Cognitive Cloud: The Definition

Cognitive Cloud: The Definition

Sergio Moreschini[1], Fabiano Pecorelli[1], Xiaozhou Li[1], Sonia Naz[1], Michele Albano[2], David Hästbacka[1], and Davide Taibi[1,3]

[1] Tampere University, Tampere, Finland
[2] Aalborg University, Aalborg, Denmark
[3] University of Oulu, Oulu, , Finland
{sergio.moreschini, fabiano.pecorelli, xiaozhou.li, sonia.naz, david.hastbacka, davide.taibi}@tuni.fi, mialb@cs.aau.dk

Abstract. Cognitive Cloud has drawn increasing attention from practitioners, academics, and funding agencies and has been adopted progressively. However, the concept remains mired in various definitions with different studies providing contrasting descriptions. Therefore, to understand the concept of cognitive cloud and to provide its definition, in this work we conduct a systematic mapping study of the literature investigating 24 papers proposing five main definitions. The main outcome of this work is a complete definition that merges all the common aspects of cognitive cloud, enabling practitioners and researchers to better understand what cognitive cloud is.

Keywords: Cloud Computing · Cognitive Cloud · Survey.

1 Introduction

From research works to international funding agencies [1] [7], the term Cognitive Cloud is getting more and more popular. However, besides the increase of its usage, its definition is still not clear, and several sources propose different ones, often contrasting each others.

In order to shed some light on the definition of Cognitive Cloud, and in particular, to understand how it is defined and how its definition evolved, in this work we propose a systematic mapping study of the literature. For this purpose, we formulate two main Research Questions (RQs) as follows.

– ***RQ1:*** *What are the definitions of cognitive cloud?*

  With this RQ we aim at understanding whether there are different definitions of cognitive cloud.

– ***RQ2:*** *How has the definition of cognitive cloud evolved?*

  Via the comparison among the different definitions, we shall observe the changes from the earliest to the latest. In this way, we shall identify how the definition of cognitive cloud has evolved through multiple fields of research.

In particular, we aim at easing the understanding of where and when different definitions have been originated. As the result of this paper, we define Cognitive Cloud as following:

> **Cognitive Cloud** is a Cloud-based system that is capable of sensing its environment, learning from it, and opportunistically and dynamically adapt its computational load as well as its outcome.

At the best of our knowledge, this is the first work that investigate the different definitions adopted, and that help to understand how Cognitive Cloud should be interpreted.

The remainder of this paper is structured as follows. Section 2 presents related reviews. Section 3 describes the research method adopted. Section 4 presents the results answering the RQs. Section 5 discusses the results while Section 6 draws the conclusion and highlights future works.

2 Background and Related Work

Together with the advance in computer hardware and software, as well as that in the big data and AI technologies, the term "cognitive" has been utilized increasingly often within the domain of computer science, as computers are, more than ever, expected to perform analogous to human brain [5]. Entering this century, "cognitive informatics" was seen an emerging and profound interdisciplinary research area which studies internal information processing mechanisms and processes of the brain, and their applications in computing and the IT industry [14]. In the 2010s, "cognitive computing" becomes the new trending research niche aiming to develop novel learning systems that act on vast amounts of data from many sources at once with coherent, unified, universal mechanism inspired by the capabilities of human brain [11]. By utilizing the theories in cognitive science, cognitive computing develops algorithms and enables machines to perform with "brain-like" cognitive intelligence [3]. Especially when entering the era of the big data, the increasing amount of unstructured data causes problems in information analysis and processing where cognitive computing can provide solutions by imitating human's way of thinking. Cognitive computing systems, combining IoT, big data analysis and cloud computing, enabled by deep learning and reinforcement learning, shall be more widely introduced in various scenarios [5].

