



# THE LAST FIVE YEARS OF BIG DATA RESEARCH IN ECONOMICS, ECONOMETRICS AND FINANCE: IDENTIFICATION AND CONCEPTUAL ANALYSIS

**THE INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY AND  
QUANTITATIVE MANAGEMENT (ITQM)**

*JOSÉ RICARDO LÓPEZ-ROBLES, MARISELA RODRÍGUEZ-SALVADOR, NADIA KARINA GAMBOA-ROSALES, SELENE RAMIREZ-ROSALES,  
MANUEL JESÚS COBO*

November 3-6, 2019 / Granada (Spain)



# TABLE OF CONTENTS

1. Introduction
2. Methodology
3. Dataset
4. Conceptual Analysis
5. Conclusions

# TABLE OF CONTENTS

1. Introduction

2. Methodology

3. Dataset

4. Conceptual Analysis

5. Conclusions

# 1. INTRODUCTION

## CONTEXT

**Big Data** has evolved from being an **emerging topic** to a growing research area in business, science and education fields.

The Big Data concept has a **multidimensional approach**, and it can be defined as a term describing the storage and analysis of large and complex data sets using a series of advanced techniques.

The professionals involved in this area of knowledge are seeking to uncover the **conceptual structure** of a research area of interest are worth and necessary.

## OBJECTIVE

The main aim of this contribution is to develop a **bibliometric analysis to evaluate the performance and conceptual evolution** of the **Big Data Research in Economics, Econometrics and Finance** from **2015 to 2019**.

The analysis is developed using **SciMAT**.

# TABLE OF CONTENTS

1. Introduction

2. Methodology

3. Dataset

4. Conceptual Analysis

5. Conclusions

## 2. METHODOLOGY

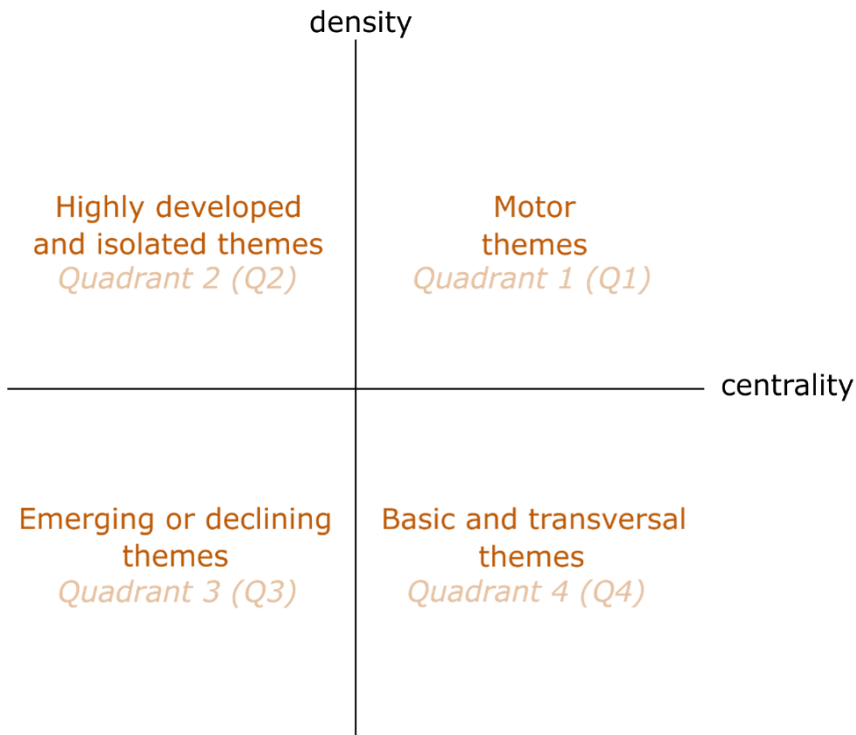
### SOFTWARE TOOL

**SciMAT** was employed to develop a longitudinal conceptual science mapping analysis based on co-words bibliographic networks.

