UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ARMA 2022

June 29-30, 2022 Valencia, Spain

4th International Conference on Advanced Research Methods and Analytics

ADE FACULTAD DE ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS. UPV

Departamento de Economía y Ciencias Sociales

Ivie

devstat

Congress UPV

4th International Conference on Advanced Research Methods and Analytics (CARMA 2022)

The contents of this publication have been evaluated by the Program Committee according to the procedure described in the preface. More information at http://www.carmaconf.org/

Scientific Editors

Josep Domenech
María Rosalía Vicente

Publisher

2022, Editorial Universitat Politècnica de València
www.lalibreria.upv.es / Ref.: 6106_01_01_01

Cover design by Gaia Leandri

ISBN: 978-84-1396-018-0 (print version)
Print on-demand

ISSN: 2951-9748

DOI: http://dx.doi.org/10.4995/CARMA2022.2022.15956

4th International Conference on Advanced Research Methods and Analytics (CARMA 2022)
This book is licensed under a Creative Commons Atribution-NonCommercial-ShareAlique-4.0 International license
Editorial Universitat Politècnica de València http://ocs.editorial.upv.es/index.php/CARMA/CARMA2022

Preface

Josep Domenech [1], María Rosalía Vicente [2]
[1] Dept. Economics and Social Sciences, Universitat Politècnica de València, Spain. [2] Dept. Applied Economics, Universidad de Oviedo, Spain

Abstract

Research methods in economics and social sciences are evolving with the increasing availability of Internet and Big Data sources of information. As these sources, methods, and applications become more interdisciplinary, the 4th International Conference on Advanced Research Methods and Analytics (CARMA) is an excellent forum for researchers and practitioners to exchange ideas and advances on how emerging research methods and sources are applied to different fields of social sciences as well as to discuss current and future challenges. This edition was celebrated in a hybrid format because of the uncertainties created by pandemic.

Keywords: *Big Data sources, Web scraping Social media mining, Official Statistics, Internet Econometrics, Digital transformation, global society.*

This work is licensed under a Creative Commons License CC BY-NC-ND 4.0
Editorial Universitat Politècnica de València

1. Preface to CARMA2022

This volume contains the selected papers of the Fourth International Conference on Advanced Research Methods and Analytics (CARMA 2022) hosted by the Universitat Politècnica de València, Spain during 29 and 30 June 2022. This fourth edition consolidates CARMA as a unique forum where Economics and Social Sciences research meets Internet and Big Data. CARMA provides researchers and practitioners with an ideal environment to exchange ideas and advances on how Internet and Big Data sources and methods contribute to overcome challenges in Economics and Social Sciences, as well as on the changes in the society after the digital transformation.

The selection of the scientific program was directed by Maria Rosalia Vicente, who led an international team of 47 scientific committee members representing institutions worldwide. Following the call for papers, the conference received 58 paper submissions from all around the globe. All submissions were reviewed by the scientific committee members under a double-blind review process. Finally, 35 papers were accepted for oral presentation during the conference, ensuring a high-quality scientific program. It covers a wide range of research topics on the Internet and Big Data, including public opinion mining, web scraping, search engine data, tourism and mobility, social behavior, data economy, or marketing and social media, among others. Additionally, 11 papers with promising work-in-progress research were selected for presentation during the conference.

The scientific program includes two keynote speakers that will review the state-of-the-art techniques and applications of the Internet and Big Data. The first keynote address is given by Giuliano Resce (University of Molise, Italy) to overview the latest digital methods for Economics and Social Sciences. The second keynote speech is delivered by Fabian Braesemann (Oxford Internet Institute, UK) and deals with the Social Data Science in the Digital Economy.

CARMA 2022 also featured two special sessions on “Big Data in Central Banks” and “Internet and Big Data in Official Statistics,” chaired by Juri Marcucci and Aidan Condron, respectively. Both sessions gave a complementary institutional perspective on how to use the Internet and Big Data sources and methods for public policy and official statistics.

