further explored” (EU Selected Submissions, 6).

Whether an EU stakeholder is for or against the proposed “press publisher right,” the consultation responses share an emphasis on how authors, journalists, and writers will be impacted. The “creator” frame appears in each cluster, underscoring the influence of the creator centric copyright orientation in the civil law legal system.

In Cluster 1, respondents such as the Norwegian Visual Artists Copyright Society fear “publishers will seek ways to circumvent authors” in order to keep remunerations for themselves (EU Selected Submissions, 191). This fear articulated in Cluster 1 of the potential strengthening of press publisher power at the expense of creator rights offers an initial starting point for understanding how unsettled the industry perspective is towards addressing news aggregators. This

uncertainty translates to an emerging narrative that often pits creators and publishers against each other — something that continues throughout the rest of the EU clusters.

Cluster 3 submissions largely focused on the economic implications of the proposed regulation, but still orient this economic analysis towards its impacts on creators versus publishers. This is exemplified in the response submitted by *Impresa*, a Portuguese publication, which explains how a press publisher right would enable it to avoid acquiring copyrights from writers. This would in turn make it easier for *Impresa* to enforce its rights when litigating copyright infringement claims against news aggregators (EU Selected Submissions, 162). In other words, *Impresa* believes the regulation would make “enforcement simpler, quicker, cheaper and with less parties involved” (ibid). There is a perceived economic efficiency associated with a press publisher right because it would allow for creator’s rights to be circumvented and a subsequent decrease in transaction costs — a belief shared across submissions belonging to Cluster 3. It is interesting that publishers view the creators of news content to be such an obstruction to exercising their rights over news aggregators. While Cluster 1 expresses a fear of creators being exploited by publishers, Cluster 3 seemingly admits this to be true and acknowledges it would allow publishers to increase their legal rights over news content written by creators such as journalists and writers.

Cluster 5 calls out the harm a press publisher right poses to creator welfare but does not focus as heavily on the publisher exploitation factor described in Cluster 1. As Cluster 5 is largely composed of stakeholders representing journalists and writers’ interests, it is expected that the “creator” frame would be present. The European Federation of Journalists expresses a concern voiced by multiple submissions in Cluster 5 that a press publisher right would hinder creators from syndicating their content across multiple publications (EU Selected Submissions, 151). In other words, once a piece of news content is published in one publication, a press publisher right would cause that publication to have proprietary access to that content, barring a creator from publishing



the piece elsewhere at a later date. Other concerns voiced in Cluster 5 hinge on this idea that a press publisher right would disrupt how creators perform their tasks such as securing interviews when future publication of the content produced is at the discretion of the publication, not the creator itself. In the words of Martin Kroll, a journalist based in Germany, “the more rights you grant to nonoriginal stakeholders in any whatsoever artistic or creative field, the more it could backfire to the original artist” (EU Selected Submissions, 134).

The sentiment expressed by Kroll that a press publisher right would “backfire” on creators is a stark contrast to Cluster 6 which claims the proposed regulation would actually enhance creator welfare. At the crux of these submissions is the belief that a press publisher right would lead to increased revenues for press publishers which would then be passed down to the creators generating content for these publications. For example, The Publishers Association believes a press publisher right would lead to “the technical environment around their content better protected and publishers’ capacity to invest in new content and technology strengthened” (EU Selected Submissions, 86). Cluster 10 not only believes remunerations will be reinvested into creators’ pockets, but also the current relationship between publishers and creators will remain the same following the passage of a press publisher right. As the Polish Chamber of Press Publishers explains, “a publisher’s right would be without prejudice to authors’ rights and will have no impact on the contractual relationship between publishers and journalists” (EU Selected Submissions, 95). This is a direct contradiction to submissions in Cluster 1 and Cluster 5 that believe creators will be exploited by publishers and not reap the benefits of remunerations.

Cluster 10 does not believe a press publisher right will help creators or publishers, instead suggesting it will harm the entire ecosystem of news. As the membership of Cluster 10 is largely comprised of technology interests, it makes sense that a total rejection of the press publisher right would occur. Google is included in Cluster 10 with a submission warning of the complexity that will

ensue when “the rights of authors and those of publishers are in different ownership” (EU Google Submission, 3). Yet, it is not just the big technology conglomerate who recognizes this possibility. The Association of Spanish Startups similarly fear “a new wave of legal uncertainty, complexity and red tape for all the businesses, large and small” if a press publisher right is passed (EU Selected Submissions, 46). Cluster 10 also includes a submission by the Scoop Foundation for Public Interest Journalism which surprisingly aligns with the technology companies by citing how it would be a disaster for many stakeholders and “break the internet” (EU Selected Submissions, 62). In comparison to the other EU clusters, Cluster 10 has the least representation of the “creator” frame and most similarity to the AU frames centered around content. One reason for this could be many organizations with a stake in protecting technology interests are headquartered in the United States, a country with a common law legal tradition that could inform an interpretation more similar to AU.

Cluster 11 is the smallest cluster and more difficult to pinpoint a shared narrative among the submissions, but there is a general neutrality expressed towards the press publisher right and hesitancy to stake any clear position towards the regulation.

Pivoting to the AU clusters, a manual text analysis revealed an emphasis on the digital platforms themselves including the focus of this thesis — Google News. This translated to stakeholders framing their submissions via the news content and how it is distributed and accessed on news aggregators. Unlike the EU clusters, there is little mention of who is creating the content and their interaction with publishers. Table 4 reports the positioning of each cluster and offers representative examples of how the content frame is employed.