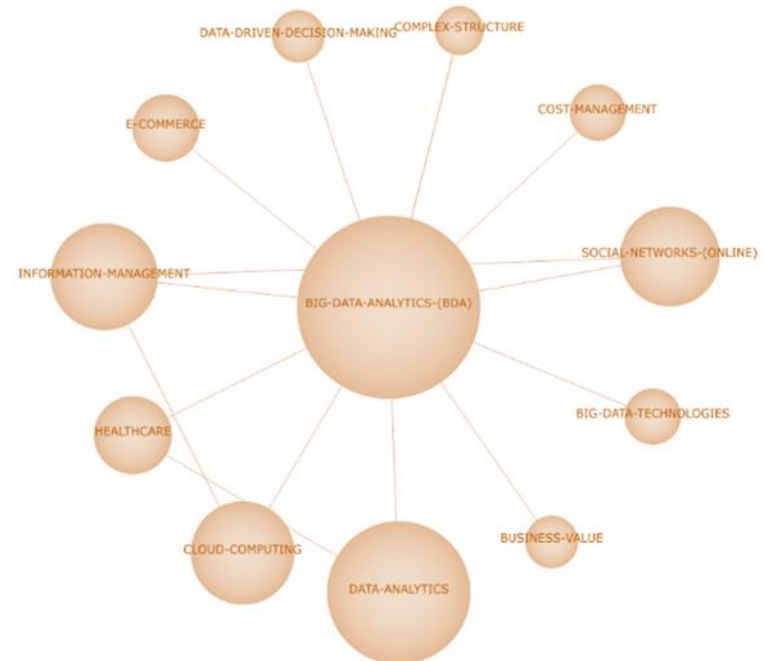
### METHODOLOGY STAGES

- 1. Detection of the research themes.** **Co-word analysis**, followed by a **clustering** of keywords to topics/themes. The similarity between the keywords is assessed using the equivalence index.
- 2. Visualizing research themes and thematic network.** **Strategic diagram** and **thematic network** (centrality and density). Research themes mapped in a two-dimensional strategic diagram and classified into four groups (Figure 1): i) motor, ii) basic/transversal, iii) highly developed-isolated, and iv) emerging/declining
- 3. Performance analysis.** Relative contribution of the research themes to the whole research field: number of published documents, number of **citations**, and different types of bibliometric indices (**h-index**).

## 2. METHODOLOGY



(a) Strategic diagram



(b) Thematic network

# TABLE OF CONTENTS

1. Introduction

2. Methodology

3. Dataset

4. Conceptual Analysis

5. Conclusions

## 3. DATASET

### CORPUS AND DATABASE

Big Data research documents published in the Scopus.

### QUERY

*TITLE-ABS-KEY ("big-data" OR "big data" OR "bigdata") AND (LIMIT-TO (SUBJAREA, "ECON")). AND (LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015)).*