On the other hand, cloud computing enables ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources rapidly provisioned and released with minimal management effort or service provider interaction, which can provide the computing and storage resource services needed by cognitive computing [10] [4]. Nowadays, the mining and analysis of large volume of data, as well as the application of cognitive computing technologies, are conducted largely on cloud computing platforms. On the other hand, cognitive computing can also effectively support the practice in cloud computing in various areas. For example, Chen et al. propose an edge cognitive computing architecture that deploys cognitive computing at the edge of the network to provide dynamic and elastic storage and computing services with dynamic cognitive service migration mechanism [6]. Abeshu and Chilamkurti apply cognitive computing in fog-to-things computing towards attack detection for IoT devices [2]. However, though the application of cognitive computing technologies facilitating cloud architectures has been investigated, no studies have been conducted

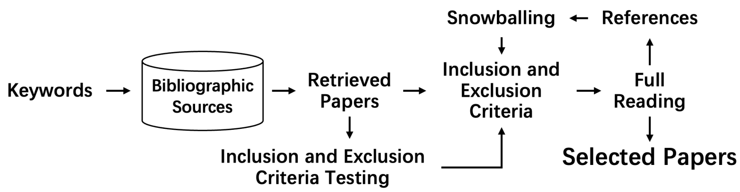


Fig. 1: The Search and Selection Process

targeting the definition of cognitive cloud and its concrete theoretical framework. Except for several works mentioning the concept of cognitive cloud, at best, no works provided its clear definition. Towards forging the first definition of an emerging concept, systematic mapping studies is an effective and well-recognized method. Using systematic mapping study, Lenarduzzi and Taibi provide a systematic definition of minimal viable product and the genealogical tree of previous definitions [9]. In this study, we follow the same approach adopted by [9] to target the definition of cognitive cloud using the same method.

3 Methodology

In this paper, we conducted a systematic mapping study of the literature related to cognitive cloud. We systematically identified, categorized, and analyzed the existing literature by following the guidelines presented by Petersen et al [13]. The aim was not only to characterize the existing definitions of cognitive cloud and other related relevant concepts but also to explore the evolution of these definitions through time. In the following, details about the process are provided.

3.1 Search Strategy

First of all, we defined a search query that could allow us to extract a comprehensive set of studies containing the targeted definition. For this reason, the search strategy consisted of the following steps: firstly, the search string definition, then, the selection of key sources to identify and shortlist the primary studies, and lastly, the data extraction and the analysis of the final results. The search strategy defines the scale for the selection of relevant bibliographical sources by defining the inclusion and exclusion criteria. Figure 1 illustrates the search strategy. We directly defined the search query as "*cognitive cloud*" in order to obtain only the most relevant results.

In order to ensure the completeness of the search results, we searched in four different bibliographical sources for the related literature, namely Scopus,[4] IEEEXplore Digital Library[5], the ACM Digital Library,[6] and Web of Science[7]. The numbers of search results from selected sources are reported in Table 1. The search was conducted on 2022-04-13. As a result, we identified 34 non-duplicated papers out of 66 in total.

[4] Scopus, `https://www.scopus.com`.
[5] IEEEXplore Digital Library `https://ieeexplore.ieee.org/`.
[6] ACM Digital Library: `https://dl.acm.org`.
[7] Web of Science database: `https://www.webofscience.com/`

Table 1: Initial Search Results by Sources

Library	Count
Scopus	28
IEEE	14
ACM	6
Web of Science	18
Non-duplicates	34

3.2 Primary Studies Selection

We shortlisted the primary studies from the initial search results by defining the inclusion and exclusion criteria. Specifically, we included the papers that are published in journals or conferences and are defining the concept of cognitive cloud. We excluded the papers that are not written in English, are not discussing the specific or relevant topic, or are duplicated. Similarly, we did not include the posters, work plans/road maps, vision papers, and papers not peer reviewed.

Following the inclusion and exclusion criteria, we selected the primary research studies as a result of two steps. In the first step, two evaluators (i.e., authors of this paper) independently read each paper's title and abstract in order to determine if it should be included or not. The papers that were rejected by both were excluded while, if there was a conflict, a third evaluator made the final decision. Out of 34 papers, we had 5 disagreements with a Cohen's kappa coefficient of 0.64, indicating a substantial agreement [8]. We recognized a total of 25 publications for the next phase at the end of this stage.

