Research Trends, Enabling Technologies and Application Areas for Big Data

Lecturer: Mr. Chan Sophal

Prepared by: SAM Lyheng
POV Phannet
LUK Chamnan

08 January, 2024

Content

ABSTRACT

Abstract: The availability of large amounts of data in combination with Big Data analytics has transformed many application domains. In this paper, we provide insights into how the area has developed in the last decade. First, we identify seven major application areas and six groups of important enabling technologies for Big Data applications and systems. Then, using bibliometrics and an extensive literature review of more than 80 papers, we identify the most important research trends in these areas. In addition, our bibliometric analysis also includes trends in different geographical regions. Our results indicate that manufacturing and agriculture or forestry are the two application areas with the fastest growth. Furthermore, our bibliometric study shows that deep learning and edge or fog computing are the enabling technologies increasing the most. We believe that the data presented in this paper provide a good overview of the current research trends in Big Data and that this kind of information is very useful when setting strategic agendas for Big Data research.

Keywords: survey; Big Data; telecommunication; image processing; smart cities; manufacturing; parallel processing; storage systems; cloud computing; deep learning

INTRODUCTION

There are two main challenges associated with these enormous amounts of data:

1. We need to provide storage systems, database systems, processing platforms, etc. to technically handle the data in a fast, cost-effective and secure way.
2. We need to develop analysis methods

METHODOLOGY

The methodology in this study consists of two steps:

1. Identifying the major application areas and the major enabling technologies for Big Data.
2. Using bibliometrics to quantify how the research interest for each of the identified application areas and enabling technologies for Big Data have developed during the last decade.

APPLICATION AREAS FOR BIG DATA

- Present a variety of application areas for Big Data, including healthcare analytics, financial services, e-commerce, supply chain management, smart cities, energy sectors, telecommunications, social media analysis, government services, and g.

- Showcase how specific technologies like Apache Hadoop, Apache Spark, and NoSQL databases are employed in these application domains.

ENABLE TECHNOLOGIES FOR BIG DATA

- Discuss key enabling technologies for Big Data, including distributed computing, NoSQL databases, data lakes, containerization, and orchestration.

- Mention the role of in-memory computing, graph databases, and hybrid cloud architectures in enhancing data processing capabilities.

TRENDS IN BIG DATA RESEARCH

- 25 May, 2022

Table 1. The number of research publications per year in the area of Big Data.

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
No. of Publications	663	2766	4817	8751	12,389	13,679	16,020	19,366	19,693	19,758