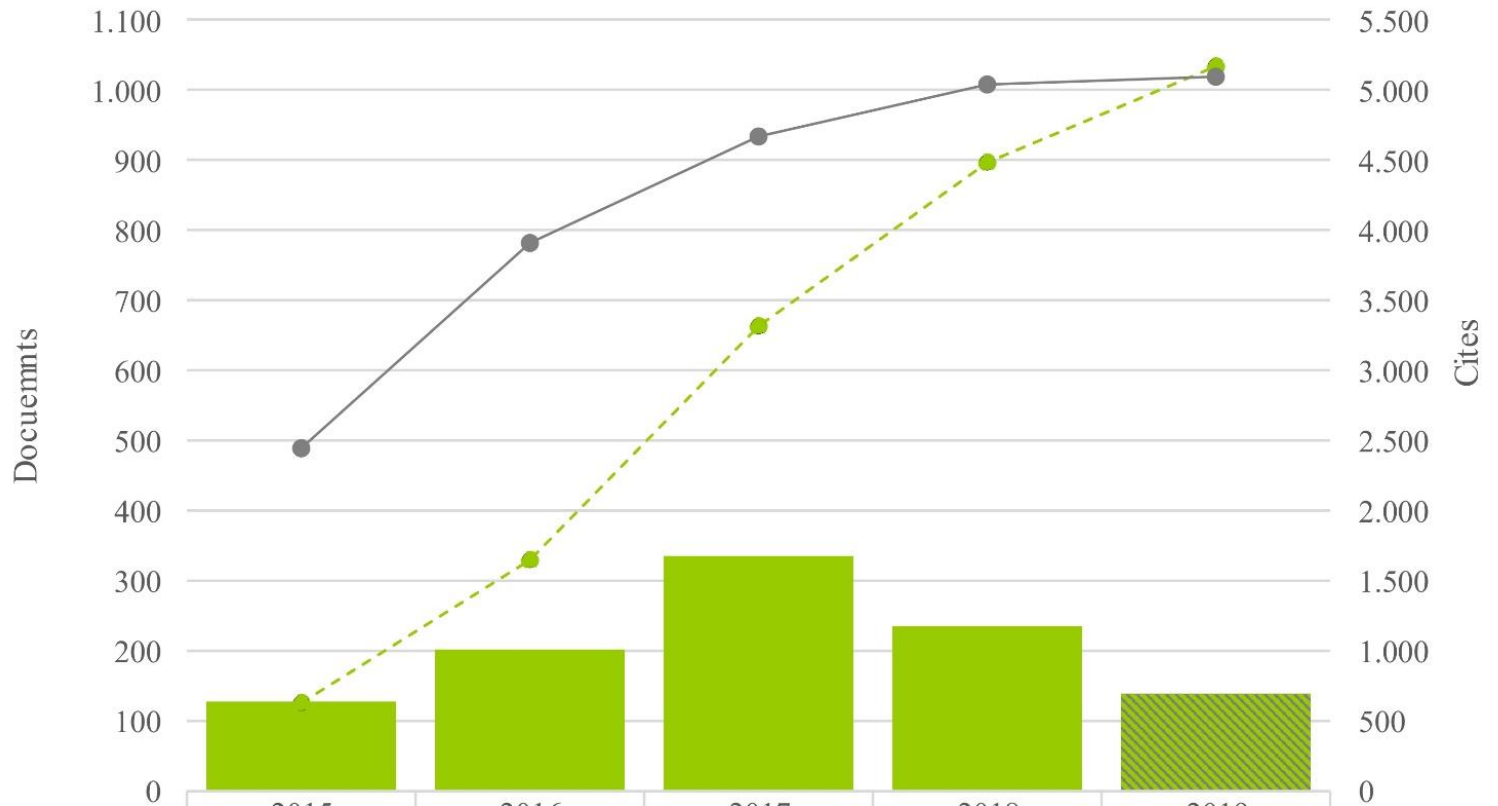
### TIME PERIOD




The corpus was evaluated in a single five-years period: 2015, 2016, 2017, 2018 and 2019.

### CORPUS SIZE

- 1,034 documents (articles, proceedings, reviews...), 5,1054 cites and 5,551 keywords.
- Citations count up to 28th July 2019.
- 2015: 126 documents, 2,445 cites and 983 keywords.
- 2016: 201 documents, 1,467 cites and 893 keywords.
- 2017: 336 documents, 759 cites and 1,417 keywords.
- 2018: 233 documents, 377 cites and 1,325 keywords.
- 2019: 138 documents, 56 cites and 933 keywords.

# DOCUMENTS AND CITATIONS BY YEAR



 Documents	126	201	336	233	138
 Documents acum	126	327	663	896	1.034
 Citations acum	2.445	3.912	4.670	5.047	5.103