**Table 4. AU Clusters | Public Interest Frame**

Cluster	Cluster Tag	Framing Examples (representative of cluster)	
2	public good	133: Australian Broadcasting Corporation	“The structure of the media market and consumer behaviour and expectations in that market have irrevocably changed. What has not changed is that for the market to thrive and deliver competition, quality, and choice to consumers, as well as provide a public good” (AU Selected Submissions, 152).
		208: Monash University	“I would encourage the inquiry to think broadly about ‘news and journalistic content’... to focus on these purposes or functions of journalism, rather than on the locations in which the journalism takes place” (AU Selected Submissions, 123).
4	accountability	158: Australian Copyright Council	“Digital platforms holding significant market power should have some accountability for the misuse of others by their platforms (whether it be the sharing or locating of infringing content)” (AU Selected Submissions, 156).
		240: Stan	“there should be greater onus on social and digital media platforms to police this issue proactively” (AU Selected Submissions, 119).
7*	consumer access to information	60: K-Monitor	“[a press publisher right] would restrict access to information and limit freedom of expression” (AU Selected Submissions, 113).
		160: Australian Communications and Media Authority	“There should be a diversity of major information sources and perspectives expressed in the public sphere to foster an informed citizenry and healthy democracy” (AU Selected Submissions, 97).
8	misuse of privacy and competition law	149: U.S. Chamber of Commerce	“Competition law should not, as some have suggested, be used as a tool to address broader social policy issues... or to promote the interests of competitors or other third parties over consumers” (AU Selected Submissions, 86).
		151: UNSW Sydney: Allens Hub	“Concealed data practices clearly raise consumer protection concerns, but they also raise competition issues” (AU Selected Submissions, 78).
		164: Australian Privacy Foundation	“The effective control of large data sets exercised by platforms, such as Google and Facebook, supports and reinforces network effects and the substantial market power possessed by platforms” (AU Selected Submissions, 65).
		173: Computer & Communications Industry Association	“The key to gaining a competitive edge in the digital economy is not the accumulation of data, but rather, the capacity to analyze and monetize data” (AU Selected Submissions, 42).
9	value of publisher content	230: Queensland County Press Assoc.	“No other media — certainly not Google or Facebook — produces the unique content our member newspapers provide” (AU Selected Submissions, 31).
		199: International Center for Law and Economics	“Newspapers and digital journalism have no normative claim over the online advertising revenue. These are markets that did not exist before they were established by Google” (AU Selected Submissions, 18).
		180: Digital Industry Trade Association	“For many journalists, these platforms form an integral part of how they identify emerging stories, distribute content, connect with audiences, and measure engagement” (AU Selected Submissions, 2).
		213: News Corp Australia	“Publishers of news and journalism are being increasingly corralled into creating content that benefits the digital platforms – rapid, undifferentiated, even copied, ad-funded content” (News Corp Submission, 88).

Cluster 2 considers journalistic content via the public good role it serves in society. For example, the Australian Broadcasting Corporation notes the central role of the media market is to “deliver competition, quality, and choice to consumers, as well as provide a public good” (AU Selected Submissions, 152). This focus on the role of content is shared by other submissions in Cluster 2 including academics at Monash University who believe an examination on the “purposes or functions of journalism” should take precedent over the “locations in which the journalism takes place” (AU Selected Submissions, 123). This is different from the EU clusters which devote a significant amount of their submissions to debating whether journalism should even be allowed to be distributed on certain platforms and not others.

In combination, Cluster 4 and Cluster 8 focus their attention on the more technical components of the proposed regulation including the absence of a safe harbor protection and data privacy concerns which are outside the scope of this thesis but do reinforce a concern for the protecting the content itself. Cluster 4 emphasizes holding news aggregators accountable for infringing upon news content. This is expressed by the Australian Copyright Council whose submission advocates for “accountability for...sharing or locating of infringing content” (AU Selected Submissions, 123). Cluster 8 is concerned with the data collection tactics of news aggregators in relation to the type of content read by users and how they interact with the platform. Cluster 8 also represents a debate about the application of competition and data privacy law, all tied to the market power of news aggregators. Cluster 8 is arguably an outlier in relation to direct use of the content frame, but the undertones of the submissions still relate to how consumers are accessing news content on the platforms.

Cluster 9 is perhaps the most explicit example of the content frame with submissions debating the value of publisher content and whether news aggregators are taking advantage of the content produced by newsrooms. The makeup of submissions in Cluster 9 features both journalism

and technology interests, demonstrating the pressure felt both stakeholders to answer whether the aggregation of news content by technology companies is acceptable. Explained by the Queensland County Press Association, “no other media — certainly not Google or Facebook — produces the unique content our member newspapers provide” (AU Selected Submissions, 31). This is a sentiment shared by press publishers across Cluster 9, including News Corp whose submission goes further to point out how news aggregators are reshaping the content landscape for the worse with the reality of “rapid, undifferentiated, even copied, ad-funded content” performing better on news aggregators than higher quality reporting (News Corp Submission, 88). Cluster 9 also includes the rebuttal of technology interests such as the Digital Industry Trade Association which attempts to highlight how news aggregators provide tools to journalists to help enhance their content and “identify emerging stories, distribute content, connect with audiences, and measure engagement” (AU Selected Submissions, 2). Included in these rebuttals from technology companies is the response of Google which defends its news aggregator product by stating how it has made “it easier for users to search for journalistic content and to connect with the publishers that have the news that most interests them” (AU Google Submission, 10). This is a contrast to how Google frames its response to the EU consultation in Cluster 10 with the “creator” frame.

As noted earlier, Cluster 7 is largely comprised of AU submissions, but does include an EU submission by K-Monitor, a Hungarian digital rights NGO. The emphasis of Cluster 7 is on consumer wellbeing and mostly includes AU stakeholders representing digital rights interests. For example, The Australian Communications and Media Authority makes a claim that the diversity of information sources and perspectives found in news content play a critical role in fostering “an informed citizenry and healthy democracy” (AU Selected Submissions, 97). This is similar to the K-Monitor submission which expresses reservations that “[a press publisher right] would restrict access to information and limit freedom of expression” (EU Selected Submissions, 113). For the purposes

of this analysis, Cluster 7 was recognized as an AU cluster because it only featured one EU submission and aligned with the content framing employed in the other clusters.

## Discussion

As predicted in the hypothesis, the combination of quantitative and manual text analysis discussed in the *Results* section reveals an alignment between EU and AU legal traditions and their respective consultation frames — “creator” and “content.” The “creator” frame prevails in the EU submissions no matter if it is *Impresa* discussing the economic implications of a press publisher right or Google warning of the threats associated with a more regulated press. Whereas the content frame emerges across AU submissions from the Digital Industry Trade Association defending the valuable tools that news aggregators provide to content generation or News Corp Australia arguing that content quality has decreased due to news aggregators. Yet, given the EU and AU consultations were completed in 2016 and 2019 respectively, the question remains if these frames continue to exist in related policy activity in the EU, AU, and elsewhere.