The conference organizing committee would like to thank all who made this fourth edition of CARMA a great success. Specifically, thanks are indebted to the authors, scientific committee members, special session organizers, invited speakers, session chairs, reviewers, presenters, sponsors, supporters, and all the attendees. Our final words of gratitude must go to the Faculty of Business Administration and Management of the Universitat Politècnica de València for supporting CARMA 2022.

2. Organizing Committee

General chair

Josep Domènech, Universitat Politècnica de València

Scientific committee chair

María Rosalía Vicente, Universidad de Oviedo

Local organization

Eduardo Cebrián Cerdá
Pilar Malagón Selma
Héctor Martínez Cabanes
Joan Manuel Valenzuela

3. Sponsors and Supporters

Universitat Politècnica de València
Facultad de Administración y Dirección de Empresas
Departamento de Economía y Ciencias Sociales
Instituto Valenciano de Investigaciones Económicas (IVIE)
DevStat — Development of Statistics. Statistics for Development

4. Scientific committee

Anto Aasa, University of Tartu
Fernando Almeida, University of Porto & INESC TEC
José A. Álvarez-Jareño, University of Valencia
María del Pilar Ángeles, Universidad Nacional Autónoma de México
Concha Artola, Banco de España
Nikolaos Askitas, IZA – Institute of Labor Economics
Seyhmus Baloglu, University of Nevada
Catherine Beaudry, Polytechnique Montreal
Silvia Biffignandi, Consultant Economic Statistics Studies (ESS)
Ludovic Calès, European Commission, Italy
Roger H.L. Chiang, University of Cincinnati
Cihan Cobanoglu, University of South Florida
Marisol B. Correia, ESGHT-Universidade do Algarve & CiTUR
Lisa Crosato, Ca' Foscari University of Venice
Pablo de Pedraza Garcia, Italy
Giuditta de Prato, EC JRC, Spain

Carlo Drago, University "Niccolò Cusano", Rome
Mohammad Falahat, Universiti Tunku Abdul Rahman (UTAR)
Juan Fernández de Guevara, Universitat de Valencia & Ivie
Rui Gaspar, Universidade Católica Portuguesa
Yolanda Gomez, DevStat
Marcos González-Fernández, Universidad de León
Peter Hackl, Vienna University of Economics and Business
Agustín Indaco, Carnegie Mellon University, Qatar
Jan Kinne, ZEW Mannheim
Marius Leckelt, University of Muenster
Caterina Liberati, University of Milano-Bicocca
Juri Marcucci, Bank of Italy
Rocío Martínez-Torres, University of Seville
Amir Mosavi, Obuda University
María Olmedilla, SKEMA Business School
Enrique Orduña-Malea, Universitat Politècnica de València
José Luis Ortega, CSIC
Viktor Pekar, OIM, Aston University
Arturo Peralta Martín-Palomino, Universidad de Castilla la Mancha
Maria Petrescu, Embry-Riddle Aeronautical University
Kostas E. Psannis, University of Macedonia
Pilar Rey del Castillo, Instituto de Estudios Fiscales
Rosa Rio-Belver, Universidad del País Vasco UPV/EHU
Anna Rosso, Università degli Studi dell'Insubria
Zhaohao Sun, PNG University of Technology
Stoyan Tanev, Carleton University
Sergio Toral Marin, Universidad de Sevilla
Konstantinos Tsagarakis, Technical University of Crete
Tiziana Tuoto, Istat and Sapienza University of Rome
Antonino Virgillito, Agenzia delle Entrate – Italian Revenue Agency
Martin R. Wolf, FH Aachen University of Applied Sciences

Index

Full papers

A visual analysis of the literature on Internet neutrality 1
Lucia Pinar Garcia, Klaudijo Klaser