# TABLE OF CONTENTS

1. Introduction

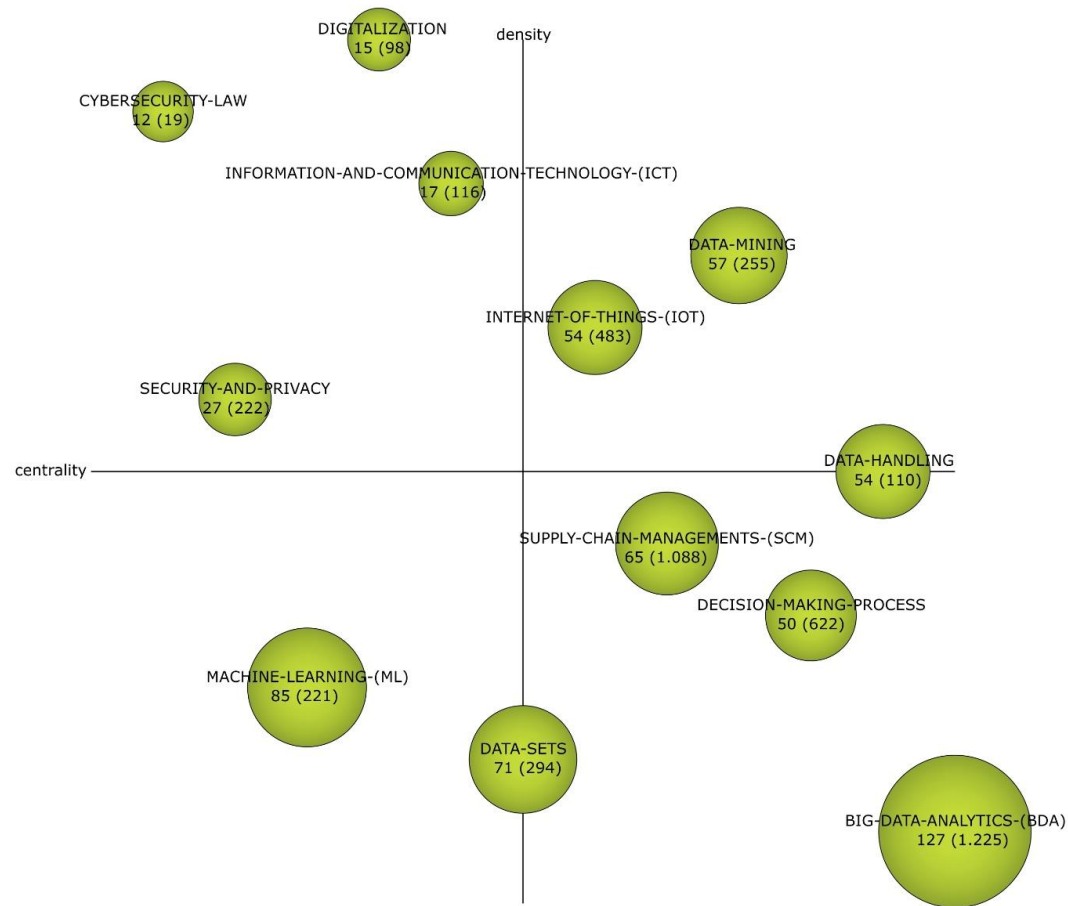
2. Methodology

3. Dataset

4. Conceptual Analysis

5. Conclusions

# 4. CONCEPTUAL ANALYSIS



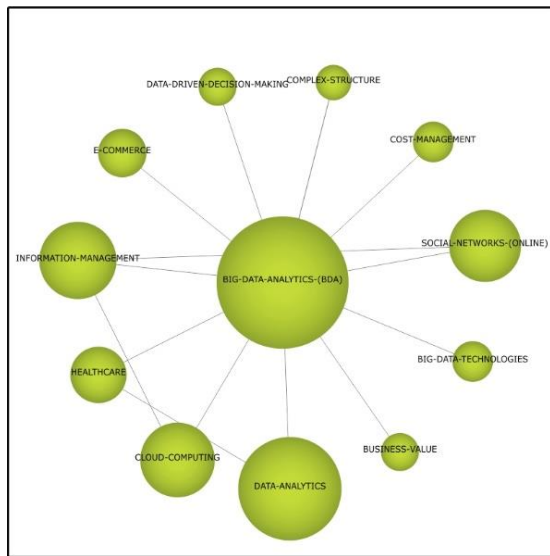
Strategic diagram 2015-2019

## 4. CONCEPTUAL ANALYSIS

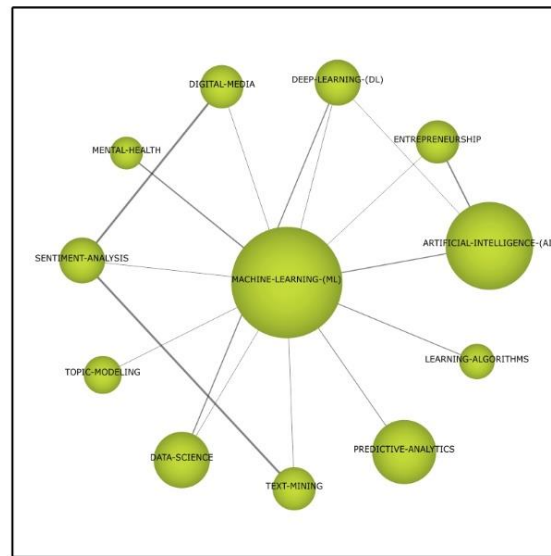
Theme	Quadrant	Documents	Citations	h-index
BIG-DATA-ANALYTICS-(BDA)	4	127	1,225	13
MACHINE-LEARNING-(ML)	3	85	221	8
DATA-SETS	3	71	294	9
SUPPLY-CHAIN-MANAGEMENTS-(SCM)	4	65	1,088	12
DATA-MINING	1	57	255	6
DATA-HANDLING	1	54	110	4
INTERNET-OF-THINGS-(IOT)	1	54	483	9
DECISION-MAKING-PROCESS	4	50	622	8
SECURITY-AND-PRIVACY	2	27	222	7