As a second step, we ran a snowballing process that included all of the papers cited in the 25 papers. We then applied the same procedure to analyze their titles and abstracts, using inclusion and exclusion criteria. This resulted in two additional papers.

All evaluators read the full-text of each of the selected 27 papers independently according to the inclusion and exclusion criteria. As a result, we finalized 24 papers.

3.3 Data Extraction Strategy

We extracted the data from the finalized 24 research papers according to our research questions. We extract the specific definition of cognitive cloud from each paper. In addition, the type of publication and the publication year are also summarized to show the trend in this research topic.

As the definitions were written in natural language, we were required to conduct a qualitative study among the authors to discover similar and dissimilar definitions. Therefore, we relied on the collective coding method to respond to our study questions. The manual identification of the aforementioned information was extracted collaboratively.

Firstly, two authors extract the definition from each paper and print them on post-it notes. Whenever the authors have disagreement on the definition selections, a third author needs to make the confirmation. When all the definitions are extracted, the authors start to post them one by one on a whiteboard, where from left to right indicates the publication year from early to recent. During the process, whenever the authors find any definition is similar or citing another definition, the two definitions shall be merged. Fig. 2 shows the extracted definitions of cognitive cloud together with the connection within as well as the year when it was published.

4 Results

Herein, we present the results summarized from the 24 selected papers via the previously described process with 10 conference papers, 10 journal papers and 4 book chapters (shown in Fig. 3). The number of studies on this topic increases slightly from 2012 to 2021. Firstly, we categorize the definitions from the selected papers by the focusing aspects of the definitions, answering RQ1. Secondly, we investigate the changes and the evolution of the definitions through time, together with the categorization, answering RQ2.

4.1 The Definitions of Cognitive Cloud

By synthesizing all the 24 definitions from the selected papers, we identify five categories of definitions with respect to their main focus when talking about "cognitive". Each group is represented by a block of a different colour, shown in Fig. 2.

The five categories are described as follows.

1. **Radio Access Technologies:** Eight of the 24 selected papers provide definitions of "cognitive cloud" related to radio access technologies (RATs), which is the underlying physical connection method for a radio based communication network that enables reliable and low-latency vehicular communications [12]. As the earliest relevant definition given on "cognitive" technologies, [SP1] emphasizes its self-management and learning capability towards the Cognitive Management system for the Opportunistic Network with RAT. [SP4] also indicates the objectives of cognitive radio access is to maximize the aggregate access and average per-client access rate of network. On the other hand, regarding cognitive cloud offloading, [SP7], [SP9] and [SP16] all mention that it uses radio interfaces used related to data transfer with fast and low-complex strategy. Furthermore, specifically on "cognitive radio" (CR), [SP10], [SP11] and [SP12] it is an intelligent communication system which accesses spectrum dynamically.
2. **Client-centric:** Two of the definitions focuses on the perspective of "Client-centric". [SP8] indicates that "cognition" of a fog architecture or a particular info-communication mechanism is the awareness of client-centric objectives and requirements (for [SP17] patients' info).
3. **Architectural:** Another perspective regarding the "cognitive" definition is the architecture. In [SP19], the authors define "cognitive computing continuum" as *distributed, opportunistic, collaborative, heterogeneous, self-managed, sensing, and learning environment, bridging the Edge and Clouds*. On the other hand, [SP2], [SP3] and [SP5] also anchor the definition in terms of architecture, indicating cognitive system is capable of environment perception, decision making regarding system provision, and adoption towards system optimization and Quality of Experience (QoE).
4. **Involving AI:** Furthermore, eight studies provide definitions in terms of the involvement of AI technologies. Regarding "cognitive computing", [SP6] indicates it is *a new paradigm of computing that enables systems to learn from multimedia data and enhance human cognitive ability to understand multimedia data and will naturally interact with humans through their senses.* [SP13] and [SP18] rephrase cognitive computing as augmented intelligence working with humans from humans' perspective. [SP14] simply maps the concept to AI. [SP15] defines cognitive technologies as *Cognitive technologies involve extracting knowledge, understanding*