Ethical Behavior and Legal Regulations in Artificial Intelligence 9
Thomas Hauer

The effects of the e-tailer's reputation, the e-tailer's familiarity, and the relevance of the e-tailer's social media communication on impulse buying 15
Yosra Akrimi

Exploring Redditors' Topics with Natural Language Processing 25
Yilang Zhao

Exploring Redditors' Communication Style 33
Yilang Zhao

Google Trends Search Information Related to Breastfeeding in the U.S. 41
Richard A. Fabes, Denise Ann Bodman, Bethany Bustamante Van Vleet, Carol L Martin

News versus Corporate Reputation: Measuring through Sentiment and financial analysis 49
Naiara Pikatza-Gorrotxategi, Izaskun Alvarez-Meaza, Rosa María Río-Belver, Enara Zarrabeitia-Bilbao

Cape Town road traffic accident analysis: Utilising supervised learning techniques and discussing their effectiveness 57
Sebnem Er, Christo Du Toit, Sulaiman Salau

Can unlisted firms benefit from market information? A data-driven approach 65
Alessandro Bitetto, Stefano Filomeni, Michele Modina

Analysis of Wellness Experiences in a Tourist Destination 73
Lourdes Cauzo-Bottala, Francisco Javier Quirós-Tomás, Myriam González-Limón, María Del Rocío Martínez-Torres

The effect on purchase intention of social media influencers recommendations 81
Miguel Gonzalez-Mohino, L. Javier Cabeza-Ramirez

What are Gen Z's and Millennials' opinions on Masculinity in Advertising: a Qualitative Research Study 91
Toms Kreicbergs, Deniss Ščeulovs

Influence of popularity on the transfer fees of football players..... 101
Pilar Malagón-Selma, Ana Debón, Josep Domenech

Study of e-commerce trends based on customer characteristics in Latvia 109
Igors Babics, Rosita Zvirgzdiņa

Social Desirability and the Willingness to Provide Social Media Accounts in Surveys. The Case of Environmental Attitudes 119
Beate Klösch, Markus Hadler, Markus Reiter-Haas, Elisabeth Lex

Evaluation of the use of influencers for the development of consumer satisfaction in the Baltic consumer goods market 129
Iveta Linina, Velga Vevere, Rosita Zvirgzdina

Collaborate for what: a structural topic model analysis on CDP data 139
Camilla Salvatore, Alice Madonna, Annamaria Bianchi, Albachiara Boffelli, Matteo Giacomo Maria Kalchschmidt

Text mining methods for innovation studies: limits and future perspectives 147
Pietro Cruciata, Davide Pulizzotto, Catherine Beaudry

The demand side of information provision: Using multivariate time series clustering to construct multinational uncertainty proxies 155
Florian Schütze

Machine Learning and MADIT methodology for the fake news identification: the persuasion index..... 165
Gian Piero Turchi, Luisa Orrù, Christian Moro, Marco Cuccarini, Monia Paita, Marta Silvia Dalla Riva, Davide Bassi, Giovanni Da San Martino, Nicolò Navarin

Monitoring the survival of subscribers in a marketing mailing list..... 173
Andrea Marletta

Digital Ethnography Redux: Interpreting Drone Cultures and Microtargeting in an era of Digital Transformation.....181
David Beesley, Gavin Mount

Applying NLP techniques to characterize what makes an online review trustworthy 189
María Olmedilla Fernández, José Carlos Romero, Rocío Martínez-Torres, Sergio Toral

Emergency Calls in the City of Vaughan (Canada) During the COVID-19 Pandemic: A Spatiotemporal Analysis.....197
Adriano O. Solis, Ali Asgary, Nawar Khan, Janithra Wimaladasa, Maryam S. Sabet

Implementing sentiment analysis to an open-ended questionnaire: Case study of digitalization in elderly care during COVID-19.....205
Ida Toivanen, Venla Räsänen, Jari Lindroos, Tomi Oinas, Sakari Taipale