INFORMATION-AND-COMMUNICATION-TECHNOLOGY-(ICT)	2	17	116	4
DIGITALIZATION	2	15	98	3
CYBERSECURITY-LAW	2	12	19	2

The first period has lower number of publications than the other periods, we could identify ten themes related to the **Big Data** research field. In this regard, we could highlight five key themes (motor theme and basic and transversal themes) of the knowledge field: *DATA-MINING*, *INTERNET-OF-THINGS-(IOT)*, *DATA-HANDLING*, *SUPPLY-CHAIN-MANAGEMENTS-(SMC)*, *DECISION-MAKING-PROCESS* and *BIG-DATA-ANALYTICS-(BDA)*.

# 4. CONCEPTUAL ANALYSIS



(a)

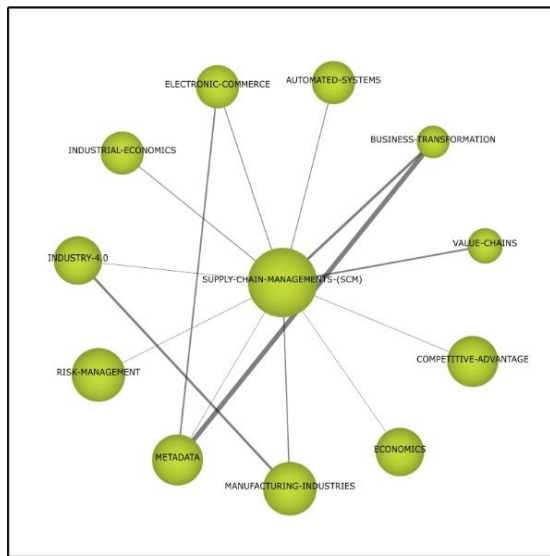


(b)

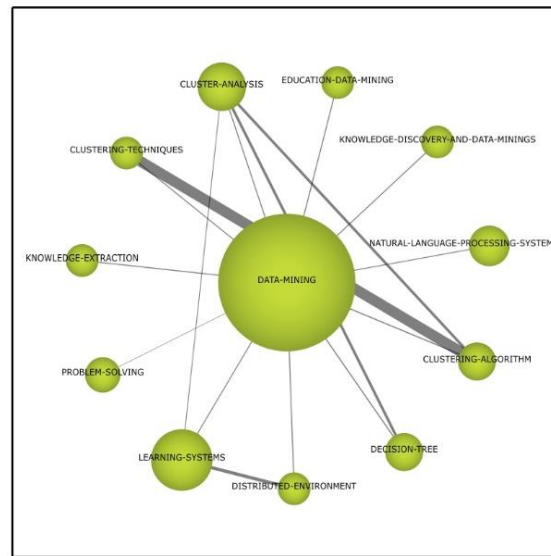


(c)

