ECAI
DIGITAL – 2020

1st Doctoral Consortium at the European Conference on Artificial Intelligence (DC-ECAI 2020)

Proceedings

29-30 August, 2020
Santiago de Compostela, Spain

Preface

Welcome to the Proceedings of the 1st Doctoral Consortium at the European Conference on Artificial Intelligence (DC-ECAI 2020)!

This Doctoral Consortium (DC) is under the umbrella of the Mentoring and Communication for Starting Researchers (MC4SR) ECAI Program. In addition to the DC, MC4SR included other events such as Meeting with a EurAI Fellow, Job Fair, and the 9th European Starting AI Researchers' Symposium (STAIRS).

The DC-ECAI 2020 provides a unique opportunity for PhD students, who are close to finishing their doctorate research, to interact with experienced researchers in the field. Senior members of the community are assigned as mentors for each group of students based on the student's research or similarity of research interests. The DC-ECAI 2020, which is held virtually this year, allows students from all over the world to present their research and discuss their ongoing research and career plans with their mentor, to do networking with other participants, and to receive training and mentoring about career planning and career options.

The Programme includes two main sessions:

- Training Session (August 29th). The focus of this session is on training transversal and communication skills. The session has 60 attendants who will learn how to present their work in oral presentations with an emphasis in the practical case of a scientific elevator pitch for both the academy and industry. The session is led by Martin Holm Mortensen (FundingBox Research).

- Mentoring Session (August 30th). This session is aimed at dissemination of the ongoing research developed by Starting AI Researchers. We have 40 speakers who will present and leverage their work to the AI community with brief 3-minute video presentations. Peers and senior AI researchers will attend discussion sessions and provide speakers with feedback and guidance on future research directions. The three best presentations are to be awarded.

We want to thank all attendants and especially to speakers and mentors for contributing to this Doctoral Consortium.

Ulises Cortés and José M. Alonso
DC-ECAI 2020 Chairs

Doctoral Consortium Chairs:

Prof. Ulises Cortés, Barcelona Supercomputing Center and Univ. Politècnica de Catalunya
Dr. José M. Alonso, CiTIUS, University of Santiago de Compostela

Publication Chair:

Dr. Alejandro Catala, CiTIUS, University of Santiago de Compostela

Awards Committee:

Prof. Ramón López de Mántaras, IIIA-CSIC
Prof. Virginia Dignum, Umea University
Prof. Luis Magdalena, Unversidad Politécnica de Madrid
Prof. Ulises Cortés, Barcelona Supercomputing Center and Univ. Politècnica de Catalunya
Dr. José M. Alonso, CiTIUS, University of Santiago de Compostela

Mentors:

Prof. Karina Gibert, Univ. Politècnica de Catalunya
Prof. Ulises Cortés, Barcelona Supercomputing Center and Univ. Politècnica de Catalunya
Dr. Pedro Meseguer, IIIA-CSIC
Dr. Jaume Manero, Barcelona Supercomputing Center
Dr. Manuel Mucientes Molina, CiTIUS, University of Santiago de Compostela
Dr. Manuel Fernández Delgado, CiTIUS, University of Santiago de Compostela
Dr. Alejandro Catala, CiTIUS, University of Santiago de Compostela
Dr. José M. Alonso, CiTIUS, University of Santiago de Compostela

Lecturer:

Martin Holm Mortensen, FundingBox Research

Doctoral Consortium - Training

The Training part in the DC-ECAI 2020 takes the form of a webinar that is split into two sessions to be held by August 29th. These sessions are devoted to training transversal and communication skills. The attendants are expected to learn how to present their work in oral presentations with an emphasis in the practical case of a scientific elevator pitch for both the academy and industry.

The session is led by Martin Holm Mortensen who is Senior Project Manager at FundingBox Research; Master of Science in European & International Politics from The University of Edinburgh and Ba. in Public Administration from Roskilde University. He has 20 years of experience in startup management as a founder, c-level manager and on the Board of startups like ComputerPeople (Nordics), CapaSystems (USA + Denmark), CopenhagenDowntown (Denmark) and Clippingpath.com (UK + Bangladesh). During the last 7 years Martin has changed focus towards helping startups with commercialization, sales and investments externally as a Project Manager, Consultant and Investment Manager in some of Europa's best accelerators and growth projects like the Fundingbox lead EU program "IMPACT Growth", the Nordic accelerator "Accelerace Management" and the UK accelerator "NBG" in Nottingham.

