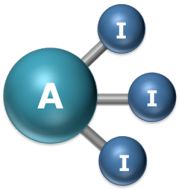

Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery

PLA2019, 7th Edition
9th - 10th April 2019
Grand Hotel Dino, Lake Maggiore, Italy

Dr Samantha Kanza & Dr Nicola Knight
University of Southampton

21/12/2021

AI3SD-Event-Series:Report-9

PLA2019, 7th Edition
AI3SD-Event-Series:Report-9
21/12/2021
DOI: 10.5258/SOTON/P0011
Published by University of Southampton

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery
This Network+ is EPSRC Funded under Grant No: EP/S000356/1

Principal Investigator: *Professor Jeremy Frey*
Co-Investigator: *Professor Mahesan Niranjan*
Network+ Coordinator: *Dr Samantha Kanza*

Contents

1 Event Details **1**

2 Event Summary and Format **1**

3 Event Background **1**

4 Talks **2**
4.1 The Lab of the Future 2
4.2 FAIR Data, AI & Semantics 2

5 Workshops **3**

6 Expert Panel **4**

7 Exhibitors **5**

8 Participants **5**

9 Conclusions **5**

10 Related Events **5**

1 Event Details

Title	PLA2019, 7th Edition	
Organisers	Paperless Lab Academy	
Dates	9th - 10th April 2019	
Programme	Programme	
No. Participants	250	
Location	Grand Hotel Dino, Lake Maggiore, Italy	
Local Chairs	Isabel Muñoz-Willery Ph.D. Roberto Castelnovo	
Committee	PLA Organsiation	
Sponsors	Partners	Lonza AgiLab
	Media Partners	List of Media Partners

2 Event Summary and Format

This event was run by the Paperless Lab Academy which is organized by NL42 Consulting and is the 7th iteration of this congress, this year hosted again at Lake Maggiore in Italy. The central theme for this edition of the conference was #eDataLifeCycle @Work – turn your laboratory into a data-driven knowledge center.

The congress was a two-day event comprised of a number of presentations, interactive workshops, and an expert panel discussion session. It was also possible to attend pre-congress training sessions on the day before the event. In the main congress the talks all ran consecutively so it was possible to attend each one, although the workshops ran 2-3 to a workshop session. There was plenty of time for networking alongside the talks with generous coffee breaks, lunch and a networking cocktail session on the first night. Alongside the main hall there was also an exhibition hall where industry vendors had stands and were able to present their offerings with respect to paperless lab solutions. This hall was open during all the networking breaks and so it was possible to talk to and receive software demos from the wide range of vendors present at this event.

3 Event Background

The Paperless Lab Academy project has been in existence since 2013 and has been running webinars and annual congresses since to further their goals of providing a learning platform for parties interested in 'consolidating, integrating and simplifying scientific data management systems and processes.' Each congress has a central theme with the talks surrounding the key aspects of that theme. The Paperless Lab Academy is free to institutions and companies which have a laboratory who are looking to potentially implement digital transformations or go paperless in their research. Attendees come from a wide range of backgrounds including pharmaceuticals, food and drink, cosmetics, chemicals, biotech and more. It also attracts a range of participants across the employment spectrum, from Laboratory managers, IT experts,

QC managers, researchers, etc.

4 Talks

A number of topics were covered over the 2 days of the conference. The conference opened with a keymore from Sergio Nasi from Boehringer Ingelheim, a LIMS company that focuses on pharmaceuticals and cancer research. Sergio detailed how his company originally tried to produce LIMS on mobile palm devices, before the world was ready for that type of technology, and emphasised the importance of understanding your user requirements before blindly developing software and hardware (lessons echoed by Peter Cohan from The Second Derivative who presented on the importance of being driven by the users needs). Sergio posed provocatively that nothing overly disruptive has happened in the world of laboratory informatics in the last 20 years, but now things are starting to change. The cost of data storage has reduced, as the capacity of computational power has increased which has driven innovation in many areas, concluding that technology has and wil continue to drive the laboratory informatics journey in the future. The main themes represented in these talks were: The Lab of the Future and Fair Data, AI & Semantics.