On April 15, 2019, the “Directive on Copyright in the Digital Single Market” passed in the EU. Nearly two years later, on February 25, 2021, AU passed the “News Media and Digital Platforms Mandatory Bargaining Code.” While both the EU and AU reforms passed, the EU approach of a press publisher right faced significant implementation issues, largely driven by the economic reality that press publishers rely on the traffic Google News drives to their own news sites. The only EU member country to implement the press publisher right so far has been France and as expected, Google refused to engage in licensing negotiations. Instead, Google offered French press publications a choice to either license their content for free or remove all their content from Google (Keller, 2020). In response, the French government abandoned the mechanism of a press publisher right altogether and began to explore alternatives to securing remunerations for press publishers. In April 2020, the French Competition Authority (FCA) became involved and accused Google’s “unilateral withdraw of snippets to be unfair and damaging to the press sector, and likely to constitute an abuse of a dominant market position” (Lomas, 2021). After Google lost its’ appeal

with the FCA, L'Alliance de la Presse d'Information Générale (APIG) which represents over 300 press publications, entered into negotiations with Google to reach an agreement on remunerations. On January 21, 2021 a framework was announced for French press publications to enter independent licensing agreements with Google (ibid). This sequence of events in France proves once again that the press publisher right passed in the EU has yet to succeed, with many academics questioning if it can ever be the vehicle for press publishers to achieve remunerations. As described by Dusollier (2020), “the compromises and bad choices made by the Directive lead to intricate provisions whose success might be illusory. This is clear for the press publications right, the utility and strength of which are disputable.” This is in contrast to AU where despite the regulation only passing a few weeks ago, has already proven to be more successful at bringing Google to the negotiating table. On February 4, 2021, before the AU regulation even officially passed, Google preemptively introduced the Google News Showcase in AU, a program for streamlining the negotiation of licensing deals with individual press publications. To date, Google has entered licensing deals with over 70 AU press publications to avoid the forced arbitration stipulated in the regulation (Beddoe, 2021).

The early success of the AU regulation has been paired with AU regulators continuing to embrace the “content” frame. Josh Frydenberg, the federal treasurer, celebrated the passage of the regulation by stating, “the Code will ensure that news media businesses are fairly remunerated for the content they generate, helping to sustain public interest journalism in Australia” (Lomas, 2021). Conversely, in the EU there has been an abandonment of the “creator” frame in favor of the “content” frame which has proved to be more successful in AU. This has been accompanied by calls from some members of the EU Parliament to amend current regulatory proposals related to the DSM to include measures similar to the AU regulation. For example, Andrus Ansip, Member of European Parliament representing Estonia noted, “we will never accept this situation when



somebody is using content ... and authors are not remunerated at all” (Espinoza and Barker, 2021). Similarly, Stéphanie Yon-Courtin, Member of European Parliament representing France stated, “it’s time to oblige online platforms to engage in fair negotiations to remunerate the news content they obtain from press publishers” (Espinoza and Barker, 2021). The emergence of EU politicians stepping away from the creator centric approach that guided previous news aggregator regulatory efforts reinforces the belief that a “content” frame is more successful in policy debates aiming to bring Google to the negotiating table.

It is not just the EU embracing the “content” frame. A review of other countries considering news aggregator regulation revealed an emphasis of content over the creator as well. For example, the “Journalism Competition and Preservation Act” reintroduced in the United States Congress on March 10, 2021 has been paired with a “content” frame. Statements by bill co-sponsor Senator Amy Klobuchar reference how the legislation “will improve the quality of reporting” (Klobuchar, 2021). In the United Kingdom, similar steps have been taken with the Digital Markets Taskforce announcing its intention to force Google and Facebook to pay news outlets for content. Additionally, the House of Lords is currently revising the “Online Harms Bills” to include news aggregator regulation policies inspired by the AU. Dr. Andrea Coscelli, CEO of the Competition and Markets Authority, who is responsible for leading these efforts in the UK has employed a content frame noting, “competitors should face a level playing field — enabling them to deliver more of the innovative products and services we value so highly” (Coulter, 2021). Lastly, in Canada there has mounting pressure from press publishers for the government to regulate news aggregators. On February 4, 2021 Canadian newspapers published a blank front page to protest news aggregators and convey how critical role their role is in news content generation (Gilmore, 2021). Simply put, the campaign communicated that the content available on Google News would not exist without their newsrooms. Heritage Minister, Steven Guilbeault has also employed a “content” frame by

emphasizing the value of news content when he stated, “news is not free and never has been” (Coulter, 2021). All of these regulatory efforts suggest the “content” frame has prevailed, raising the question why this is the case.

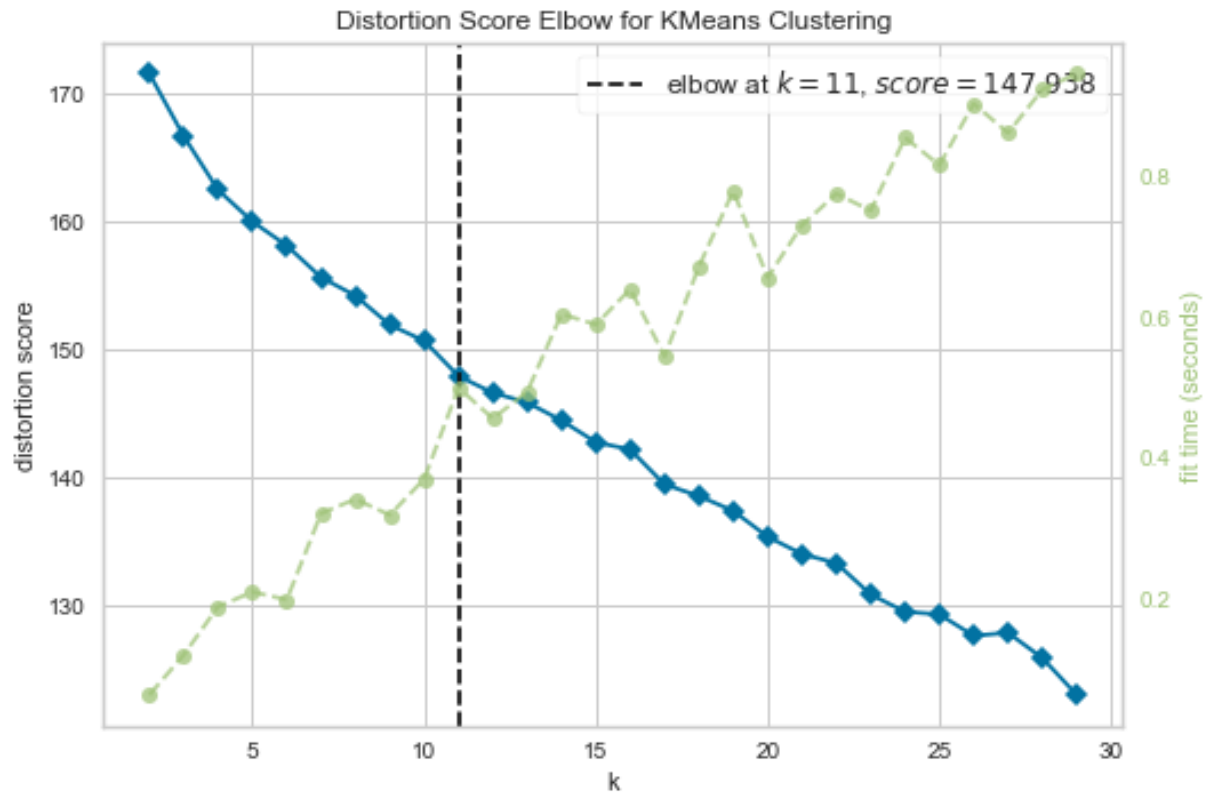
The reason for the success of the “content” frame can be attributed to the fact that Google does not create content itself. Unlike the “creator” frame which is rather inward looking and does not easily lend itself to arguments defending the public interest, the “content” frame has proven to be a powerful tool to reinforce how a healthy press publishing sector is necessary for creating the news content that informs the population. While it is certainly debatable if the application of competition law is appropriate and if a true market power imbalance exists between press publishers and news aggregators, what is not debatable is that the newsrooms create content and Google does not. This fact has carried stakeholders in AU and France much further than prior attempts to prove copyright infringement or defend the ownership of a particular written work, motivating the adoption of the “content” frame.