Changes in corporate websites and business activity: automatic classification of corporate webpages.....213
Joan Manuel Valenzuela Rubilar, Josep Domenech, Ana Pont

Understanding the effects of Covid-19 on P2P hospitality: Comparative classification analysis for Airbnb-Barcelona. 221
Juan Pablo Argente Del Castillo Martínez, Isabel P. Albaladejo

Simulating the inconsistencies of Google Trends data.....229
Eduardo Cebrián, Josep Domenech

Covid 19 and lodging places 237
Estefania Ruiz-Martinez, Francisco Porras-Bernardez, Georg Gartner

Cracking the Code of Geo-Identifiers: Harnessing Data-Based Decision-Making for the Public Good.....245
Patricia Snell Herzog

Non-conventional data and default prediction: the challenge of companies' websites 253
Lisa Crosato, Josep Domenech, Caterina Liberati

Automated Information Retrieval from the Bibliographic Metadata: A Way to Facilitate the Systematic Literature Review.....259
Marie Vítová Dušková, Martin Víta

Abstracts

Leveraging mobile network data to understand pandemic-era population human mobility 267
Aidan Condron

Big Data and Official Statistics: Challenges and Applications at Statistics Netherlands...268
Piet J. H. Daas

Quality Guidelines for the Acquisition and Usage of Big Data with additional Insights on Web Data.....269
Alexander Kowarik, Magdalena Six

Suggested Framework for Big Data Analysis of Enterprise Websites. A Case Study for Web Intelligence Network.....270
Jacek Maslankowski, Dominika Nowak

The Web Intelligence Hub – A tool for integrating web data in Official Statistics.....271
Fernando Reis

Exploration and experience with new web data sources. A Case Study for innovative tourism statistics.....272
Galya Stateva, Marek Cierpial-Wolan

An interpretable machine learning workflow with an application to economic forecasting.....275
Marcus Buckmann, Andreas Joseph

Textual analysis of a Twitter corpus during the COVID-19 pandemics 276
Valerio Astuti, Marta Crispino, Marco Langiulli, Juri Marcucci

Opportunities and risks in the residential sector during a green transition: House prices, energy renovations and rising energy prices.....277
Alessandro T. Martinello, Niels F. Møller

Assessing the green transition priorities of SMEs: A large scale web mining approach....278
Josep Domenech, Maria Rosalia Vicente, Hector Martinez Cabanes, Pablo De Pedraza

Unsupervised Learning for the Analysis and Detection of Fraud in the Insurance Industry 279
José A. Alvarez-Jareño, José Manuel Pavía

Policy indicators from private online platforms.....280
Jose Vila, Jose Luis Cervera-Ferri, Yolanda Gomez

Investigating mechanisms for compensating for an inability to touch products: the role of brand and situational involvement 281
Lili Zheng, Michel Plaisent, Prosper Bernard

Report on Amazon´s Project: Statistical evaluation on socio-economic variables across Germany 282
Sebastian De La Serna

Research on the Construction of Agro-Ecological Park Under the Background of Smart Agriculture 283
Yingying Cai, Jian Chen, Yunhan Gaoge, Haiyan Lv

Analyzing the Natural Language Processing technology field using Tech mining 284
Gaizka Garechana, Rosa María Río-Belver, Izaskun Álvarez-Meaza, Enara Zarrabeitia

Análisis bibliométrico de la economía experimental y progreso del campo de la investigación 285
Myriam González-Limón, Asunción Rodríguez-Ramos, Cristina Maldonado

Evaluating E-Learning systems success to understand student's performance during Covid Pandemic. 286
Eliseo Bustamante, Mónica Martínez-Gómez, César Berna-Escriche

A comparative study of Bitcoin's Price fluctuations by Twitter sentiments 287
Sadia Bruce