4.1 The Lab of the Future

Following the theme of technology driving innovation, there were a number of presentations detailing digital transformations of laboratories and visions of the 'Lab of the Future'. Pat Pijanowski from Accenture detailed their vision of unleashing the power of digital technologies in the laboratory, suggesting that the laboratory should function 'as a node in a highly connected and collaborative network of laboratories' and likened the journey to the lab of the future to that of grocery shopping where digital technologies are being readily developed. A scenario was proposed where one could walk into a laboratory with their list of things (experiment plan), be recognised for who they are (user identification), walk up to an instrument, setup an experiment and come back to download the results later. Pat emphasised that we are closer to this than we might think! Rik Pepermans from Unilever presented on the Digital Knowledge Revolution, using examples of digital holiday booking systems to demonstrate how knowledge can be transformed from peoples heads into code in order to digitally transform a process. On a similar vein Michael Shanler from Gartner presented on 'The Digital Laboratory: build a path to the future' and laid out criteria for digital labs of the future, including the use of semantic web technologies, digital twins, IoT connected devices, and increasing intelligent automation whereby systems can learn from their data. Naturally, part and parcel of that was also using standards and data principles such as FAIR.

4.2 FAIR Data, AI & Semantics

Markus Ruffener from Lonza presented a case study on historical datasets, demonstrating how a global setup of datasets resulted in multiple instances of the same dataset, and the issues that caused. He emphasised the importance of all working from the same datasets rather than having many copies where things can get lost or changed. Eric Little, the Chief Data Officer for Osthus continued the theme of FAIR with his presentation on "FAIR Data, Data Catalogs and the Foundations of AI". Eric detailed the importance of FAIR, how you need to be able to find data and access it, but you also need it to be interoperable which is arguably the hardest bit of FAIR to achieve, noting that reusable should come naturally if the first three principles are adhered to. Eric presented a very favourable view of leaving data where it is, and embracing your data silos. In alignment with Markus, he noted that we have far too much data and that companies and academics alike are spending far too much time moving their data around, making copies of it, and trying to work out how to use it and integrate it

together, which is both a time consuming and expensive task. Eric promoted the idea of a data catalog for your data silos. The data catalog would describe what is in each different data source and how to get to them, thus enabling searching for the data. So the catalog is really just a portal to searching the data, but the references and links need to be correct or the links won't work. Eric and his colleagues are working on methods to achieve this combining Semantic Web technologies (knowledge graphs), with statistics and machine learning. Eric concludes by stating that Semantics is the missing link when discussing AI, and in order to make systems like this work we need Semantics and Statistics.

5 Workshops

14 workshops were run across the two days with a variety of topics. A number of the workshops were run by sponsors or vendors were exhibiting at the event.

- **AgiLab – Unify your GXP and NON-GXP Laboratory processes to support the complete development cycle:** This workshop demonstrated how to use the same system to manage different processes, with a large focus on the advantages of using Agile methods over the traditional waterfall methods for development, and on re-using and combining existing technology where possible.
- **Lonza – Practical applications of Data Integrity and Audit Trail Review:** This workshop looked at how to understand your audit trail and review the data contained in it.
- **Gilson – The Connected Scientist: More Productive Research through Smarter Lab Tools:** This workshop focused on traceability and reproducibility with IoT products.
- **Labware – Demystifying the Cloud: How will it help my laboratory?:** This workshop looked at building the bridge between business and computing.
- **Wega Informatik – Agile Validation in GMP Projects:** This workshop also focused on agile approaches for GxP/GAMP projects.
- **Agaram Technologies – Pragmatic Approach to Achieve Data Integrity & Compliance:** This workshop focused on the common data integrity challenges and how these can be mitigated.
- **Agilent Technologies – Harnessing the power of modern IT to deliver laboratory efficiency:** This workshop discusses how laboratories can be improved with the changes in modern technology.
- **Dexstr – How INSIGHT ENGINES designed for life science can increase the efficiency of your lab?:** This workshop demonstrated the INSIGHT engine, showing how it takes in lots of different data sources, and extoling the INSIGHT principles: FOLLOW (be aware of your data and able to identify it), SEARCH (be able to search through your data and find what you need), EXPLOIT (be able to reorganise your documents to suit your needs) and SHARE (be able to share all your data internally and externally).
- **Dassault Systemes – The Truly Connected Lab is Here Today:** This workshop provides a demonstration of using the BIOVIA laboratory management system 'One LAB' to integrate the entore laboratory environment.
- **Osthus – Exploiting FAIR Data & Data Catalogs with LeapAnalysis: A New Approach to Virtualized Data Integration Across Federated Data Sources:** This workshop explained the new tool Leap Analysis and demonstrated how it worked,