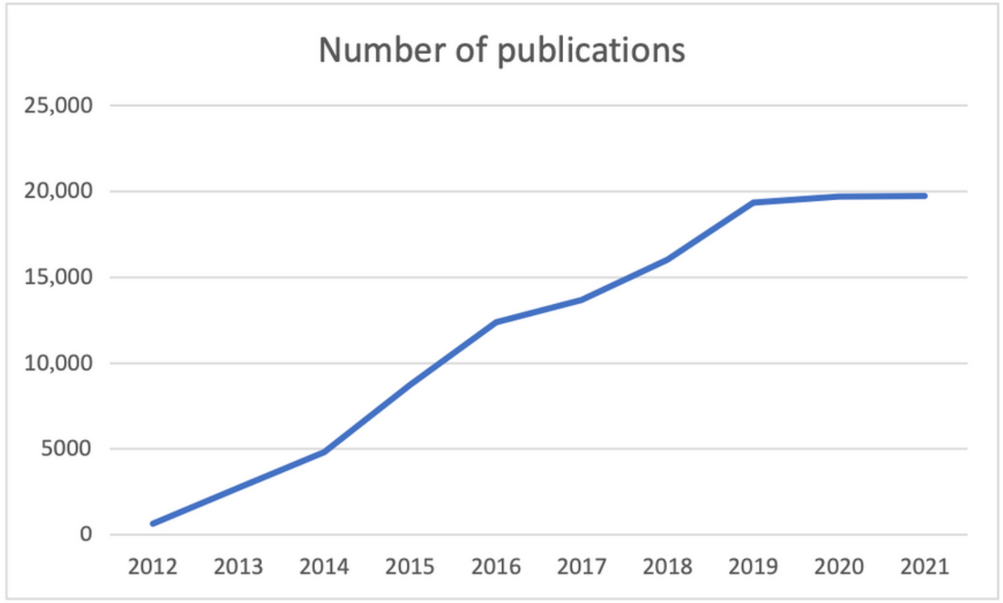

Figure 1. The number of research publications per year in the area of Big Data.

TRENDS IN BIG DATA RESEARCH

- 25 May, 2022

Table 2. The number of research publications per year for the seven application areas for Big Data that we have identified.

Application Area\Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Telecommunication	29	85	169	370	658	550	737	720	675	628
Manufacturing	11	63	180	384	546	886	1180	1540	1912	2323
Smart Cities	8	28	69	166	327	476	747	1014	1096	1281
Image Processing	6	92	154	452	606	714	903	1147	1218	1083
Social Media	46	232	402	844	1223	1367	1548	2063	2040	2008
Agriculture or Forestry	7	47	97	188	341	396	615	882	1076	1359
Finance	10	54	111	265	390	531	672	987	1313	1588

TRENDS IN BIG DATA RESEARCH

- 25 May, 2022

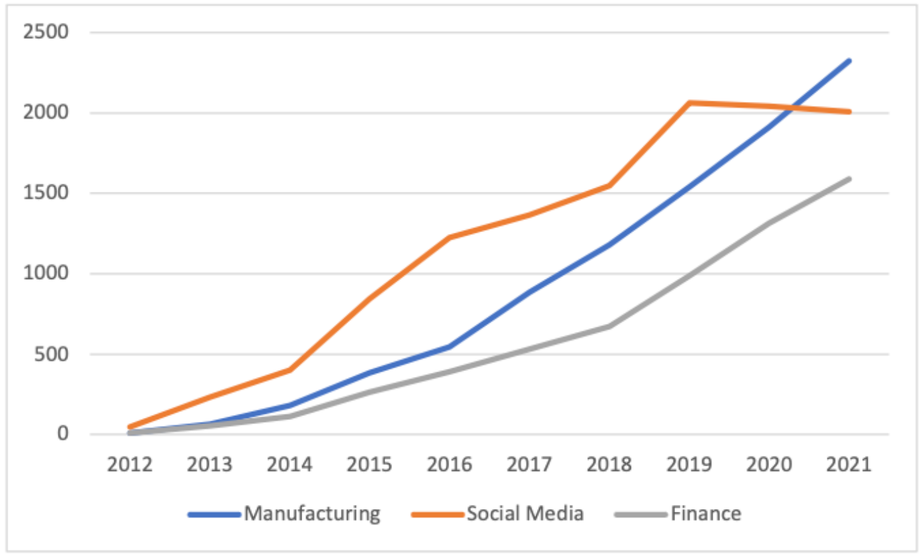

Figure 2. The number of research publications per year for the Big Data application areas manufacturing, social media and finance.