However, the implications associated with the abandonment of the “creator” frame must also be considered. While the twenty-first century has been termed the “content revolution,” there is an emerging next wave of self-published journalism focused on the creator. Led by new journalism enterprise models, journalists are leaving the mastheads of traditional press publications to reach their audiences directly. For example, Substack is a service that enables writers to draft, edit, and send email newsletters to free or paid subscribers on their own. Its’ value proposition is not grounded in content, but rather the creator itself. This intention is confirmed by Substack CEO, Chris Best, who aims “to allow writers and creators to run their own personal media empire” (Fatemi, 2021). The service currently has more than 250,000 paying subscribers with the top ten newsletters collectively generating over \$7 million in annualized revenue (ibid). The business model of Substack has proven to be quite enticing to creators as explained by Substack author, Judd

Legum, “it’s not about gaming the Google algorithm or the Facebook algorithm. Instead, it’s about writing compelling content that wins hearts and minds” (ibid). Moreover, many career journalists who have migrated to Substack have said to have earned more recognition and money than their previous reporting jobs (ibid). Many of the most popular newsletters on Substack share an attribute of newsworthiness including “Letters from an American” by Heather Cox Richardson who contextualizes the current political landscape in history, “The Bitcoin Forecast” by Willy Woo which offers Bitcoin pricing projections, and “Platformer” by Casey Newton which reports on social networks and democracy.

If the “creator” frame is abandoned to accommodate the “content” frame, then regulators may risk harming tomorrow’s wave of self-published journalistic endeavors. Arguably it is the niche, self-published creators and emerging journalism enterprises like Substack that are addressing the threats posed by Google News in a more crafty and inventive way than traditional press publishers relying on a paycheck from Google. So, while the “content” frame may be a mightier weapon against Google News today, this does not mean it won’t harm innovative efforts that rely on the freedom of the internet to distribute original reporting by creators. If there is anything that the history of copyright and the press has demonstrated, it is that protecting written works is a never-ending battle that requires a dynamic tactical approach. From the “Copyright Act of Queen Anne” to Murdoch’s accusations of theft against Google News, to the emergence of competition law regulation in AU, all of these developments reinforce the notion that how copyright is framed today may change tomorrow.

## Appendix 1. Elbow Plot



## Appendix 2. Cluster Membership

#	ID	Organization Name	Stakeholder Type	Gov
	15	European Visual Artists	Collective Management Organization	EU
	87	Norwegian Visual Artists Copyright Society	Collective Management Organization	EU
	109	Visual Copyright Society in Sweden	Collective Management Organization	EU
1	76	Association of Dutch Designers	Trade Association (Other)	EU
	81	Platform Makers	Writer	EU
	84	The Authors Guild of the Netherlands	Writer	EU
	114	Creators' Rights Alliance	Writer	EU
	208	Monash University: Margaret Simons	Education or Research Institution	AU
	133	ABC	Broadcaster	AU
	153	Walkley Foundation	Journalist	AU
2	178	Croakey Health Media	Scientific Publisher	AU
	137	Andrew Jaspán	Journalist	AU
	144	The Judith Neilson Institute for Journalism and Ideas	Education or Research Institution	AU
	229	Public Interest Journalism Initiative	Press Publisher	AU
	110	European Broadcasting Union	Broadcaster	EU
	33	Finnish Authors' Copyright Society	Collective Management Organization	EU
	85	Vrijschrift	Consumer	EU
	29	Denmarks Electronic Research Library	Education or Research Institution	EU
	34	Finnish IT Center for Science	Education or Research Institution	EU
	67	University of Dublin: Trinity College Library	Education or Research Institution	EU
	118	Joint Information Systems Committee	Education or Research Institution	EU
	40	Oulu University Library	Library/Cultural Heritage Institution	EU
	61	National Széchényi Library of Hungary	Library/Cultural Heritage Institution	EU
	63	Consortium of National & University Libraries	Library/Cultural Heritage Institution	EU
	44	Union of Finnish Writers	Writer	EU
	65	Local Ireland	Press Publisher	EU
3	95	Impresa	Press Publisher	EU
	111	Association of Photographers	Professional Photographer	EU
	116	Getty Images	Professional Photography	EU
	125	The British Photographic Council	Professional Photographer	EU
	106	Spanish Copyright Office	Public Authority	EU
	47	Explore Media	Startup	EU
	2	AMEC / FIBEP	Trade Association (Media)	EU
	48	Federation of Media Watch Companies	Trade Association (Media)	EU
	52	German Association for Law and Informatics	Trade Association (Technology)	EU
	16	European Writers' Council	Writer	EU
	44	Union of Finnish Writers	Writer	EU
4	158	Australian Copyright Council	NGO (Copyright)	AU
	161	Australian Film & TV Bodies	Broadcaster	AU