Doctoral Consortium - Mentoring

The Mentoring part in the DC-ECAI 2020 is organized into five Mentoring Tracks (MT), to be held by August 30th. MTs have chairs who lead the session, introducing each speaker and handling questions and answering. There are 8 speakers in each track. Each speaker has a slot of 15 min for a brief personal introduction which is followed by an informal conversation regarding current and future research work.

- MT1: Knowledge Representation and Reasoning (Chair: Karina Gibert)

    - Munyque Mittelmann, "Auction Description Language (ADL): a general framework for representing auction-based markets"
    - Periklis Mantenoglou, "Online probabilistic interval-based event calculus"
    - Fernando E. Casado, "Learning from the individuals and the crowd in a society of devices"
    - Gianluca Zaza, "Monitoring cardiovascular risk by facial video processing and fuzzy rules"
    - Kenneth Skiba, "Towards a ranking-based semantics using system Z"
    - Ignacio Huitzil, "Advanced management of fuzzy semantic information"
    - Gönül Ayci, "Handling uncertainty and risk in privacy prediction"
    - Marcelo de Souza, "On the automatic configuration of algorithms"

- MT2: Agents, Planning and Scheduling (Chair: Jaume Manero)

    - Francisco J. Gil-Gala, "Evolving priority rules for scheduling problems by means of hyper-heuristics"
    - Justin Svegliato, "A metareasoning framework for planning and execution in autonomous systems"
    - Marcel Tiator, "Smart object segmentation to enhance the creation of interactive environments"
    - Norah Aldahash, "Improving agent interaction safety through environmental control and regulation"
    - Stanislav Sitanskiy, "Recognition and learning of agent behaviour models"
    - Gabriele Sartor, "Integrating open-ended learning and planning for long-term autonomy"
    - Jeferson José Baqueta, "A computational trust model for task delegation"
    - Anas Shrinah, "Verification and validation of planning-based autonomous systems"

- MT3: Natural Language Technology (Chairs: Jose M. Alonso and Alejandro Catala)

    - Kervadec Corentin, "Improving visual reasoning in visual question answering systems"
    - Weilai Xu, "Stylistic dialogue generation in narratives"
    - Aaron Keesing, "Improving robustness of emotion recognition in multi-accent and multilingual speech"
    - Andrea Cascallar Fuentes, "Fuzzy quantified protoforms for data-to-text systems: a new model with applications"
    - Chaina Santos Oliveira, "A two-level item response theory model to evaluate automatic speech recognition systems"
    - Damián Ariel Furman, "Hate speech analysis via argumentation schemes mining"
    - Yago Fontenla-Seco, "A framework for the automatic generation of natural language descriptions of processes"

- Carlos Andres Lopez Jaramillo, "A model for adopting the omnichannel strategy from a context-aware computing and natural language processing approach"

- MT4: Explainable and Trustworthy AI (Chairs: Pedro Meseguer and Ulises Cortés)

  - Gunay Kazimzade, "Structures and infrastructures around AI: unbiasedness, politics and metrics in data-driven socio-technical systems"
  - Ilse Verdiesen, "Accountability and control over autonomous weapon systems: a framework for comprehensive human oversight"
  - Joyjit Chatterjee, "Explainable AI for intelligent decision support in operations and maintenance of wind turbines"
  - Raúl Antonio del Águila Escobar, "OBOE: an explainable text classification framework"
  - Ettore Mariotti, "Understandable deep learning analysis for very high energy astroparticle physics"
  - Conor Hennessy, "Explaining Bayesian networks in natural language"
  - Ilia Stepin, "Argumentation-based interactive factual and counterfactual explanation generation"
  - Lucas Morillo-Méndez, "Age-related individual differences and their implications for the design of cognitive interactive systems"

- MT5: Big Data and Deep Machine Learning (Chairs: Manuel Mucientes Molina and Manuel Fernández Delgado)

  - Aiste Gerybaite, "Big data in IoE: investigating IT approaches to big data in healthcare whilst ensuring the competing interests of the right to health and the right to privacy"
  - Jorge García-González, "Deep learning neural networks to detect anomalies in video sequences"
  - Russa Biswas, "Embedding based link prediction for knowledge graph completion"
  - Ana Vieira, "Using machine learning for the personalized healthcare of vascular diseases"
  - Hemant Rathore, "Adversarial attacks on Android malware detection models using reinforcement learning"
  - Noelia Rico, "Ranking rules as a tool for reducing the impact of the distance in machine learning methods"
  - Armand Vilalta, "Semantic embeddings in deep convolutional neural networks"
  - Bartosz Piotrowski, "Reinforcement learning for saturation-based theorem provers"

The 40 speakers had provided previously brief 3-minute video presentations which were evaluated by the Awards Committee in terms of:

- Clarity: Did the presentation clearly describe the impact and/or results of the research, including conclusions and outcomes?