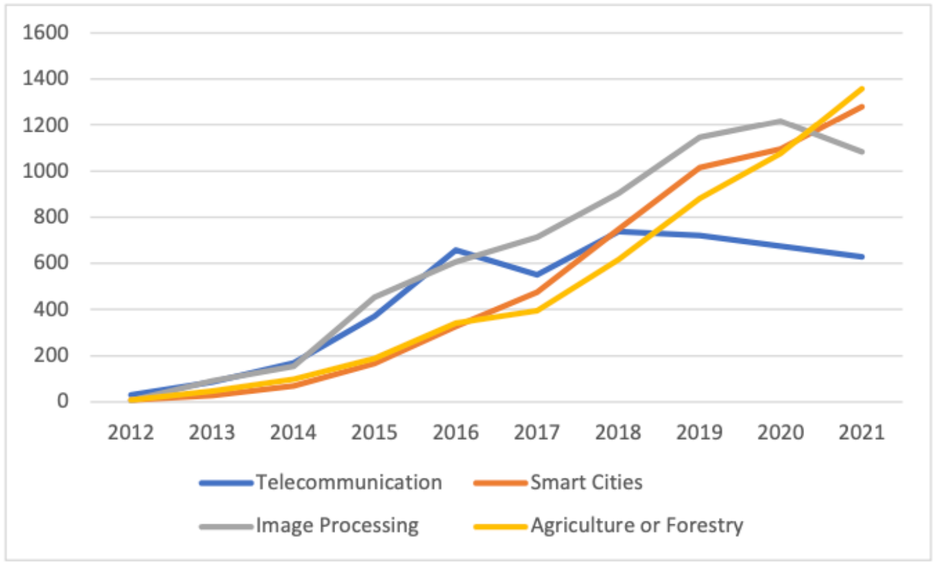

Figure 3. The number of research publications per year for the Big Data application areas of telecommunication, smart cities, image processing and agriculture or forestry.

TRENDS IN BIG DATA RESEARCH

- 25 May, 2022

Table 3. The number of research publications per year for the six enabling technologies for Big Data that we identified.

Enabling Tech.\Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
High-Performance Computing or Supercomputing	69	203	382	753	1009	1012	1105	1378	1051	1008
Deep Learning	0	7	29	143	388	835	1674	3195	3958	4434
Cloud Computing	147	635	1078	2020	2880	3149	3594	4223	3644	3704
Parallel Processing, Distributed Processing or GPU	81	289	480	858	1373	1187	1239	1685	1221	867
Storage Systems or Database Systems or Data Lakes	115	411	639	1017	1214	1312	1313	1452	1187	1035
Edge Computing or Fog Computing	0	3	5	25	81	167	501	744	957	1123

TRENDS IN BIG DATA RESEARCH

- 25 May, 2022

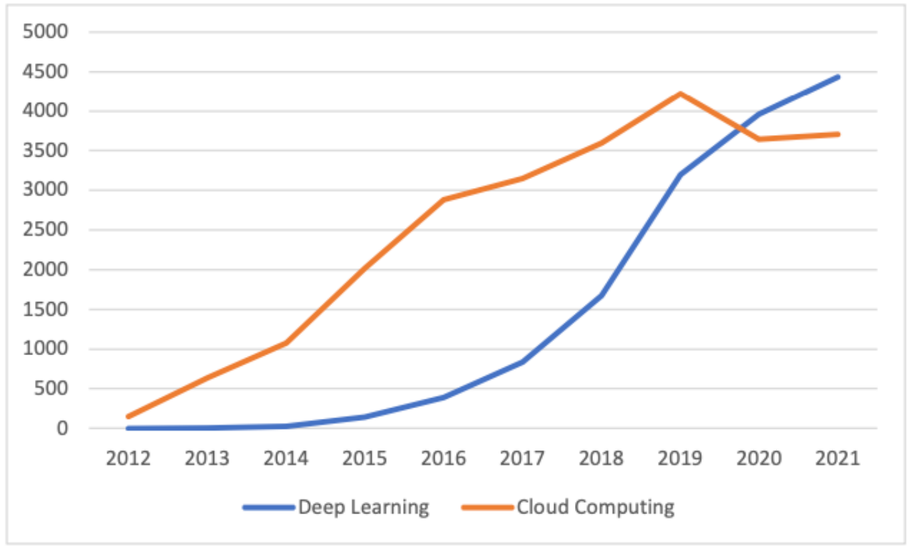

Figure 4. The number of research publications per year for the Big Data-enabling technologies deep learning and cloud computing.

