1- TITLE:

AI-enabled Processes: The Age of Artificial Intelligence and Big Data

2- NAME AND CONTACT INFORMATION OF THE ORGANIZERS

Prof. Amin Beheshti, Scientia Prof. Boualem Benatallah, Prof. Michael Sheng, Prof. Fabio Casati, Prof. Hamid-Reza Motahari Nezhad, Prof. Jian Yang, and Prof. Aditya Ghose.

3- ABSTRACT:

Business processes, i.e., set of coordinated tasks and activities carried out manually/automatically to achieve a business objective or goal, are central to the operation of public and private enterprises. Modern processes are often extremely complex, data-driven and knowledge-intensive. In such processes, it is not sufficient to focus on data storage/analysis; and the knowledge workers will need to collect, understand and relate the big data (from open, private, social and IoT data islands) to process analysis. Today, the advancement in Artificial Intelligence (AI) and Data Science has the potential to transform business processes in fundamental ways; by assisting knowledge workers in communicating analysis findings, supporting evidences and to make decisions. This tutorial gives an overview on Services in Organizations, Business, and Society. We introduce notions of Data Lake as a Service and Knowledge Lake as a Service, and discuss their role in analyzing data-driven and knowledge-intensive processes in the age of Artificial Intelligence and Big Data. We introduce the novel notion of AI-enabled Processes and discuss methods for building intelligent Data Lakes and Knowledge Lakes as the foundation for Process Automation and Cognitive Augmentation in Processes. The tutorial also points out challenges and research opportunities.

4- REQUISITE BACKGROUND OF THE TARGET AUDIENCE AND OVERVIEW OF THE TOPIC COVERED

With data science continuing to emerge as a powerful differentiator across industries, almost every organization is now focused on understanding their business and transform data into actionable insights. For example, governments derive insights from vastly growing private, open and social data for improving government services, such as to personalize the advertisements in elections, improve government services, predict intelligence activities and to improve national security and public health. In this context, organizing vast amount of data gathered from various private/open data islands, i.e., Data Lake, will facilitate dealing with a collection of independently-managed datasets (from relational to NoSQL), diversity of formats and non-standard data models.

The notion of a Data Lake has been coined to address this challenge and to convey the concept of a centralized repository containing limitless amounts of raw (or minimally curated) data stored in various data islands. The rationale behind a Data Lake is to store raw data and let the data analyst decide how to cook/curate them later. While Data Lakes, do a great job in organizing big data and providing answers on known questions, the main challenges are to understand the potentially interconnected data stored in various data islands and to prepare them for analytics.

Previously, we presented the notion of Knowledge Lake, i.e., a contextualized Data Lake. The term Knowledge refers to a set of facts, information, and insights extracted from the raw data using data curation techniques such as extraction, linking, summarization, annotation,

enrichment, classification and more. The Knowledge Lake will provide the foundation for big data analytics by automatically curating the raw data in the Data Lake and to prepare them for deriving insights.

This tutorial gives an overview on Services in Organizations, Business, and Society; and introduce data-driven and knowledge-intensive processes in the age of Artificial Intelligence and Big Data. We introduce the novel notion of AI-enabled Processes and discuss methods for building intelligent Data Lakes and Knowledge Lakes as the foundation for Process Automation and Cognitive Augmentation in Processes. The tutorial also points out challenges and research opportunities.

5 - THE TUTORIAL LENGTH:

1 hour 50 minutes

6- TUTORIAL OUTLINE: USING A BULLET LIST, WITH ESTIMATED PRESENTATION TIME FOR EACH ITEM, IN THE ORDER OF PRESENTATION

- ***Overview: Services in Organizations, Business, and Society (10 min)***
  - In this part we provide an overview on Services in Organizations, Business, and Society; and introduce main topics in Business Process Management.
- ***Data Lake as a Service:*** *Data-centric Processes* ***(20 min)***
  - In this part we provide an introduction to data and its cross-cutting aspects (e.g., provenance and versioning) and then introduce the notion of Big Data and how it is different from large datasets.
  - We also introduce data curation, i.e., tasks that transform raw data (unstructured, semi-structured and structured data sources, e.g., text, video, image data sets) into curated data (contextualized data and knowledge that is maintained and made available for use by end-users and applications).
- ***Knowledge Lake as a Service:*** *Knowledge-Intensive Processes* ***(20 min)***
  - In this part we introduce the novel notion of Knowledge Lake, i.e., a contextualized Data Lake, and discuss big-data methods for building Knowledge Lakes being assets for big-data applications.
- ***AI-enabled Processes (45 min)***
  - In this part we introduce the notion of Artificial Intelligent, and discuss how AI can facilitate Process Automation and Cognitive Augmentation in Processes.
  - We also discuss topics including: Processes-aware chatbots, Intelligent RPA, and Document-centered processes
- ***Motivating Scenario: Police Investigation in the Age of AI and Big Data (15 min)***
  - Modern police investigation processes are often extremely complex, data-driven and knowledge-intensive. In such processes, it is not sufficient to focus on data storage and data analysis; and the knowledge workers (e.g., investigators) will need to collect, understand and relate the big data (scattered across various systems and generated by many Internet enabled devices such as CCTVs, Police cars, camera on officers on duty and more) to process analysis. In this part we use this motivating scenario to explain how AI-enabled Processes can assist police investigators in communicating analysis findings, supporting evidences and to make decisions.