	190	Gladwin Legal	Legal Institutions	AU
	240	Stan	Technology Company	AU
4	233	QUT School of Law: Nicolas Suzor	Education or Research Institution	AU
	235	Redbubble	Technology Company	AU
	242	StartupAUS	Startup	AU
	160	Australian Digital Alliance	NGO (Copyright)	AU
	185	Foxtel	Broadcaster	AU
	21	The European Grouping of Societies of Authors and Composers	Collective Management Organization	EU
	97	Portuguese Society of Authors	Collective Management Organization	EU
	123	Publishers Licensing Society Limited	Collective Management Organization	EU
	13	European Federation of Journalists	Journalist	EU
5	45	Union of Journalists in Finland	Journalist	EU
	77	Dutch Journalist Association	Journalist	EU
	62	ProArt Hungarian Copyright Alliance	NGO (Copyright)	EU
	113	British Association of Picture Libraries and Agencies	Trade Association (Information Science)	EU
	59	Association of Private Law Lecturers	Trade Association (Other)	EU
	54	Martin Kroll	Writer	EU
	92	Polish Society of Authors and Composers	Writer	EU
	126	The Publishers Association	Book Publisher	EU
	19	International Federation of Reproduction Rights Organizations	Collective Management Organization	EU
	98	Visapress	Collective Management Organization	EU
	112	Authors Licensing and Collecting Society	Collective Management Organization	EU
	28	Charles University: Institute of Formal and Applied Linguistics	Education or Research Institution	EU
	11	European Copyright Society	NGO (Copyright)	EU
	23	Union of Publishers of Bulgaria	Press Publisher	EU
	25	Europapress Holding	Press Publisher	EU
	26	Stryria Media Group	Press Publisher	EU
6	36	Finnish Newspapers Association	Press Publisher	EU
	39	Medialitto	Press Publisher	EU
	93	The Polish Chamber of Press Publishers	Press Publisher	EU
	96	Portuguese Press Association	Press Publisher	EU
	101	Ringier Romania	Press Publisher	EU
	102	Ringier Axel Springer Slovakia	Press Publisher	EU
	115	European Publishers Council	Press Publisher	EU
	120	News Media Association	Press Publisher	EU
	122	Professional Publishers Association	Press Publisher	EU
	104	Podemos	Public Authority	EU
	55	Springer Nature	Scientific Publisher	EU
	117	International Association of Scientific, Technical and Medical Publishers	Scientific Publisher	EU
	9	EURIMAG	Trade Association (Other)	EU

	90	Creative Poland Association	Trade Association (Other)	EU
	7	DigitalEurope	Trade Association (Technology)	EU
	69	Anitec	Trade Association (Telecommunications)	EU
	42	The Finnish Association of Non-Fiction Writers	Writer	EU
	124	Society of Authors	Writer	EU
	157	Australian Communications and Media Authority	Public Authority	AU
	219	Office of the eSafety Commissioner	Public Authority	AU
	232	Queensland University and University of Sydney (Terry Flew and Tim Dwyer)	Education or Research Institution	AU
7	60	K-Monitor	NGO (Digital Rights)	EU
	132	5Rights Foundation	NGO (Digital Rights)	AU
	170	Centre for Media Transition	Education or Research Institution	AU
	228	Public Health Association of Australia	Consumer	AU
	227	Public Health Advocacy Institute of WA	Consumer	AU
	139	Association for Data-driven Marketing & Advertising	Advertising and Public Relations	AU
	231	Queensland Law Society	Legal Institutions	AU
	172	Communications Alliance	Trade Association (Media)	AU
	156	Australian Chamber of Commerce and Industry	Public Authority	AU
	175	Consumer Policy Research Centre	Consumer	AU
	220	Office of the Victorian Information Commissioner	Public Authority	AU
	136	American Bar Association Section of Antitrust Law	Legal Institutions	AU
	218	Office of the Australian information Commissioner	Public Authority	AU
	147	TotallyAwesome	Startup	AU
8	151	UNSW Sydney: Allens Hub for Technology, Law, and Innovation (Katharine Kemp & Rob Nicholls)	Education or Research Institution	AU
	149	U.S. Chamber of Commerce	Public Authority	AU
	150	UN Special Rapporteur on the Right to Privacy	Public Authority	AU
	152	UNSW Sydney: Kayleen Manwaring	Education or Research Institution	AU
	164	Australian Privacy Foundation	NGO (Digital Rights)	AU
	174	Consumer Action Law Centre	Consumer	AU
	201	Internet of Things Alliance Australia	Trade Association (Telecommunications)	AU
	206	Melbourne Law School: Jeannie Marie Paterson	Education or Research Institution	AU
	168	BSA The Software Alliance	Trade Association (Technology)	AU
	197	Iconoclast Tech	Education or Research Institution	AU
	222	Oracle	Technology Company	AU
	159	Australian Data Privacy Certification Register	NGO (Digital Rights)	AU
	173	Computer & Communications Industry Association	Trade Association (Technology)	AU
	142	Tania Katsanis	Startup	AU
	221	oOhlmedia	Advertising and Public Relations	AU
9	148	Twitter	Technology Company	AU
	198	Innovation Defense Foundation	Trade Association (Technology)	AU
	202	Intran Australia	NGO (Digital Rights)	AU

	226	Provincial Press Group	Press Publisher	AU
	163	Australian Press Council	Press Publisher	AU
	162	Australian Lottery and Newsagents' Association	Press Publisher	AU
	214	NewsMediaWorks	Press Publisher	AU
	215	Nine	Broadcaster	AU
	181	DuckDuckGo	Technology Company	AU
	183	Fairfax Media	Press Publisher	AU
	184	Food4U	Startup	AU
	186	Free TV	Trade Association (Media)	AU
	193	Guardian Australia	Press Publisher	AU
	195	Harley Comrie	Journalist	AU
	199	International Center for Law and Economics	Education or Research Institution	AU
	212	News & Media Research Centre	Education or Research Institution	AU
	239	Special Broadcasting Service	Broadcaster	AU
	166	Australian Radio Network	Broadcaster	AU
	182	Facebook	Technology Company	AU
	241	Star News Group	Press Publisher	AU
	138	Arnold Bloch Leibler	Legal Institutions	AU
9	177	Country Press Australia	Press Publisher	AU
	204	McPherson Media Group	Press Publisher	AU
	224	Post Newspapers	Press Publisher	AU
	145	The North Western Courier	Press Publisher	AU
	134	Access Now	NGO (Digital Rights)	AU
	146	Tony Healy	Technology Company	AU
	189	Getty Images	Professional Photography	AU
	141	Australian Associated Press	Press Publisher	AU
	180	Digital Industry Group Incorporated	Trade Association (Technology)	AU
	191	Global Antitrust Institute	Education or Research Institution	AU
	200	Internet Australia	Trade Association (Telecommunications)	AU
	210	Netflix	Technology Company	AU
	211	Network Ten	Broadcaster	AU
	213	News Corp Australia	Press Publisher	AU
	216	NSW Business Chamber	Public Authority	AU
	217	NSW Young Lawyers	Legal Institutions	AU
	223	Outdoor Media Association	Press Publisher	AU
	225	Progressive Policy Institute	Education or Research Institution	AU
	116	Getty Images	Professional Photography	AU
	135	Ad Standards	Public Authority	AU
	143	The Communications Council	Advertising and Public Relations	AU
	203	James Meese	Education or Research Institution	AU
	205	Media, Entertainment and Arts Alliance	Trade Association (Media)	AU
	207	Microsoft	Technology Company	AU