Analysis of factors involved in the teaching-learning system in a state of emergency 288
Alba Lira Pérez Avellaneda, Diana Cueva, Mónica Martínez-Gómez

Refusing to be safe. The Social Network Communication of deniers. 289
Rosario D'Agata, Simona Gozzo

Topic modeling in court rulings 290
Juan Diego Cuenca Camacho

Automated real estate valuation disruption in the Smart Cities context 291
Andrea San José Cabrero, Gema María Ramírez Pacheco

4th International Conference on Advanced Research Methods and Analytics (CARMA2022)
Universitat Politècnica de València, València, 2022
DOI: http://dx.doi.org/10.4995/CARMA2022.2022.14936

A visual analysis of the literature on Internet neutrality

Klaudijo Klaser[1], Lucía Desamparados Pinar García[2]
[1]Università degli Studi di Trento, Dipartimento di Economia e Management, via Inama 5, 38122, Trento, Italy. [2]University of Valencia, ERI-CES and Department of Economic Analysis, Av. de los Naranjos, s/n, 46022, Valencia, Spain.

Abstract

Internet neutrality – a principle against the discrimination between Internet data packages – has been one of the most debated Internet regulation policies in the last decade. However, this debate seems to be very fragmented and there is not a global comprehension of the direction of it. In this paper we try to fill this gap circumscribing the literature on Internet neutrality. Through the open-source software VOSviewer we provide a visual analysis of the relationship between the 50 most relevant words occurring in the abstracts of scientific publications on the topic in the last 15 years.

Keywords: *Internet, Literature Analysis, Net Neutrality, VOSviewer.*

This work is licensed under a Creative Commons License CC BY-NC-SA 4.0
Editorial Universitat Politècnica de València

1. Introduction

Internet (or net) neutrality has been the most relevant Internet regulatory policy of the last decade (Jacobides, 2020). The principle of net neutrality requires that all Internet data packages, regardless of their content, origin, destination**,** or type of equipment used, should be equally treated. (Wu**,** 2003). In general terms we can say that the debate on the need to set the obligation of net neutrality is between: on the one hand the Content Providers with strong market power (e.g. Amazon, Google, Meta, Netflix) which defend it under the argument of maintaining Internet as an open and global network that fosters innovation; and on the other hand the Internet Service Providers (e.g. Orange, Movistar, TIM, Vodafone) which argue that net neutrality discourages investments in maintenance and extension of network capacity because free rider behavior from the Content Providers side is allowed.

The academic results about the argument are often inconclusive and the regulatory policy sometimes contradictory[1]. In other words, the literature about Internet neutrality seems to be very fragmented and there is a lack of comprehension of its global direction. Therefore, the inconsistency between the results often obtained in differents fields of literature may explain the contentious and polemic debate that still continues on whether net neutrality is necessary and how to enforce it.

In this paper, through an analysis of the words occurring in the abstracts of scientific publications, we provide a global snapshot of the academic literature on net neutrality. We show that the debate is very interdisciplinary, involving law, economics, engineering and political sciences. However, in our analysis we found two dominant macro-areas of study in the literature: the the economic debate and the legal perspective. We interpret this compartmentalization of fields as an absence of synergy between economic and legal outcomes, which adds complexity and generates confusion in the net neutrality debate.

In Section 2 we present the methodology employed to provide a visual analysis of the net neutrality literature. Section 3 presents the main results. Finally, conclusions are provided.

2. Method

On the database Web of Science[2] we looked for scientific publications which had the locution “internet neutrality” or “net neutrality” either in the title or in the abstract section.

[1] In 2015, the European Union incorporated regulations on net neutrality like the one established in the US between 2015 and 2017. In the European Union, since 2016, it is the Body of European Regulator for Economic Communications BEREC that manages the guidelines about net neutrality. In general terms, BEREC prohibits any type of quality discriminatory practice such as prioritization. However, certain financial discrimination practices such as zero-rating are allowed on a case-by-case basis. More recently, in September 2021, the Court of Justice of the European Union, by analyzing the cases of Vodafone and Deutsche Telekom in Germany, decided that zero rating offers (use of applications without data consumption) violate the net neutrality principle. This is a clear signal about how still confusing and hectic the debate is.