- Significance: Is the Thesis significant to the field? Are the communicated results coherent?

- Communication: Did the presenter capture and maintain their audience's attention?

Two weeks in advance to the DC-ECAI 2020 Mentoring session, the Awards Committee screened the 10 finalists in the DC-ECAI 2020 Best Presentation Contest:

- Periklis Mantenoglou (NCSR Demokritos)

- Corentin Kervadec (Universite de Lyon)

- Fernando E. Casado (CiTIUS-USC)

- Ettore Mariotti (CiTIUS-USC)

- Norah Aldahash (University of York)

- Weilai Xu (Bournemouth University)

- Gabriele Sartor (Universita degli Studi di Torino)

- Ignacio Huitzil (University of Zaragoza)

- Noelia Rico (University of Oviedo)

- Armand Vilalta (Barcelona Supercomputing Center)

The three winners will be announced in the DC-ECAI 2020 closing session. In addition to a diploma, the winners will be awarded with a 3-months research stay at CiTIUS-USC, including travel and accommodation expenses, as well as 700 eur per month for subsistence allowance.

Doctoral Consortium - Program (*)

(*) Time is expressed in GMT+2/CEST.

August 29th, 2020

9:00–9:10   **Welcome and Opening**

9:10–10:30   **DC Training (Session 1)**

10:30–10:45 **Break**

10:45–12:15 **DC Training (Session 2)**

August 30th, 2020

13:45–14:00 **Introduction to Mentoring**

14:00–15:00 **DC Mentoring (Session 1)**

14:00–14:15   Munyque Mittelmann (MT1)

                   Francisco J. Gil-Gala (MT2)

                   Kervadec Corentin (MT3)

                   Gunay Kazimzade (MT4)

                   Aiste Gerybaite (MT5)

14:15–14:30   Periklis Mantenoglou (MT1)

                   Justin Svegliato (MT2)

                   Weilai Xu (MT3)

                   Ilse Verdiesen (MT4)

                   Jorge García-González (MT5)

14:30–14:45   Fernando E. Casado (MT1)

Marcel Tiator (MT2)

Aaron Keesing (MT3)

Joyjit Chatterjee (MT4)

Russa Biswas (MT5)

14:45–15:00   Gianluca Zaza (MT1)

Norah Aldahash (MT2)

Andrea Cascallar Fuentes (MT3)

Raúl Antonio del Águila Escobar (MT4)

Ana Vieira (MT5)

15:00–15:30   Break

15:30–16:30   DC Mentoring (Session 2)

15:30–15:45   Kenneth Skiba (MT1)

Stanislav Sitanskiy (MT2)

Chaina Santos Oliveira (MT3)

Ettore Mariotti (MT4)

Hemant Rathore (MT5)

15:45–16:00   Ignacio Huitzil (MT1)

Gabriele Sartor (MT2)

Damián Ariel Furman (MT3)

Conor Hennessy (MT4)

Noelia Rico (MT5)

16:00–16:15   Gönül Ayci (MT1)

                  Jeferson José Baqueta (MT2)

                  Yago Fontenla-Seco (MT3)

                  Ilia Stepin (MT4)

                  Armand Vilalta (MT5)

16:15–16:30   Marcelo de Souza (MT1)

                  Anas Shrinah (MT2)

                  Carlos Andrés López Jaramillo (MT3)

                  Lucas Morillo-Méndez (MT4)

                  Bartosz Piotrowski (MT5)

16:30–17:00   Concluding Remarks and Farewell

Table of Contents

Big Data in IoE: investigating IT approaches to Big Data in healthcare whilst ensuring the competing interests of the right to health and the right to privacy

Aiste Gerybaite[1]

Abstract. Big Data has been vastly applied in sports domain while Big Data in IoE in healthcare has only recently started gaining traction. Sophisticated IoE devices are now able to process Big Data in order to monitor undesirable events through real-time alerting such as crash of vital signs with wearable wireless sensors, domestic accidents involving elderly people, or assist in predicting and monitoring outbreaks of various diseases which may put at risk public health as a whole. Due to the multidimensional nature of healthcare emergencies, the use of Big Data in such emergencies poses a number questions, not only with respect to the precise definition of an emergency, the agencies involved, the procedures used, but also questions with respect to securing privacy and the right to health in such emergencies without hindering the potential benefits of the development of Big Data solutions within healthcare sector.