7- SHORT BIOS OF THE PRESENTERS INCLUDING THEIR EXPERTISE RELATED TO THE TUTORIAL (NO MORE THAN 100 WORDS PER PRESENTER)

Prof. Amin Beheshti, *Macquarie University, Sydney, Australia*

- Prof. Amin Beheshti is a Full Professor of Data Science and the Director of AI-enabled Processes (AIP) Research Centre, School of Computing, Macquarie University. Amin is also the head of the Data Analytics Research Lab and Adjunct Academic in Computer Science at UNSW Sydney. Amin completed his Ph.D. and Postdoc in Computer Science and Engineering at UNSW Sydney and hold a Master and Bachelor in Computer Science both with First Class Honours. In addition to his contribution to teaching activities, Amin extensively contributed to research projects; where he was the R&D Team Lead and Key Researcher in the 'Case Walls & Data Curation Foundry' and 'Big Data for Intelligence' projects. Amin has been recognized as a high-quality researcher in Big-Data/Data/Process Analytics and has been invited to serve and served as Keynote Speaker, General-Chair, PC-Chair, Organisation-Chair, and program committee member of top international conferences. He is the leading author of the book entitled "Process Analytics", co-authored with other high-profile researchers in UNSW and IBM research, recently published by Springer. Amin was able to secure over $9.8 Million Research Grants for AI-Enabled, Data-Driven, and Intelligence-Led projects.

Scientia Prof. Boualem Benatallah, *University of New South Wales, Sydney, Australia*

- Prof. Boualem Benatallah is Scientia Professor at the University of New South Wales Sydney, Australia. His research interests lie in the areas of Web service protocols analysis and management, enterprise services integration, large scale and autonomous data sharing, process modelling and service oriented architectures for pervasive computing. He has several ARC (Australian Research Council) funded projects in these areas. He was a visiting scholar at Purdue University (USA), Visiting Professor at INRIA-LORIA (France), Visiting Professor at Blaise Pascal University (Clermont-Ferrand, France). He has been a General chair, PC chair and Program Committee member of several top conferences. He has been keynote and tutorial speaker on Web Services at several workshops and conferences. He has published widely in top international journals and conferences.

Prof. Michael Sheng, *Macquarie University, Sydney, Australia*

- Prof. Michael Sheng is a full Professor and Head of Department of Computing at Macquarie University. Michael holds a PhD degree in computer science from the University of New South Wales (UNSW) and did his post-doc as a research scientist at CSIRO ICT Centre. Michael has more than 320 publications as edited books and proceedings, refereed book chapters, and refereed technical papers in top journals and conferences. Michael is the recipient of the ARC Future Fellowship (2014), Chris Wallace Award for Outstanding Research Contribution (2012), and Microsoft Research Fellowship (2003). He is a member of the IEEE and the ACM.

Prof. Fabio Casati, *Servicenow, Silicon Valley , USA*

- Prof. Fabio Casati is a Principal Machine Learning Engineer at Servicenow, on leave from the University of Trento. Fabio focuses on several phases of the AI solution lifecycle, ranging from designing algorithms that can be effective at scale, to AI devOps in a PaaS context, and to helping customers be successful in the adoption of AI in their workflows. Previously, he was Professor at the University of Trento, where he worked on crowdsourcing and hybrid human-machine computations, and researcher at Hewlett-Packard Labs, working on business process intelligence. He co-authored a best-selling book on Web services and is the author of over 250 peer-reviewed papers

Prof. Hamid-Reza Motahari Nezhad, *UpBrains AI, Inc., Silicon Valley, USA*

- Prof. Hamid Motahari is Founder and CEO of UpBrains AI, Inc., which is a startup with a focus on intelligent business process automation for the enterprise. He is an Honorary Professor of Computer Science position at Department of Computing at Macquarie University, Australia. Before starting UpBrains AI, Hamid served as Head of AI Science at EY AI Lab, Palo Alto, in EY Global Technology Innovation where he was leading the global AI Science team working in Document Intelligence and Enterprise AI. His background and interest lie in applied AI, document intelligence, conversational AI, intelligent RPA, and AI in business process automation. Prior to EY AI Lab, Hamid was serving as Research Lead for Cognitive Services at IBM Research, and was a member of IBM Academy of Technology. He has published more than 100 scholarly papers in various scholarly conferences. Hamid has served as chair, invited speaker and organizer of multiple data analytics, and AI-focused conferences.

Prof, Jian Yang, *Macquarie University, Sydney, Australia*

- Prof. Jian Yang is a full professor and director of research of Department of Computing at Macquarie University. Her research interests include data driven business process management, data and service integration, big data analytics and graph mining, and their applications in smart city and ageing and aged care. Her research is supported by ARC (Australian Research Council), government agencies, and industry. She was a visiting professor at Beijing University of Post and Telecommunication, Jilin University of Economics and Finance. She has been a Steering Community member of ICSOC since 2013. She has been a General chair, PC chair and Program Committee member of ICSOC and BPM. He has been keynote and tutorial speaker on Web Services at ICSOC and several workshops and conferences. She has been publishing extensively in top international journals and conferences.

