PRISME Forum

Pharmaceutical R & D Information Systems Management Executives

PRISME Forum

TECHNICAL MEETING

If data is the currency of R&D, how do we recognize its value?

PRISME Forum Chair:

Olivier Gien, *Sanofi*

PRISME Forum Technical Meeting Chair:

Dan Chapman, *UCB*

May 17-18, 2017

Berlin, Germany

Host: Bayer AG

Download and stay up-to-date with our

PRISME Forum Spring 2017 Technical Meeting App:

http://my.yapp.us/PRISMETECH

Meeting Venue

All sessions will be hosted by Bayer AG and held at their Berlin Campus, located at:
Müllerstraße 178, 13353 Berlin, Germany.

Hotel

Swissôtel Berlin, Augsburger Str. 44, 10789 Berlin, Germany.

Contacts

Program Coordinator:		+44.77.68.173.518/jcmwise@prismeforum.org
Secretariat	Office:	+1.224.938.9523/logistics@prismeforum.org
	On site, SMS:	+1.312.622.1234

PRISME Forum Technical Meeting Advisory Committee

Olivier Gien (ex officio)
Chair, PRISME Forum
Global Head, Clinical IT, *Sanofi*

John Apathy
Vice President, R&D Informatics, *Celgene*

Dan Chapman
Chair, PRISME Forum Technical Meeting
Director, Discovery Research IM, *UCB*

Edsel Calliste-David
Global Lead, Business Relationship & Solutions, IS, *Astellas*

Lars Greiffenberg
Director R&DIR and Translational Informatics, *AbbVie*

Martin Leach
VP, R&D IT, GOIT, EDMA, Alexion *Pharmaceuticals*

James McGurk
Sr. Director, Data Architecture, R&D Informatics
Daiichi-Sankyo

Simon Roach
SVP, R&D IT, *GlaxoSmithKline*

Alex Schuleit
Senior R&D Business IT Development Manager, *H. Lundbeck A/S*

Susie Stephens
Senior Director, R&D BT, *Pfizer, Inc*

Ashok Upadhyay
Senior Director, Global R&D IT, *Otsuka Pharmaceutical Development & Commercialization*

John Wise
Program Coordinator, *PRISME Forum*

PRISME Forum Host

The PRISME Forum Technical Meeting Advisory Committee would like to thank Bayer AG for hosting the 2017 PRISME Forum Spring meeting.

PRISME Forum Statement of Compliance

"All meetings, working groups and communications will be open to all Members and any records thereof will be non-confidential and available for inspection by any Member. The Members acknowledge that discussing any commercially sensitive topics, including costs, volumes, inventories, sales level methods, channels of distribution, access to future products, markets, current or future prices, profitability, ***contract pricing or trading terms*** is prohibited. The Members of PRISME will strictly comply with all laws relevant to their activities, including US state and federal anti-trust laws and European competition laws."

If Data Is the Currency of R&D, How Do We Recognize Its Value?

Analysis and visualization of internal and externally hosted data is critical to the on-going success of Pharma R&D as it searches for patterns, trends and sequences in the R&D processes to discover, develop and bring to market new drugs for unmet patient needs.

Data is the key currency for bio-pharmaceutical R&D. R&D generates, consumes, analyses and represents data for operational purposes as well as generating insights. The processes of R&D are data driven and should be clearly understood to provide clarity about the data backing the science and claims being made for medicines. The importance of data provenance, governance, curation, sharing and tracking should be a cultural norm for all scientists within R&D. In the rush to new scientific advances, concerns of data management, controls and processes are sometimes viewed as of secondary importance. Data without context is not useful and if used to make decisions, damaging. As the body of data grows, how does one understand and interpret how much governance is 'good enough'?

Data analytics is a continuing strategic theme for the biopharmaceutical sector as the industry is faced with an explosion of biomedical data being generated and made available for R&D use. Along with the increasing volumes of publicly available data, new entrants have been observed in the life science industry, including non-traditional companies such as Facebook, Amazon and Google. How will the biopharmaceutical industry ensure that these new and exciting technologies will process these data intelligently?