[2] Entered the 15th of February 2022.

This search produced a total of 368 available contributions, of which 264 articles, 52 proceeding papers and 26 book chapters. Only 296 pubblications resulted supplied with an abstract field[3]. For our analysis we focused on the co-ocurrences of words within the abstracts.

Using a minimum threshold of occurrence of 10 times – counting also if a word appeared more than once within the same abstract – we identified a total of 175 possible relevant terms. Starting from that list of words we brainstormed and selected the most important ones given the context, that is Internet neutrality, according to standard procedures in the field (Caputo et al. 2021 and Donthu et al. 2021). Thus, even if some terms occurred very frequently, we excluded from the visual analysis 92 words which had no relevance given the context or which had a too general or ambiguous meaning, in the sense that they could be easily coupled with several other words in several different ways. For instance, we excluded words such as "analysis", "content", "decision", "internet", "issue" or "paper".

In a second stage of refinement of the remaining 83 terms left we grouped the words with a coincidental or identical meaning. For example, we merged "commission" and "fcc" with "federal communications commission", "Europe" with "European Union" and "law" with "legislation". In the same way we decided to merge "costumer", "consumer", "end user", "internet user" and "user" into one single category. This further refinement left us with 48 unique words with no ambiguity or synonyms plus other 10 words derived from merging 35 similar terms.

For the visual analysis of the network between the identified words we used the open-source software VOSviewer (van Eck and Waltman 2010, 2014).

3. Results

We first start with some general indexes and then we move to the network analysis.

3.1 General bibliometric indexes

Figure 1 shows the publication and citation history of the 368 contribution identified on Web of Science starting from 2006. As it is possible to notice from the bar chart, the literature on Internet neutrality constantly grew until 2018, when we can observe the pick for both the mentioned variables. FUrthermeore, the scientific production on the theme had a significant drop during the Covid years. On the one hand this is might be a natural trend of many fields of research, because the focal center moved on the health emergency. On the other hand, this conspicuous drop is surprising given the importance that Internet had

[3] Since the number of abstracts was relatively low we decided not to exclude any publication from the following analysis.

during the lockdown periods to carry out fundamental socio-economic activities and the consequent increase in traffic volumes (Feldmann et al. 2021).

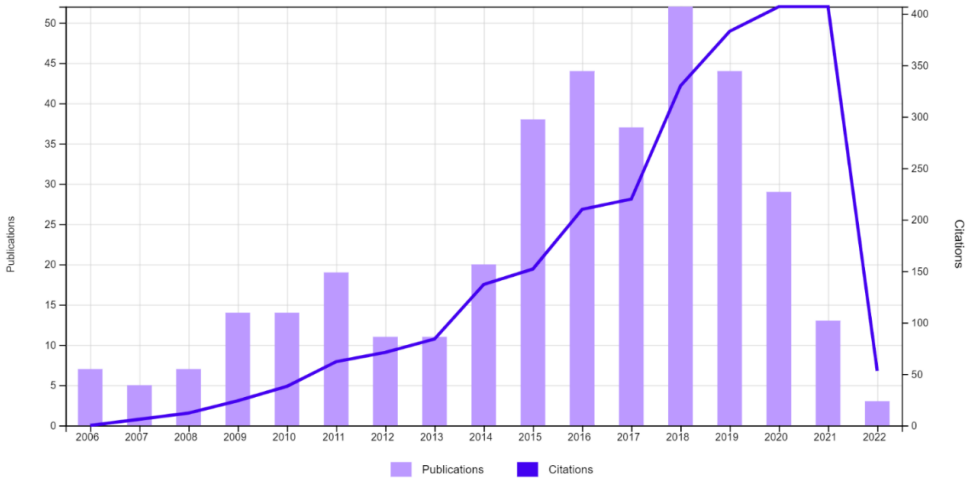

Figure 1. Publications and citations over time, 2006-2022. Source: Web of Science, Citation report