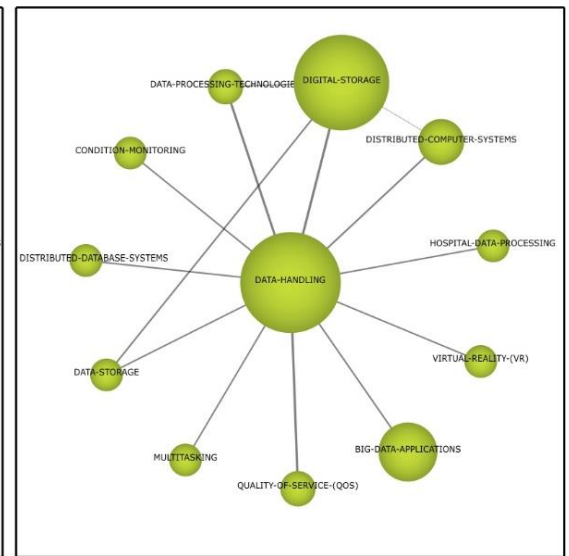
# 4. CONCEPTUAL ANALYSIS



(d)

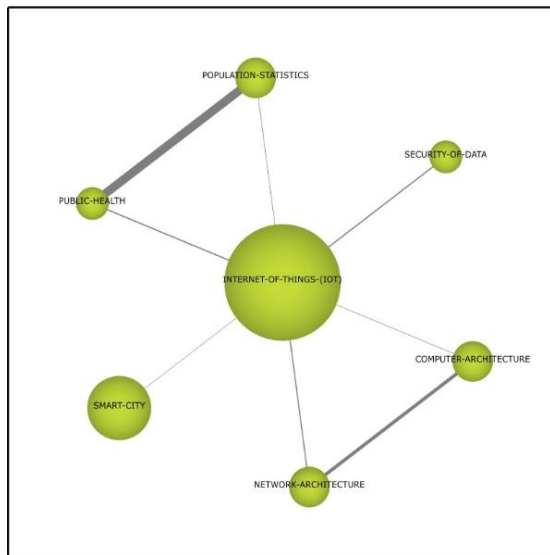


(e)

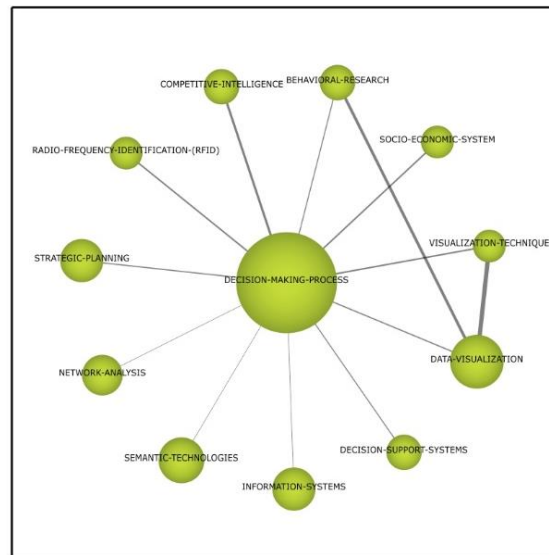


(f)

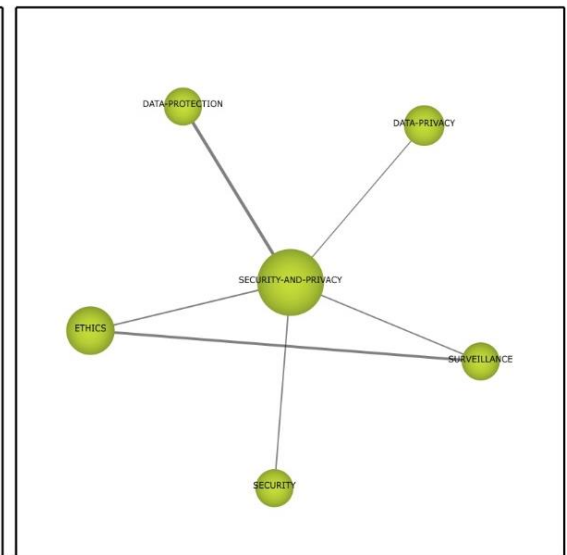
# 4. CONCEPTUAL ANALYSIS



(g)