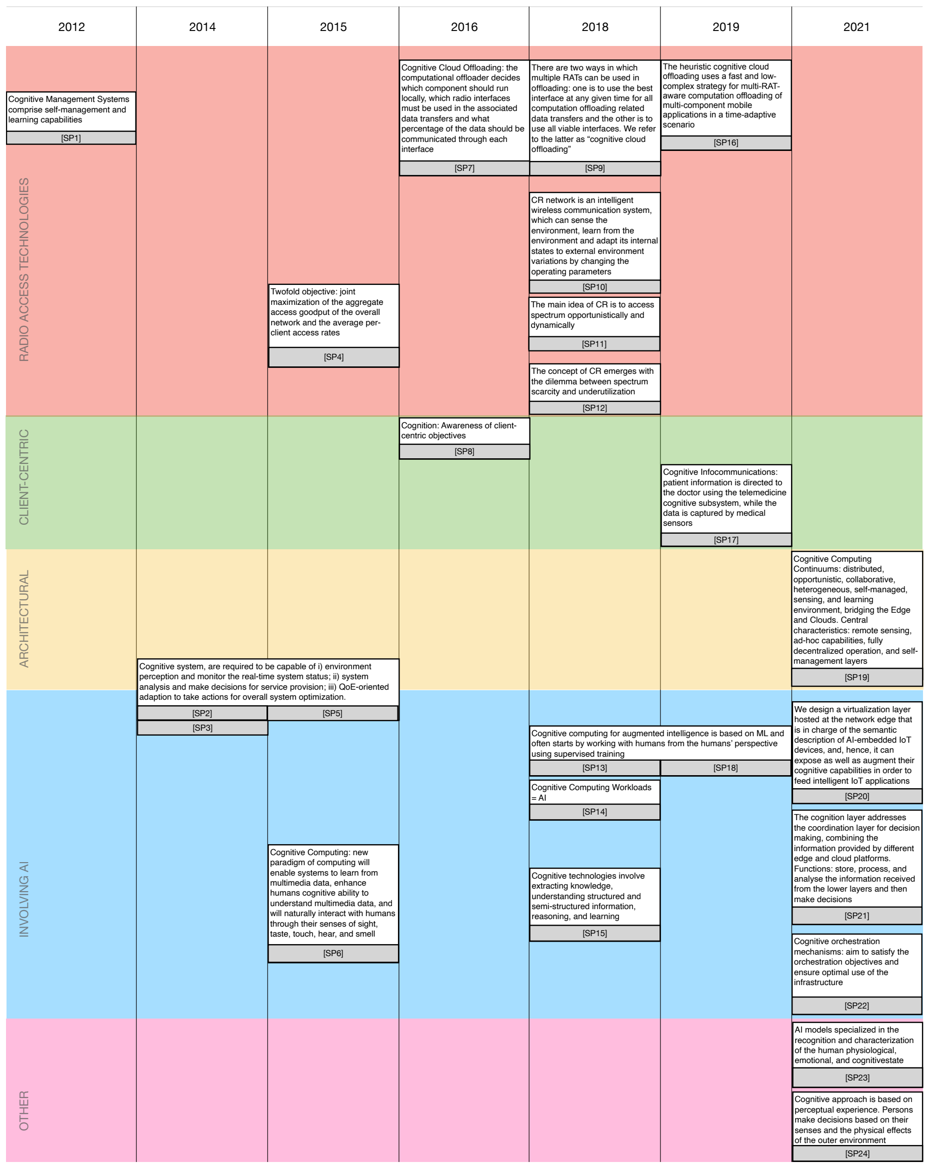
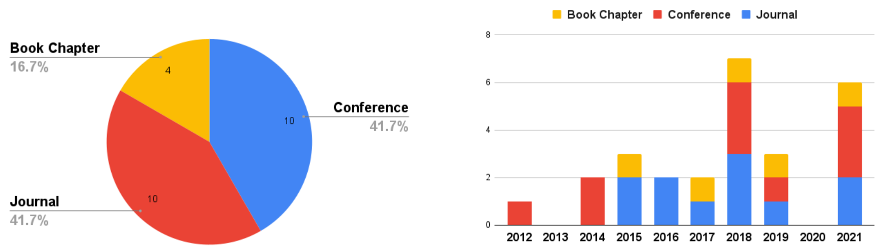


Fig. 2: Definitions of cognitive grouped by year and concepts. Each column represent a different year while the coloured blocks represents different aspects.