Prof. Aditya Ghose, *University of Wollongong, Wollongong, Australia*

- Prof. Aditya Ghose is Director of the Decision Systems Lab at the University of Wollongong, a Fellow of the Institution of Engineers (Australia) and current President of the Service Science Society of Australia. He holds a PhD and MSc in Computing Science from the University of Alberta, Canada (he also spent parts of his PhD candidature at the University of Illinois at Urbana-Champaign and the University of Tokyo), and a Bachelor of Computer Science and Engineering from Jadavpur

University, Kolkata, India. His research has been funded by the Australian Research Council (3 Discovery Projects and 5 Linkage Projects), the Japanese Research Institute for Advanced IT, the Canadian NSERC and via engagements with organisations such as Bluescope Steel, BHP Building Products, DXC, IBM Research, Xerox Research, the CRC for Smart Services, Infosys Labs, the Australian Defence Science and Technology Group (DSTG), Data61 and BAE Systems.

8- A LIST OF UP TO 20 MOST IMPORTANT REFERENCES TO BE COVERED IN THE TUTORIAL

1. Beheshti, A., Benatallah, B., Nouri, R., Chhieng, V.M., Xiong, H., Zhao, X.: CoreDB: a data lake service. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017, Singapore, November 06 - 10, 2017.
2. Beheshti, A., Benatallah, B., Nouri, R., Tabebordbar, A.: CoreKG: a knowledge lake service. PVLDB, 11(12), 2018.
3. Beheshti A., Benatallah B., Tabebordbar A., Motahari-Nezhad H., Barukh M., and Nouri R., "DataSynapse: A Social Data Curation Foundry". Distributed and Parallel Databases (DAPD) Journal, 37(3): 351-384, 2019.
4. Barukh, M. Zamanirad S., Báez M., Beheshti A., Benatallah B., Casati F, Yao L., Sheng M., Schiliro F.: Cognitive Augmentation in Processes. Next-Gen Digital Services, 123-137, 2021.
5. Beheshti A., Yakhchi S., Mousaeirad S., Ghafari S., Goluguri S., Edrisi M: Towards Cognitive Recommender Systems. Algorithms 13(8): 176, 2020.
6. Beheshti, A., Benatallah, B., Nezhad, H.: ProcessAtlas: A scalable and extensible platform for business process analytics. Softw., Pract. Exper. 48(4), 2018.
7. Beheshti, A., Schiliro, F., Ghodratnama, S., Amouzgar, F., Benatallah, B., Yang, J., Sheng, Q., Casati, F., Motahari-Nezhad, H., iProcess: Enabling IoT Platforms in Data-Driven Knowledge-Intensive Processes, BPM Forum, 2018.
8. Beheshti, S., Tabebordbar, A., Benatallah, B., Nouri, R.: On automating basic data curation tasks. In: WWW, 2017.
9. Beheshti, S., Benatallah, B., Motahari-Nezhad, H.R.: Scalable graph-based OLAP analytics over process execution data. Distributed and Parallel Databases 34(3), 379-423, 2016.
10. Beheshti, S., Benatallah, B., Sakr, S., Grigori, D., Motahari-Nezhad, H.R., Barukh, M.C., Gater, A., Ryu, S.H.: Process Analytics - Concepts and Techniques for Querying and Analyzing Process Data. Springer Book, 2016.
11. Beheshti, S., Benatallah, B., Nezhad, H.R.M.: Enabling the analysis of cross-cutting aspects in ad-hoc processes. In: CAiSE. 2013.
12. Tene, Omer, and Jules Polonetsky. "Big data for all: Privacy and user control in the age of analytics." Nw. J. Tech. & Intell. Prop. 11, 2012.
13. Terrizzano, Ignacio G., et al. "Data Wrangling: The Challenging Yourney from the Wild to the Lake." CIDR. 2015.
14. Santiputri, M., Ghose, A.K. and Dam, H.K., 2017. Mining task post-conditions: Automating the acquisition of process semantics. Data & Knowledge Engineering, 109, pp.112-125.
15. Santipuri, M., Ghose, A., Dam, H.K. and Roy, S., 2017, November. Goal orchestrations: Modelling and mining flexible business processes. In Proc. of ER-2017 (pp. 373-387), 2017.
16. Gou, Y., Ghose, A. and Dam, H.K., 2019, June. GameOfFlows: Process Instance Adaptation in Complex, Dynamic and Potentially Adversarial Domains. In Proc. of CAISE-2019 (pp. 479-493), 2019.
17. Dan G. Tecuci, Ravi Palla, Hamid R. Motahari Nezhad, Nishchal Ahuja, Alex Monteiro, Tigran Ishkhanov, Nigel P. Duffy: DICR: AI Assisted, Adaptive Platform for Contract Review. AAAI: 13638-13639, 2020.