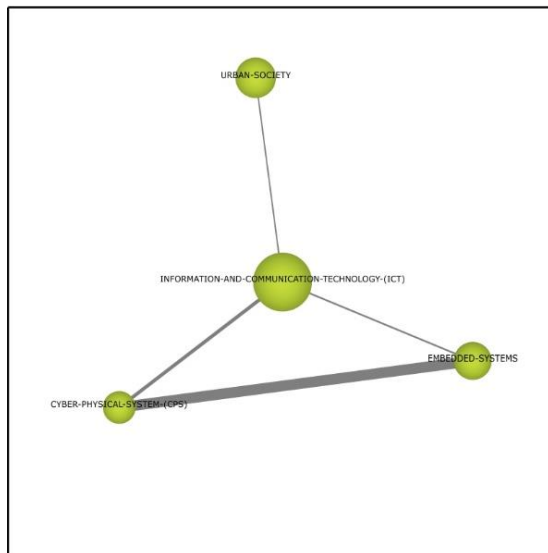


(h)

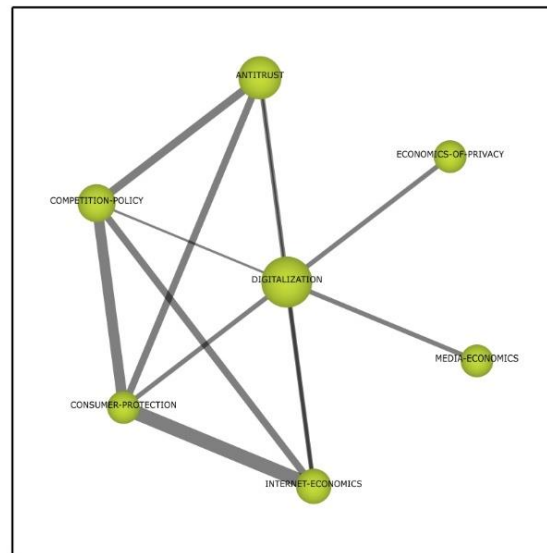


(i)

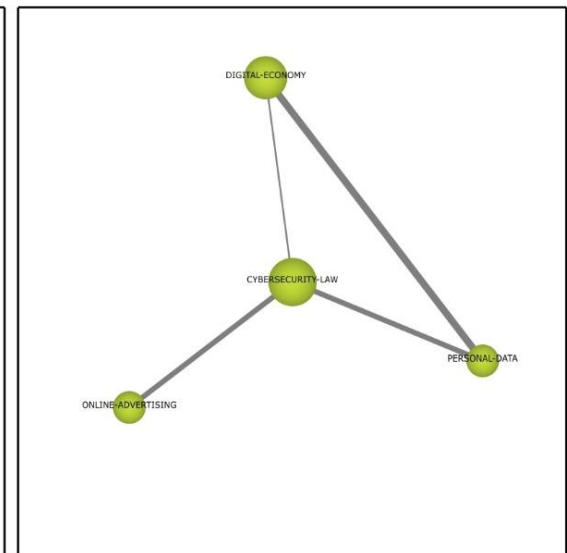
# 4. CONCEPTUAL ANALYSIS



(j)



(k)



(l)

# TABLE OF CONTENTS

1. Introduction

2. Methodology

3. Dataset

4. Conceptual Analysis

5. Conclusions

# CONCLUSIONS

## SUMMMARY

- An amount of 1,034 documents (articles, proceedings and reviews) were retrieved from the Scopus.
  
- The corpus was evaluated in a single five-years period.
  - 2015: 126 documents, 2,445 cites and 983 keywords.
  - 2016: 201 documents, 1,467 cites and 893 keywords.
  - 2017: 336 documents, 759 cites and 1,417 keywords.
  - 2018: 233 documents, 377 cites and 1,325 keywords.
  - 2019: 138 documents, 56 cites and 933 keywords.
  