	237	Simon Thompson	Legal Institutions	AU
	238	Southern Cross Austereo	Press Publisher	AU
	165	Australian Publishers Association	Press Publisher	AU
	171	Commercial Radio Australia	Broadcaster	AU
	230	Queensland Country Press Association	Press Publisher	AU
	169	Carol O'Donnell	Education or Research Institution	AU
	179	David Fagan	Journalist	AU
	194	Guillaume Roger	Education or Research Institution	AU
	209	National Association for the Visual Arts	Trade Association (Other)	AU
	234	REA Group	Advertising and Public Relations	AU
	236	Seven West Media	Press Publisher	AU
	140	Audited Media Association of Australia	Trade Association (Media)	AU
	192	Google	Technology Company	AU
	196	Harvard Kennedy School: Mitchell Watt & Hubert Wu	Education or Research Institution	AU
	155	Australian Associated Press	Press Publisher	AU
	167	Australian Society of Authors	Writer	AU
	187	Freedom Publishers Union	Press Publisher	AU
	188	George Mason University: Mercatus Center (Christine McDaniel & Danielle Parks)	Education or Research Institution	AU
	154	Worldview Exchange	Technology Company	AU
	66	Public Relations Institute of Ireland	Advertising and Public Relations	EU
	130	Microsoft Press	Book Publisher	EU
	56	VG Media	Collective Management Organization	EU
	82	SURFmarket	Collective Management Organization	EU
	4	COMMUNIA International Association on the Public Domain	Consumer	EU
	10	European Consumer Organization	Consumer	EU
	12	European Digital Rights	Consumer	EU
	68	Altroconsumo	Consumer	EU
	99	Association for Technology and Internet	Consumer	EU
	128	Center for Democracy & Technology	Consumer	EU
10	0	University of Vienna: Media Governance and Industries Research Lab	Education or Research Institution	EU
	53	Helmholtz Open Science Coordination Office	Education or Research Institution	EU
	73	Nexa Center for Internet & Society	Education or Research Institution	EU
	94	Wroclaw University of Technology	Education or Research Institution	EU
	127	University of Strathclyde	Education or Research Institution	EU
	27	Association of Library and Information Professionals of the Czech Republic	Library/Cultural Heritage Institution	EU
	31	The Danish Library Association	Library/Cultural Heritage Institution	EU
	32	Council for Finnish University Libraries	Library/Cultural Heritage Institution	EU
	35	Finnish Library Association	Library/Cultural Heritage Institution	EU
	38	Helsinki University Library	Library/Cultural Heritage Institution	EU
	43	The National Library of Finland	Library/Cultural Heritage Institution	EU

78	European Bureau of Library, Information and Documentation Associations	Library/Cultural Heritage Institution	EU
79	FOBID Netherlands Library Forum	Library/Cultural Heritage Institution	EU
83	The Association of European Research Libraries	Library/Cultural Heritage Institution	EU
89	Cracow University of Technology Library	Library/Cultural Heritage Institution	EU
108	The International Association of Music Libraries, Archives and Documentation Centres	Library/Cultural Heritage Institution	EU
20	Partnership for Copyright & Society (SA&S)	NGO (Copyright)	EU
80	Kennisland	NGO (Copyright)	EU
119	Libraries and Archives Copyright Alliance	NGO (Copyright)	EU
22	Wikimedia	NGO (Digital Rights)	EU
24	OpenMedia	NGO (Digital Rights)	EU
91	Modern Poland Foundation	NGO (Digital Rights)	EU
107	Spanish Digital Economy Association	NGO (Digital Rights)	EU
121	OpenForum Europe	NGO (Digital Rights)	EU
57	yeebase media GmbH	Press Publisher	EU
86	Scoop Foundation for Public Interest Journalism	Press Publisher	EU
88	300Polityka	Press Publisher	EU
10	105 Spanish Association of Periodical Publishers	Press Publisher	EU
	1 Allied for Startups	Startup	EU
	30 Doable	Startup	EU
103	Association of Spanish Startups	Startup	EU
179	David Fagan	Journalist	EU
129	Google	Technology Company	EU
131	Mozilla	Technology Company	EU
	6 Copyright for Creativity	Trade Association (Information Science)	EU
100	Kosson	Trade Association (Information Science)	EU
	5 Computer and Communications Industry Association	Trade Association (Technology)	EU
	8 EDiMA	Trade Association (Technology)	EU
41	The Federation of Finnish Technology Industries	Trade Association (Technology)	EU
64	ICT Ireland	Trade Association (Technology)	EU
14	European Internet Services Providers Association	Trade Association (Telecommunications)	EU
70	Assinform	Trade Association (Telecommunications)	EU
71	Confindustria Digitale	Trade Association (Telecommunications)	EU
72	Employers Association of Telecommunication Operators	Trade Association (Telecommunications)	EU
	3 Association of Belgian Publishers	Book Publisher	EU
	17 Federation of European Publishers	Book Publisher	EU
	18 Flemish Book Publishers Association	Book Publisher	EU
11	46 Actes Sud	Book Publisher	EU
	49 French Publishers Association	Book Publisher	EU

	50	Hachette Livre	Book Publisher	EU
	51	Les Editions du Rouergue	Book Publisher	EU
11	58	Association of Greek Publishers and Booksellers	Book Publisher	EU
	74	The Italian Publishers Association	Book Publisher	EU
	75	Copyright and Communication Consulting Agency / Latvian Authors Association	Collective Management Organization	EU