1   PROJECT DESCRIPTION

1.1   Research problem

The correlation between Big Data and data protection is undeniable. Big Data in healthcare brings about several data protection and privacy challenges. Firstly, many of the classic fair information principles and data protection principles are being challenged by Big Data. In fact, van Der Sloot argues that the current data protection regime, based on the principle of data minimization, undermines Big Data's potential as in Big Data scenarios the objective is to collect as much data as possible and store data for as long as possible (for example, in healthcare research Big Data analytics are used to find trends and treatments that have the highest rate of success in the real world for cancer treatment). Further, Big Data tests the principles of *ratione personae* and *ratione materiae*[1]. In the first instance, Big Data do not focus on specific individuals but rather on groups of individuals/everyone, while in the latter, it becomes unclear whether a particular right (right to data protection) is even involved in certain Big Data scenarios. This raises a question on whether data protection and privacy in Big Data scenarios should take into account the collective dimension of data protection rights rather than focus on the individual dimension of such rights[2]. Additionally, since Big Data processing is often transnational, the rapidly changing technology may circumvent the applicable data protection provisions. Consequently, many questions arise whether the definitions of personal data, anonymized data, sensitive personal data are still tenable in the age of Big Data where ICT approaches tend to be more and more dynamic while the law remains faithful to its static nature.

In healthcare sector, the research to date focuses on the development of the ICT solutions for the sector (various medical devices, tracing devices) or on the regulatory requirements applicable to the sector. Yet, Big Data, privacy and data protection issues tend to be overlooked. In particular, the healthcare sector tends to overlook the two dimensions of healthcare emergencies, the public health and the individual dimension, and their implications to data protection and privacy. The public dimension of healthcare emergencies refers to emergencies such as global outbreaks of diseases (such as Covid-19, SARS etc). The latter, instead, refers to loss of vital signs by an individual which would not qualify as a public health emergency but, nevertheless, could have devastating impact on an individual's wellbeing.

The compendium of data aggregated by various actors in such emergencies though the use various health-monitoring, medical devices, or data generated and captured through various tracing apps and other healthcare software systems proliferates both concerns and opportunities. Indeed, a number of research focuses on developing WBAN, Primary Mobile Device (PMD), and Internet using Bluetooth, ZigBee IEEE 802.15.4, or 3G/LTE/Wi-Fi communication technologies-based ICT solutions[3]. Nonetheless, integrating real-time and similar health monitoring data has significant ethical[4], legal and social implications on individual's informational privacy due to the abstract and value based nature of law which is built on compromise. Instead, "Big Data is empirical, algorithmic, and deterministic. Also, Big Data is inherently acontextual. Big Data cannot interpret itself, nor can it discern the indeterminate boundaries of legal principles[5]".

Since the use of Big Data depends on the context it is used in, from the perspective of healthcare Big Data Shameer argues that the ownership rights of the data subject and the control of how patients may opt-in and opt-out are important for building real-time and similar health monitoring systems[6]. Such systems would need to be secure in order to ensure high degree of protection of personal and sensitive data in order to adhere to data protection requirements. This is particularly pertinent in the European scenario where the General Data Protection Regulation (GDPR) is applicable. In this

---

[1] CIRSFID, University of Bologna/ Law Department of the University of Torino/ Department of Computer Science of the University of Luxembourg; email: aiste.gerybaite@unito.it

regard, an in-depth analysis of the scope and application of articles 5,6 and 9 of the GDPR is required.

Likewise, ransomware and hacking attacks do not only work on data that a healthcare organisation already holds but also when various devices and health apps transmit, collect and share the data between one another or with a healthcare organisation. Thus, the processing of such personal and sensitive data may lead to data privacy violations. According to Cohen privacy violations can be divided into two categories, consequentialist and deontological concerns[7]. Whilst the first point concerns the negative consequences that the data breach carries with it, the second one does not depend on experiencing the negative consequences (ie. publication of sensitive information on an open web, which inadvertently allows third parties to gain access to such data).Whilst most research encountered addresses the potential breach in terms data privacy under the existing regulatory regime, few assess to what extent the existing ICT solutions in healthcare already provide for data privacy and data security.