(a) (b)

Fig. 3: Selected Papers: a) by Types. b) by Years and Types.

structured and semi-structured information, reasoning, and learning. Furthermore, [SP20], [SP21] and [SP22] indicate the cognition layer of a network edge is AI-embedded coordination for decision making by combining the information from different platforms towards orchestration objectives.

5. **Others:** [SP23] and [SP24] refer cognitive computing to AI models for the recognition and characterization of human physiological, emotional and cognitive states.

4.2 The Evolution of the Cognitive Cloud Definition

The first definition regarding cognitive management system, which was later adapted and connected to the cloud, is the one that anchors it related to radio access technologies in 2012 [SP1]. Then after two years, cognitive system were mentioned in different contexts ([SP2], [SP3]) where, however, the meaning of "cognitive" is similar regarding the system's learning capability. In the following year, the anchoring of cognitive systems concept in the perspectives of radio access technologies, AI involvement and architecture continued. Regardless the contexts, the indication of entities being "cognitive" meaning intelligent computing on large volume of data supporting smart decision making and optimization remains ([SP4] - [SP6]). Meanwhile, the perception of the environment and the interaction with human senses is added in terms of AI involvement. In 2016, the perspective of client-centric awareness was brought up ([SP8]) when the adoption of "cognitive" in cloud offloading in terms of RATs continued ([SP7]).

In 2018, the concept of "cognitive radio" started to draw more attention academically, where the notion of "cognitive" means being capable of sensing environment and learning opportunistically and dynamically remain ([SP10] - [SP12]). On the other hand, the adoption of "cognitive" towards cloud offloading continued in 2018 and 2019 where smartness and awareness in computation is emphasized ([SP9], [SP16]). Meanwhile, the concept of "cognitive computing" started to anchor the "cognitive" notion more to the perspective of AI ([SP13] - [SP15], [SP18]). Especially when entering 2021, when the notion of "cognitive" was more connected to the cloud computing domain, the AI involvement perspective of the concept became dominant ([SP20] - [SP22]).

5 Discussion

Five groups of definitions of Cognitive Cloud have been proposed in the last years. However, only few have been used or extended in other works.

Our results show that Cognitive Cloud cannot be considered a re-branding of a wider concept such as multi-agent systems, which does not specifically include cloud systems, and therefore would deserve a dedicated definition.

Based on the analysis conducted in this work, we can propose a new definition of cognitive cloud, combining the most frequently mentioned aspects. Therefore, we define cognitive cloud as: *a Cloud-based system that is capable of sensing its environment, learning from it, and opportunistically and dynamically adapt its computational load as well as its outcome.*

This definition of Cognitive Cloud can be seen as a synthesis of the multi-aspects of the cognitive concept in the domain of cloud computing. It emphasizes that for cloud entities being "cognitive" indicates the capabilities of environment sensing and learning, and opportunistical and dynamic adaptation of computational load. This definition shall simply clarify what Cognitive Cloud is and can help the researchers and industrial practitioners better understand this concept and its theoretical contexts.

5.1 Threats to Validity

We are aware that our work is subject to threats to validity. The terms Cognitive, Cloud are sufficiently stable to be used as search strings. In order to assure the retrieval of all papers on the selected topic, we searched broadly in general publication databases, which index most well-reputed publications.

To improve the reliability of this work, we defined search terms and applied procedures that can be replicated by others. Since this is a mapping study and no systematic review, the inclusion/exclusion criteria are only related to whether the topic of Cognitive Cloud is present in a paper or not, as suggested by [13].