Table 1 summarizes instead the main fields of literature of the publicaitons according to the Web of Science categorization[4]. The research fields that mostly engage with the Internet neutrality debate are communication, telecommunication, information science, law, economics. Other 57 categories were present, but only less than nine contributions per category belonged to those areas. Therefore, they are not shown in Table 1.

3.2 Visual analysis of the network

Focusing on the relationship between the words present in the abstracts of the identified publications, we can notice (Fig. 2)[5] that already after the first round of refinement – thati is the elimination of non-significant words – we obtain a clear split of the 50 most relevant words in two macro-clusters (areas of research), mainly connected through the very central node represented by the term "neutrality".

[4] One publication can belong to more than one field, so there is the possibility of multiple counting.

[5] For this visualization we used a binary counting of words, which counts only the number of documents in which a term occurs. The double or triple appearance of the same word within the same abstract is counted as one. Out of the 66 total words so identified we selected the 50 most relevant.

Table 1 – Web of Science main categories of the identified contributions

Field	Times
Communication	70
Telecommunications	66
Information Science Library Science	62
Law	60
Economics	49
Engineering Electrical Electronic	42
Computer Science Information Systems	33
Computer Science Hardware Architecture	27
Computer Science Theory Methods	23
Computer Science Software Engineering	20
Management	19
Polittical Science	13

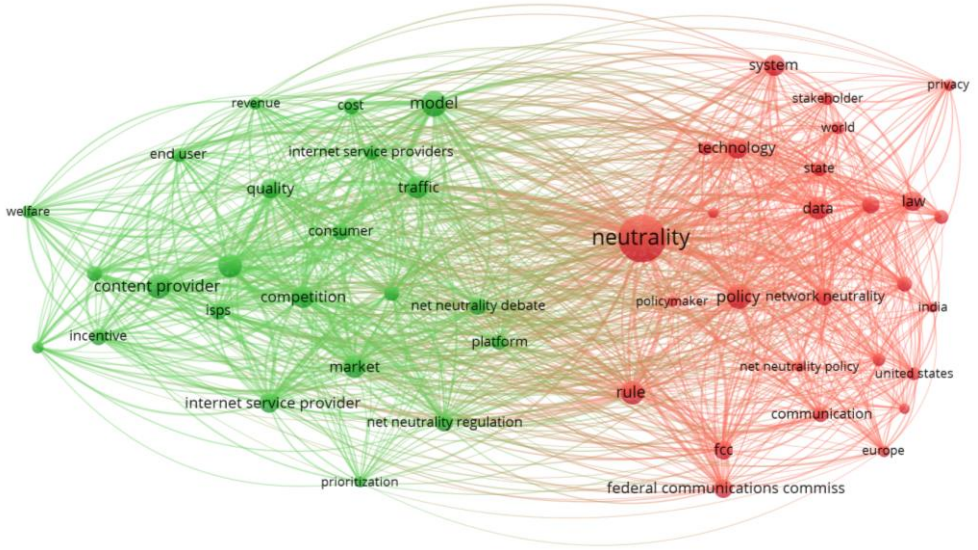

Figure 2. Network visualization of the 50 most relevant words, no merge of synonyms. Source: VOSviewer

The left cluster (green) contains words like "content provider", "internet service provider(s)", "market", "model" and "competition". This is a clear indication that one important part of the literature on Internet neutrality is basicallly focused on the normative (meant in an economic sense) configuration of the Internet market and the economic relationship between the main market actors. This is confirmed also by the co-presence of

Para seguir leyendo, inicie el proceso de compra, click aquí