showing how disparate data sources located in different places can be linked to via this system which then auto matches the data to relevant ontologies and facilitates SPARQL queries over the different sources.

- **Waters – The Process of Moving from Paper to Paperless (Louise Tobin – Teva):** This workshop showcases real life examples of moving from paper to paperless by using an ELN.
- **Global Value Web – The True Paperless Lab: Organizing for Today, Executing for Tomorrow:** This workshop looks at how laboratory performance can be improved using new technologies and by standardising lab processes.
- **Novatek International – Cleaning Validation in a Pharmaceutical Manufacturing Environment:** This workshop discusses the different methods and products that can be used to identify contamination and validate cleaning in the pharmaceutical environment.
- **PerkinElmer – Smarter Questions, Faster Answers: Accelerating your research with Signals Platform:** This workshop demonstrated the PerkinElmer Signals platform and how it can be used to make data driven decisions using diverse data sources.

6 Expert Panel

Following on from the talks and workshops there was a panel forum hosted by Peter Boogaard (Industrial Lab Automation). On the panel were five experts:

- Rik Pepermans – Unilever
- Michael Shandler – Gartner
- Mark Fish – Accenture
- Jernej Grmas – Novartis/Sandoz
- Sergio Nasi – Boehringer Ingelheim

This panel discussed three main questions, where to start with FAIR, what are the most important issues in the laboratory, and which technologies the panelists believed had the highest potential impact for the future.

With respect to FAIR, the response focused on creating one process that every dataset can use, using australian sheep shaving analogy whereby there is one hole and all the sheep that come through get shaved. At some point in the future there will be one process for FAIR that everyone can use.

There were a number of issues identified as very important in the lab, predominantly that there are different drivers depending on the type of laboratory which makes it difficult to identify one singular key issue. However, data integrity, reproducibility, and proper training were identified as key issues.

There was no solid agreement on the technology for the highest potential impact for the future, as this also depends on the situation. However Cloud and SAAS were identified as a potential for high impact, as were AI and Semantic Web technologies, and ultimately whatever leads to further digitalisation.

7 Exhibitors

Alongside the talks and workshops there was an exhibitors hall, which contained vendor stalls from a wide range of companies including lab automation, data management and equipment companies. This allowed ample opportunity for attendees to discover the solutions that the vendors had available along with potential software demonstrations.

8 Participants

There were many participants at this event, a majority of which were from industry and worked across different parts of the laboratory lifecycle, with a strong focus on those who worked in the digital / technology areas. There were also some academics there alongside our team who were interested in research data management and the digitisation of scientific research.

9 Conclusions

The main conclusions of this event is that there are a lot of people talking about FAIR and Semantics and these standards are becoming more prevelant in the world of digital laboratory transformation. At the heart of our laboratories is the vast amount of data that underpin them, and if we are going to see a real digital transformation then this data needs to be curated, transformed and made available in a way that is useful to others.

10 Related Events

Paperless Lab runs yearly and their latest events can be found here:
https://www.paperlesslabacademy.com/.