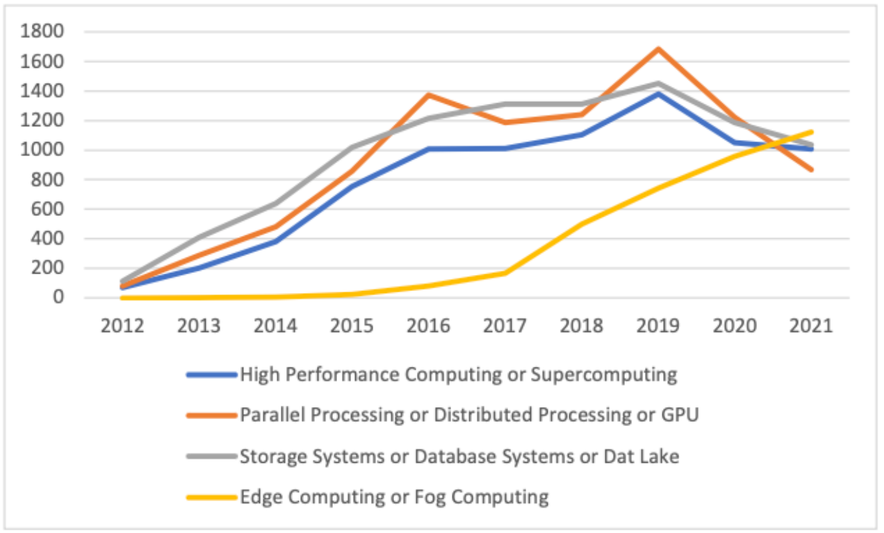

Figure 5. The number of research publications per year for the Big Data-enabling technologies of high-performance computing or supercomputing; parallel processing, distributed processing or GPU; storage systems, database systems or data lakes and edge computing or fog computing.

TRENDS IN BIG DATA RESEARCH

- 25 May, 2022

Table 4. The number of research publications in Big Data per year for different geographic regions.

Countries\Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
China	69	588	962	1970	3340	3862	4960	5831	8330	9263
USA	268	921	1523	2542	3063	3052	3139	3656	2917	2227
Other	326	1257	2332	4239	5986	6765	7921	9879	8446	8268

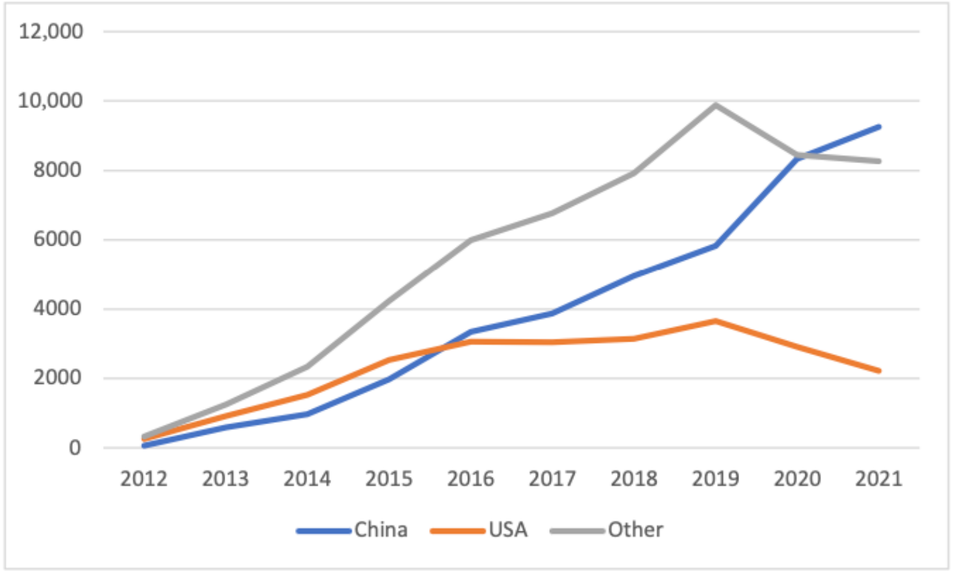

Figure 6. The number of research publications in Big Data per year for different geographic regions.