- The impact achieved is summarized in the following indicators:
  - Average citations per publication: 3.40
  - Sum of Times Cited (without self-citations): 5,104 (4,584)
  - Citing articles (without self-citations): 2,954 (2,609)
  - h-index: 30 publications

# CONCLUSIONS

## MAIN CONCLUSION

- The size of literature related to **Big Data** research field showed a noticeable increase in the past decade (2015-2019). Given the large volume of publications and citations received in this field, it is expected that the use of these will be seen as part of other knowledge fields.
- The main themes used in the **Big Data** literature are: *BIG-DATA-ANALYTICS-(BDA)*, *MACHINE-LEARNING-(ML)*, *DATA-SETS*, *SUPPLY-CHAIN-MANAGEMENTS-(SCM)* and *DATA-MINING*.

## FUTURE WORKS

- Evaluate the evolution and application of the research themes in other disciplines.
- Study the relationship to other disciplines as **Business Intelligence** and **Competitive Intelligence**.



ITQM 2019

THANK YOU

THE INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY AND  
QUANTITATIVE MANAGEMENT (ITQM)

**Acknowledgments:** This work was supported by European Union FEDER funds (TIN2016-75850-R). Additionally, the authors acknowledge the support by DGRI-Dirección General de Relaciones Exteriores (Mexico) and CONACYT-Consejo Nacional de Ciencia y Tecnología (Mexico) to carry out this research.

# REFERENCES

Oussous A, Benjelloun FZ, Lahcen AA, Belfkih S, Big Data technologies: A survey, Journal of King Saud University-Computer and Information Sciences 2018; 30:431-448.

López-Robles JR, Otegi-Olaso JR, Porto Gómez I, Cobo MJ, 30 years of intelligence models in management and business: A bibliometric review, Int J Inf Manage 2019; 48:22-38.

López-Robles JR, Otegi-Olaso JR, Arcos R, Gamboa-Rosales NK, Gamboa-Rosales H, Mapping the structure and evolution of JISIB: A bibliometric analysis of articles published in the Journal of Intelligence Studies in Business between 2011 and 2017, J. Intell. Stud. Bus. 2018; 8.

López-Robles JR, Otegi-Olaso JR, Gamboa-Rosales NK, Gamboa-Rosales H, Cobo MJ, 60 Years of Business Intelligence: A Bibliometric Review from 1958 to 2017, New Trends in Intelligent Software Methodologies, Tools and Techniques: Proceedings of the 17th International Conference SoMeT\_18, IOS Press; 2018, p. 395.

McAfee A, Brynjolfsson E, Davenport TH, Patil DJ, Barton DC, Big data: the management revolution, Harv. Bus. Rev. 2012; 90:60-68.

López-Robles JR, Otegi-Olaso JR, Porto Gómez I, Gamboa-Rosales NK, Gamboa-Rosales H, Robles-Berumen H, Bibliometric Network Analysis to Identify the Intellectual Structure and Evolution of the Big Data Research Field, International Conference on Intelligent Data Engineering and Automated Learning, Springer; 2018, p. 113-120.

López-Robles JR, Guallar J, Otegi-Olaso JR, Gamboa-Rosales NK, El profesional de la información (EPI): bibliometric and thematic analysis (2006-2017), El profesional de la información 2019; 28:e280417.

Cobo MJ, SciMat: herramienta software para el análisis de la evolución del conocimiento científico. Propuesta de una metodología de evaluación, Granada: Universidad de Granada; 2012.

Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F, SciMAT: A new science mapping analysis software tool, J. Am. Soc. Inf. Sci. Technol. 2012; 63:1609-1630.

Martínez MA, Herrera M, López-Gijón J, Herrera-Viedma E, H-Classics: Characterizing the concept of citation classics through H-index, Scientometrics 2014; 98:1971-1983.

Gutiérrez-Salcedo M, Martínez MA, Moral-Munoz JA, Herrera-Viedma E, Cobo MJ, Some bibliometric procedures for analyzing and evaluating research fields, Appl Intell 2018; 48:1275-1287.