---

## References

- ACCC. (2018). Digital Platforms Inquiry Issues Paper. Australian Competition and Consumer Commission. [https://www.accc.gov.au/system/files/DPI%20-%20Issues%20Paper%20-%20Vers%20for%20Release%20-%202025%20F..\\_%20%28006%29.pdf](https://www.accc.gov.au/system/files/DPI%20-%20Issues%20Paper%20-%20Vers%20for%20Release%20-%202025%20F.._%20%28006%29.pdf)
- ACCC. (2020, July 30). *Draft news media bargaining code* [Text]. Australian Competition and Consumer Commission. <https://www.accc.gov.au/focus-areas/digital-platforms/draft-news-media-bargaining-code>
- Athey, S., Mobius, M., & Pal, J. (n.d.). *The Impact of Aggregators on Internet News Consumption*. 37.
- AU Google Submission. (2018)  
<https://www.dropbox.com/s/acsjfgmw2y0cs0o/AU%20Google%20Submission.pdf?dl=0>
- AU News Corp Submission. (2018).  
<https://www.dropbox.com/s/42q446e3yi9oqre/AU%20Selected%20Submissions.pdf?dl=0>
- AU Selected Submissions. (2019).  
<https://www.dropbox.com/s/402f55byt47solq/AU%20Selected%20Submissions.pdf?dl=0>
- Auchard, E. (2007, April 7). AFP, Google News settle lawsuit over Google News. *Reuters*. <https://www.reuters.com/article/us-google-afp-idUSN0728115420070407>
- Baldwin, P. (2014). The Battle between Anglo-American Copyright and European Authors' Rights. In *The Copyright Wars* (pp. 14–52). Princeton University Press. <https://doi.org/10.2307/j.ctt6wq0z2.4>
- Beddoe, K. (2021, February 4). Google News Showcase launches in Australia. Google.  
<https://blog.google/products/news/google-news-showcase-launches-australia/>
- Bharat, K. (2011, May 6). Google News and the Coverage of Bin Laden. Google News Blog.  
<https://news.googleblog.com/2011/05/google-news-and-coverage-of-bin-laden.html>
- Bochkay, K., Chychyla, R., & Leone, A. (2020). Using Python for Text Analysis in Accounting Research. *Foundations and Trends in Accounting*, 223.
- Bracha, O. (2010). The Adventures of the Statute of Anne in the Land of Unlimited Possibilities: The Life of a Legal Transplant. *Berkeley Technology Law Journal*, 25(3), 1427–1473.
- Calzada, J., & Gil, R. (2020). What Do News Aggregators Do? Evidence from Google News in Spain and Germany. *Marketing Science*, 39(1), 134–167. <https://doi.org/10.1287/mksc.2019.1150>
- Chiou, L., & Tucker, C. (2017). Content aggregation by platforms: The case of the news media. *Journal of Economics & Management Strategy*, 26(4), 782–805. <https://doi.org/10.1111/jems.12207>
- Chyi, H. I., & Sylvie, G. (1998). Competing With Whom? Where? And How? A Structural Analysis of the Electronic Newspaper Market. *Journal of Media Economics*, 11(2), 1–18.  
[https://doi.org/10.1207/s15327736me1102\\_1](https://doi.org/10.1207/s15327736me1102_1)

- CompTIA. (2020, May 20). U.S. Exports of Technology Totaled Close to \$340 Billion in 2019, Supporting 918,500 Jobs, According to CompTIA. PR Newswire. <https://www.prnewswire.com/news-releases/us-exports-of-technology-totaled-close-to-340-billion-in-2019--supporting-918-500-jobs-according-to-comptia-301062686.html>
- Coulter, M. (2021). Facebook Faces News Battles in 7 Other Countries After Australia. *Business Insider*. <https://www.businessinsider.com/facebook-australia-news-ban-battles-7-other-countries-2021-2#canada-has-promised-to-introduce-new-regulations-2>
- Cozens, C. (2005, March 21). AFP sues Google over copyrighted content. *The Guardian*. <http://www.theguardian.com/technology/2005/mar/21/media.newmedia>
- Crampton, T. (2007, February 13). Brussels court rules against Google in copyright case— Business— International Herald Tribune (Published 2007). *The New York Times*. <https://www.nytimes.com/2007/02/13/business/worldbusiness/13iht-google.4584306.html>
- Dellarocas, C., Sutanto, J., Calin, M., & Palme, E. (2016). Attention Allocation in Information-Rich Environments: The Case of News Aggregators. *Management Science*, 62(9), 2543–2562. <https://doi.org/10.1287/mnsc.2015.2237>
- Dusollier, S. (2020). The 2019 Directive on Copyright in the Digital Single Market: Some Progress, a Few Bad Choices, and an Overall Failed Ambition (SSRN Scholarly Paper ID 3695839). Social Science Research Network. <https://papers.ssrn.com/abstract=3695839>
- Entman, R. M. (1991). Framing U.S. Coverage of International News: Contrasts in Narratives of the KAL and Iran Air Incidents. *Journal of Communication*, 41(4), 6–27. <https://doi.org/10.1111/j.1460-2466.1991.tb02328.x>
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Entman, R. M. (2010). Media framing biases and political power: Explaining slant in news of Campaign 2008. *Journalism*, 11(4), 389–408. <https://doi.org/10.1177/1464884910367587>
- Espinoza, J., & Barker, A. (2021). EU ready to follow Australia's lead on making Google and Facebook pay for news. *Los Angeles Times*. <https://www.latimes.com/world-nation/story/2021-02-08/eu-australia-google-facebook-tech-news>
- EU Google Submission. (2016). <https://www.dropbox.com/s/1b80fvgb7efjnqv/EU%20Google%20Submission.pdf?dl=0>
- EU Selected Submissions. (2016). <https://www.dropbox.com/s/y28knoch41ikvc/EU%20Selected%20Submissions.pdf?dl=0>