As for the analysis procedure, since our analysis only uses descriptive statistics, the threats are minimal. However, we are aware that the synthesis of the definition might be subjective. To mitigate this threat, the analysis was done collaboratively, using a collecting coding method, and discussing with all the authors about inconsistencies. The Kohen K index about our disagreement also confirms the quality of the qualitative analysis performed.

6 Conclusion

In this work, we proposed a systematic mapping study on the definition of Cognitive Cloud to obtain an overview of existing definitions and understand how they evolved. We identified 24 main articles and we found promising results in terms of state-of-the-art definitions, definition of Cognitive Cloud and its key aspects. Starting from the initial definitions proposed in 2012, we found five groups of definitions, considering cognitive in the contexts of radio access technologies, client-centric awareness, architectural, involving AI and AI for human cognition. In conclusion, we suggest to complement existing definitions into a common one, that merges both aspects explicitly and enable both practitioners and researchers to better understand the concept of continuum.

Bibliography

[1] Funding & tender opportunities: Cognitive cloud: Ai-enabled computing continuum from cloud to edge (ria). https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-cl4-2022-data-01-02, accessed: 2022-06-16

[2] Abeshu, A., Chilamkurti, N.: Deep learning: the frontier for distributed attack detection in fog-to-things computing. IEEE Communications Magazine **56**(2), 169–175 (2018)

[3] Appel, A.P., Candello, H., Gandour, F.L.: Cognitive computing: Where big data is driving us. In: Handbook of Big Data Technologies, pp. 807–850. Springer (2017)

[4] Armbrust, M., Fox, A., et al.: Above the clouds: A berkeley view of cloud computing. Tech. rep., Technical Report UCB/EECS-2009-28, Uni. of California (2009)

[5] Chen, M., Herrera, F., Hwang, K.: Cognitive computing: architecture, technologies and intelligent applications. Ieee Access **6**, 19774–19783 (2018)

[6] Chen, M., Li, W., Fortino, G., Hao, Y., Hu, L., Humar, I.: A dynamic service migration mechanism in edge cognitive computing. ACM Transactions on Internet Technology (TOIT) **19**(2), 1–15 (2019)

[7] NIST Cloud Computing Standards Roadmap (July 2013), [online] https://www.nist.gov/system/files/documents/itl/cloud/NIST_SP-500-291_Version-2_2013_June18_FINAL.pdf

[8] Emam, K.E.: Benchmarking kappa: Interrater agreement in software process assessments. Empirical Software Engineering **4**(2), 113–133 (1999)

[9] Lenarduzzi, V., Taibi, D.: Mvp explained: A systematic mapping study on the definitions of minimal viable product. In: Euromicro SEAA. pp. 112–119 (2016)

[10] Mell, P., Grance, T., et al.: The nist definition of cloud computing. Tech. rep., Computer Security Division, Information Technology Laboratory (2011)

[11] Modha, D.S., Ananthanarayanan, R., Esser, S.K., Ndirango, A., Sherbondy, A.J., Singh, R.: Cognitive computing. Communications of the ACM **54**(8), 62–71 (2011)

[12] Naik, G., Choudhury, B., Park, J.M.: Ieee 802.11 bd & 5g nr v2x: Evolution of radio access technologies for v2x communications. IEEE Access **7**, 70169–70184 (2019)

[13] Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic mapping studies in software engineering. In: EASE 12. pp. 1–10 (2008)

[14] Wang, Y.: On cognitive informatics. Brain and Mind **4**(2), 151–167 (2003)

Appendix A: the Selected Papers

[SP1] Georgakopoulos, A. et al. (2012). Cognitive cloud-oriented wireless networks for the Future Internet. In 2012 WCNCW (pp. 431-435). IEEE.

[SP2] Cai, W. et al. (2014). Environment Perception for Cognitive Cloud Gaming. In International Conference on Cloud Computing (pp. 3-13). Springer, Cham.

[SP3] Cai, W. et al. (2014). Resource management for cognitive cloud gaming. In 2014 ICC (pp. 3456-3461). IEEE.