DISCUSSION

Table 5. The number of research publications related to Big Data per year for AI and IoT.

Concept\Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Artificial Intelligence or AI	192	1068	1804	3580	4930	5497	6830	8605	9436	9353
Internet of Things or IoT	13	91	270	727	1307	1924	2999	4026	4442	5341

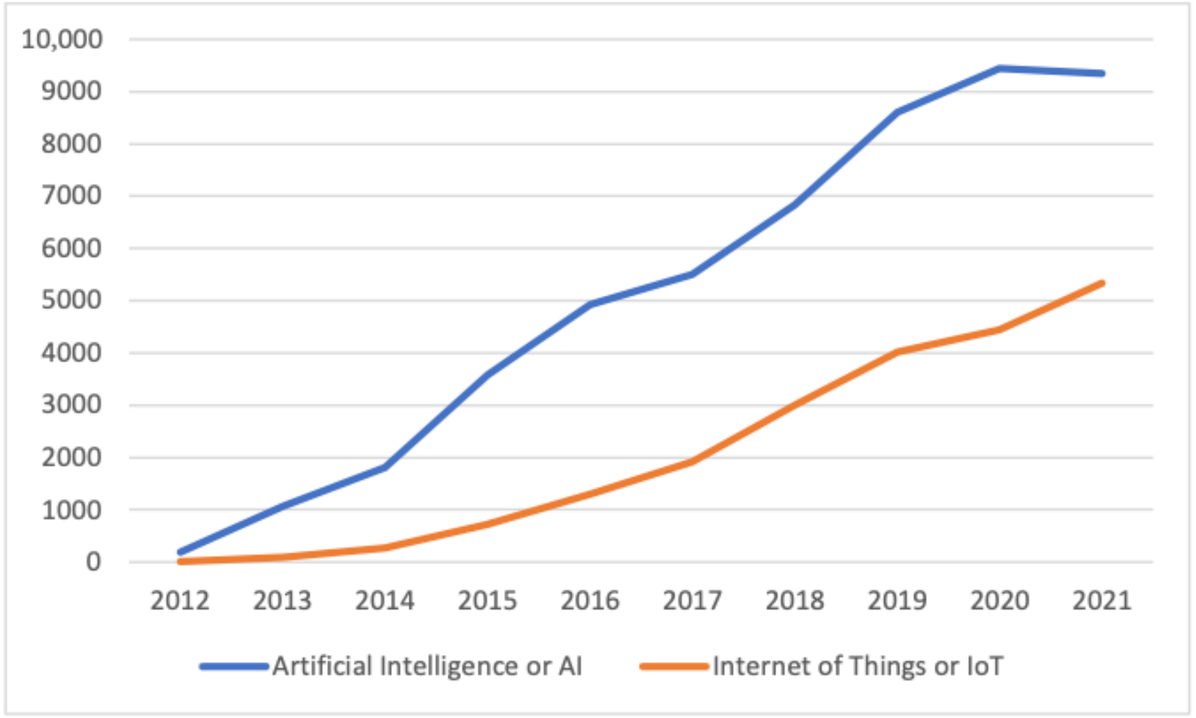

Figure 7. The number of research publications related to Big Data per year for AI and IoT.

CONCLUSION

In the last decade, Big Data research has boomed but recently plateaued. Seven key application areas, led by manufacturing and agriculture/forestry, have seen rapid growth in research. Technologies like deep learning and edge/fog computing are rising.

The USA and China dominate Big Data research, with China's interest growing faster. The link between Big Data and AI is strong, while the IoT is driving interest in manufacturing, smart cities, and agriculture/forestry.

The challenge lies in efficiently handling vast data while deriving value from it. Understanding these trends is crucial for shaping future Big Data research agendas.

Thanks for listening