- European Commission. (2015, May 6). *A Digital Single Market for Europe: Commission sets out 16 initiatives to make it happen* [Text]. European Commission. [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_15\\_4919](https://ec.europa.eu/commission/presscorner/detail/en/IP_15_4919)
- European Commission. (2018, June 22). Frequently Asked Questions on Copyright Reform [Text]. European Commission. <https://ec.europa.eu/digital-single-market/en/faq/frequently-asked-questions-copyright-reform>
- European Council. (2019, April 15). *EU adjusts copyright rules to the digital age*. European Council - Council of the European Union. <https://www.consilium.europa.eu/en/press/press-releases/2019/04/15/eu-adjusts-copyright-rules-to-the-digital-age/>
- Fatemi, F. (2021, January 21). The Rise Of Substack—And What’s Behind It. *Forbes*. <https://www.forbes.com/sites/falonfatemi/2021/01/20/the-rise-of-substack-and-whats-behind-it/>
- Fortt, J. (2008, April 22). *What’s Google News worth? \$100 million*. *Fortune*. <https://fortune.com/2008/07/22/whats-google-news-worth-100-million/>
- Frosio, G. (2020). Reforming the C-DSM Reform: A User-Based Copyright Theory for Commonplace Creativity. *IIC - International Review of Intellectual Property and Competition Law*, 51(6), 709–750. <https://doi.org/10.1007/s40319-020-00931-0>
- Gilmore, R. (2021, February 4). A world without news: Papers publish blank front pages to highlight industry struggles—National | Globalnews.ca. *Global News*. <https://globalnews.ca/news/7617779/canada-news-industry-newspapers-blank-struggles-journalism/>
- Harms, L. T. C., & World Intellectual Property Organization. (2018). *A casebook on the enforcement of intellectual property rights*.
- Huang, J. S., Yang, M. J., & Chyi, H. I. (2013). Friend or foe? Examining the relationship between news portals and newspaper sites in Taiwan. *Chinese Journal of Communication*, 6(1), 103–119. <https://doi.org/10.1080/17544750.2012.753502>
- Isbell, K. A. (2010). *The Rise of the News Aggregator: Legal Implications and Best Practices* (SSRN Scholarly Paper ID 1670339). Social Science Research Network. <https://doi.org/10.2139/ssrn.1670339>
- Jacie Yang, M., & Iris Chyi, H. (2011). Competing with Whom? Where? and Why (Not)? an Empirical Study of U.S. Online Newspapers’ Competition Dynamics. *Journal of Media Business Studies*, 8(4), 59–74. <https://doi.org/10.1080/16522354.2011.11073531>
- Jeon, D.-S. (2018). *Economics of News Aggregators*. 22.
- Keller, P. (2020, April 30). The press publishers right will fail—To support the media we should tax information aggregators instead. A Shared Digital Europe. <https://shared-digital.eu/as-predicted-the-new-press-publishers-right-is-a-failure-lets-make-information-aggregators-pay-news-media-producers-for-real/>

- Khan, L. M. (2016). Amazon's Antitrust Paradox. *Yale Law Journal*, 126(3), 710–805.
- Klobuchar, A. (2021). Senator Klobuchar and Representative Cicilline Introduce Legislation to Protect Journalism in the United States. U.S. Senator Amy Klobuchar. <https://www.klobuchar.senate.gov/public/index.cfm/2021/3/senator-klobuchar-and-representative-cicilline-introduce-legislation-to-protect-journalism-in-the-united-states>
- Klüver, H., Mahoney, C., & Opper, M. (2015). Framing in context: How interest groups employ framing to lobby the European Commission. *Journal of European Public Policy*, 22(4), 481–498. <https://doi.org/10.1080/13501763.2015.1008550>
- Lee, A. M., & Chyi, H. I. (2015). The Rise of Online News Aggregators: Consumption and Competition. *International Journal on Media Management*, 17(1), 3–24. <https://doi.org/10.1080/14241277.2014.997383>
- Lomas, N. (2021, January 21). Google inks agreement in France on paying publishers for news reuse. TechCrunch. <https://social.techcrunch.com/2021/01/21/google-inks-agreement-in-france-on-paying-publishers-for-news-reuse/>
- Lomas, N. (2021, February 25). After Facebook's news flex, Australia passes bargaining code for platforms and publishers. TechCrunch. <https://social.techcrunch.com/2021/02/25/after-facebooks-news-flex-australia-passes-bargaining-code-for-platforms-and-publishers/>
- Mullooly, O. (2020, March 6). The new press publishers' right: Will Big Tech push back? Arthur Cox. <https://www.arthurcox.com/knowledge/the-new-press-publishers-right-will-big-tech-push-back/>
- Murdoch, R. (2009, December 8). Journalism and Freedom. *Wall Street Journal*. <https://www.wsj.com/articles/SB10001424052748704107104574570191223415268>
- Newman, N. (2020). *Reuters Institute Digital News Report 2020*. 112.
- Picard, R. G. (1989). *Media economics: Concepts and issues* (Vol. 22). SAGE Publications, Incorporated.
- Román, D. (2014, December 16). Google's Shutdown of Spanish News Service Watched Elsewhere in Europe. *Wall Street Journal*. <https://online.wsj.com/articles/google-shuts-spanish-news-service-ahead-of-new-law-1418728149>
- Rose, B. (2014, December 24). Document Clustering with Python. GitHub. <http://brandonrose.org/clustering?ref=dzone#Tf-idf-and-document-similarity>
- Schonhardt-Bailey, C. (2005). Measuring Ideas More Effectively: An Analysis of Bush and Kerry's National Security Speeches. *PS: Political Science and Politics*, 38(4), 701–711.
- Scott, M. (2016, September 13). E.U. Rules Look to Unify Digital Market, but U.S. Sees Protectionism (Published 2016). *The New York Times*. <https://www.nytimes.com/2016/09/14/technology/eu-us-tech-google-facebook-apple.html>

- Searle, N. (2020). Business Models and Copyright Reform: The Legal Business Model. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3599891>
- Senftleben, M., Kerk, M., Buiten, M., & Heine, K. (2017). New Rights or New Business Models? An Inquiry into the Future of Publishing in the Digital Era. *IIC - International Review of Intellectual Property and Competition Law*, 48(5), 538–561. <https://doi.org/10.1007/s40319-017-0605-y>
- Sterling, G. (2011, July 16). *Belgian Newspapers Claim Retaliation By Google After Copyright Victory*. Search Engine Land. <https://searchengineland.com/belgian-newspapers-claim-retaliation-by-google-after-copyright-victory-85924>
- Sullivan, D. (2009, November 24). *Under The Hood: Google News & Ranking Stories*. Search Engine Land. <https://searchengineland.com/google-news-ranking-stories-30424>
- The Evolution of Search*. (2011, November 27). [https://www.youtube.com/watch?feature=player\\_embedded&v=mTBShTwCnD4&t=122](https://www.youtube.com/watch?feature=player_embedded&v=mTBShTwCnD4&t=122)
- Tryhorn, C. (2009, December 3). We're no thieves – despite what Rupert Murdoch claims, says Google. *The Guardian*. <https://www.theguardian.com/technology/2009/dec/03/were-no-thieves-murdoch-google>
- Upstill, T. (2018, May 8). *The new Google News: AI meets human intelligence*. [https://blog.google/products/news/new-google-news-ai-meets-human-intelligence/?utm\\_source=tw&utm\\_medium=feed&utm\\_campaign=io18](https://blog.google/products/news/new-google-news-ai-meets-human-intelligence/?utm_source=tw&utm_medium=feed&utm_campaign=io18)
- Wolff, M., & Wolff, M. (2008). *The man who owns the news: Inside the secret world of Rupert Murdoch* (1st ed.). Broadway Books.