[SP4] Cordeschi, N. et al. (2015). Reliable adaptive resource management for cognitive cloud vehicular networks. IEEE Trans. on Vehicular Tech., 64(6), 2528-2537.

[SP5] Shi, W. (2015). QoE guarantee scheme based on cooperative cognitive cloud and opportunistic weight particle swarm. Electrical and Computer Engineering, 2015.

[SP6] Baughman, A. K. et al. (2015). Disruptive innovation: Large scale multimedia data mining. In Multimedia Data Mining and Analytics (pp. 3-28). Springer, Cham.

[SP7] Mahmoodi, S. E. et al. (2016). A time-adaptive heuristic for cognitive cloud offloading in multi-RAT enabled wireless devices. IEEE Transactions on Cognitive Communications and Networking, 2(2), 194-207.

[SP8] Chiang, M., et al. (2016). Fog and IoT: An overview of research opportunities. IEEE Internet of things journal, 3(6), 854-864.

[SP9] Mahmoodi, S. E. et al. (2018). Cognitive Cloud Offloading Using Multiple Radios. In Spectrum-Aware Mobile Computing (pp. 23-33). Springer, Cham.

[SP10] Wu, X. et al. (2018). Phase-compensation-based cooperative spectrum sensing algorithm for cognitive cloud networks. In 2018 ICOIN (pp. 755-759). IEEE.

[SP11] Wang, L. et al. (2018). Cooperative Spectrum Sensing Algorithm Based on Phase Compensation in Cognitive Cloud Networks. ICUFN (pp. 143-147).

[SP12] Huang, H. et al. (2018). On the Performance of Cognitive Cloud Radio Access Networks in the Presence of Hardware Impairment. In APSIPA ASC (427-431).

[SP13] Marshall, T. E. et al. (2018). Cloud-based intelligent accounting applications: accounting task automation using IBM watson cognitive computing. Journal of Emerging Technologies in Accounting, 15(1), 199-215.

[SP14] Jann, J. et al. (2018). IBM POWER9 system software. IBM Journal of Research and Development, 62(4/5), 6-1.

[SP15] Kloeckner, K. et al. (2018). Building a cognitive platform for the managed IT services lifecycle. IBM Journal of Research and Development, 62(1), 8-1.

[SP16] Mahmoodi, S. E. et al. (2019). Time-Adaptive and Cognitive Cloud Offloading Using Multiple Radios. In Spectrum-Aware Mobile Computing (pp. 49-66).

[SP17] Garai, Á. et al. (2019). Revolutionizing healthcare with IoT and cognitive, cloud-based telemedicine. Acta Polytechnica Hungarica, 16(2), 163-181.

[SP18] Amato, F. et al. (2019). A federation of cognitive cloud services for trusting data sources. In CISIS (pp. 1022-1031).

[SP19] Ferrer, A. J. et al. (2021). Towards a Cognitive Compute Continuum: An Architecture for Ad-Hoc Self-Managed Swarms. In CCGrid (pp. 634-641).

[SP20] Campolo, C. et al. (2021). Virtualizing AI at the distributed edge towards intelligent IoT applications. Journal of Sensor and Actuator Networks, 10(1), 13.

[SP21] Vermesan, O. et al. (2021). Internet of Vehicles–System of Systems Distributed Intelligence for Mobility Applications. In Intelligent Technologies for Internet of Vehicles (pp. 93-147). Springer, Cham.

[SP22] Kretsis, A. et al. (2021). SERRANO: Transparent Application Deployment in a Secure, Accelerated and Cognitive Cloud Continuum. In MeditCom (pp. 55-60)

[SP23] Bacciu, D. et al. (2021). TEACHING-Trustworthy autonomous cyber-physical applications through human-centred intelligence. In 2021 COINS (pp. 1-6).

[SP24] Zhang, H. et al. (2021). Knowledge-based systems for blockchain-based cognitive cloud computing model for security purposes. International Journal of Modeling, Simulation, and Scientific Computing, 2241002.