In conclusion, the current state of the art undoubtedly indicates a gap between the appropriate translation and application of fundamental legal research into concrete scenarios with specific ICT technologies used in healthcare sector. As the ethical-legal research world focuses on fostering a high-level discussion on Big Data, healthcare and IoE, the ICT sector wonders what all this means to their specific scenario. Finally, the researcher believes that through the proper analysis of case-specific scenarios a further path for the regulation of Big Data and data protection may be paved.

2   RESEARCH OBJECTIVES

The research project investigates ICT approaches to the processing of Big Data in healthcare emergencies focusing on the two dimensions of such emergencies. In particular, the research focuses on applications used to assist and manage pandemics, and on wearables used for tracking person's vital signs and how data is processed and treated in healthcare emergencies. Such analysis permits to provide a multidimensional look at the core issues from both legal and technical standpoints for Big Data in healthcare emergencies with possible solutions for the issues. The objective therefore is to set ground for future regulatory regime which would facilitate the sharing of personal and personal sensitive data in a secure manner that would allow to implement ICT solutions within healthcare sector. Finally, the researcher believes that certain existing ICT tools used in healthcare may already be capable of guaranteeing data protection and privacy.

The research project will engage a discussion on the two angles of Big Data, the regulatory side and the technical side:

1. on the regulatory side the project analyses what part data protection legislation plays within the development of Big Data applications in healthcare sector. The research further looks into the precise notions of an emergency from both empirical and legal standpoints; the notions of data privacy vs data security; and securing privacy and the right to health in such emergencies without hindering the potential benefits of the development of Big Data solutions within healthcare sector.
2. The research considers to what extent the current design guidelines for wearables and various applications in healthcare take into account the applicable data protection framework.
3. The research explores how to balance the competing rights and interests of the affected parties with respect to data privacy and

the right to health. In particular, through looking at the specific Big Data applications within healthcare, the research attempts to answer the question of how the benefits of Big Data in healthcare can be captured in a way that respects fundamental principles of ethics and the rights to privacy and health.

In essence, the approach taken within research aims to look at the Big Data in healthcare and regulations with a novel perspective, ie. evaluate specific ICT solutions within the specific legislative framework. If the research is successful, the proposed research would benefit a wide range of stakeholders including legislators, scholars, practitioners, and consumers. The project is aimed and bridging the gap between the law and the ICT solutions within healthcare sector with respect to the EU applicable regulations and the ICT tools developed within the healthcare sector.

3   AUTHOR

Aiste Gerybaite is a doctoral researcher in the Law, Science and Technology Joint Doctorate - Rights of Internet of Everything, funded by Marie Skłodowska-Curie Actions. Her doctoral research focuses on Big Data for Health in IoE in emergency situations from the perspective of the competing interests of the right to health and the right to privacy.

ACKNOWLEDGEMENTS

REFERENCES

[1] B. Van der Sloot and A. De Groot, *The handbook of privacy studies: An interdisciplinary introduction*. Amsterdam University Press, 2018.

[2] L. Floridi, 'Open Data, Data Protection, and Group Privacy', *Philos. Technol.*, vol. 27, no. 1, pp. 1–3, Feb. 2014.

[3] M. M. Rathore, A. Ahmad, A. Paul, J. Wan, and D. Zhang, 'Real-time Medical Emergency Response System: Exploiting IoT and Big Data for Public Health', *J. Med. Syst.*, vol. 40, no. 12, p. 283, Dec. 2016.

[4] L. Floridi *et al.*, 'Key Ethical Challenges in the European Medical Information Framework', *Minds Mach.*, vol. 29, no. 3, pp. 355–371, Sep. 2019.

[5] B. van der Sloot and S. van Schendel, 'Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study', *J. Intellect. Prop. Inf. Technol. Electron. Commer. Law*, vol. 7, 2016.

[6] K. Shameer, M. A. Badgeley, R. Miotto, B. S. Glicksberg, J. W. Morgan, and J. T. Dudley, 'Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams', *Brief. Bioinform.*, vol. 18, no. 1, pp. 105–124, Jan. 2017.

[7] W. N. Price and I. G. Cohen, 'Privacy in the age of medical big data', *Nat. Med.*, vol. 25, no. 1, pp. 37–43, Jan. 2019.