Advances in data analytics do not appear to have reduced the effort required to prepare the data for use. A recent survey of data scientists[1] suggests that data preparation is the most time consuming and least enjoyable aspect of their jobs. As the volumes and variety of data increase, the veracity, provenance, integrity – indeed the simple 'usefulness' of data becomes increasingly difficult to assess.

The focus of this PRISME forum Technical Meeting will be to explore classic and modern techniques and technologies that will aid biopharma R&D in collating, organizing, filtering, preparing and sharing data to provide the best opportunities to our scientists and operations in generating further value from our data – not least by collaboration.

[1] *http://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#4b32a76f7f75*

PROGRAM

All sessions will be held at the Bayer AG, Muellerstrasse 178, 13353 Berlin, Germany

WEDNESDAY, 17th May 2017		
18:15	*Gather in hotel lobby for departure to PRISME Forum Group Reception (Business Meeting delegates and Technical Meeting delegates)*	
THURSDAY, 18th May 2017		
08:00	*Gather in the hotel lobby for departure to the meeting venue*	
08:30	*Check-in; poster installation*	
08:45	**Welcome Notes**	**Errol Sandler,** Treasurer/Board Secretary, *PRISME Forum*
08:50	**Introductions**	**Dan Chapman,** Technical Meeting Chair, *PRISME Forum;* Director, Discovery Research IM, *UCB*
SESSION I: Keynote Presentations		**Chair: Dan Chapman, Director, Discovery Research IM, *UCB***
09:00	**The Data Catalog: The Key to Managing Data Big and Small**	**April Reeve**, Director, Enterprise Information Strategy & Architecture, *Celgene*
09:35	**AETIONOMY - Data, Disease Modeling and Reasoning**	**Martin Hofmann-Apitius,** Head, Department of Bioinformatics, *Fraunhofer SCAI*; Professor, *University of Bonn*
10:10	*Coffee Break*	
SESSION II: Perspectives: *Data Quality*		**Chair: Lars Greiffenberg, Director, R&D IT and Translational Informatics, *AbbVie***
10:40	*Brief introductions by panelists followed by moderated discussion*	
	Acquiring the Data - Instrumentation Perspective	**Eduardo Gonzalez-Couto,** BioInformatics Products Strategist & Manager, *PerkinElmer*
	No Data Left Behind: Using Fixed Instruction Sets to Provide Computational Closure Around Scientific Instrumentation and Data Acquisition	**DJ Kleinbaum,** Co-Founder, *Emerald Cloud Lab*
	Organization - Standards & Ontologies - The Allotrope Foundation Perspective	**Christoph Schaub,** Director, Product Supply IT Innovation, *Bayer AG* **Edith Scheringer,** Head, IT RDM-E Chemistry Infrastructure & Structural Analytics, *Boehringer-Ingelheim*
	IMI 2 Call 9 (EQIPD) Data Quality in Preclinical R&D	**Karsten Wicke,** Associate Director, *AbbVie*
	Moderated Panel Discussion/Q&A	
SESSION III A: Posters		**Chair: Simon Roach, SVP, R&D IT, *GlaxoSmithKline***
11:40	*Introduction - Poster Rotations (three 15-minute rotations)*	
P1	**Development and Use of Collaborative Curation Approaches**	**Alexandra Grebe de Barron,** Senior Information Architect, *Bayer AG*
P2	**Three Dimensions of Enterprise Clinical Informatics: Metadata, Events, Repository**	**Peter Smilansky,** SVP, Product Strategy, *EDETEK*
P3	**Data Quality and MDM – From Zero to Production in 9 Months**	**Khushi Sharma,** Account Executive, *Ataccama* **Tomas Bolek,** Presales Team Leader, *Ataccama*
P4	**Empowering Data Management and Analysis with Semantic Microservices**	**Robert Greenwood,** CEO, *SciBite*
P5	**Data Standards – The eTRIKS Perspective**	**Ibrahim Emam,** Bioinformatician, Data Science Institute, *Imperial College London*
12:30	*Lunch*	
SESSION III B: Posters		**Chair: Simon Roach, SVP, R&D IT, *GlaxoSmithKline***
13:30	*Poster Session (Remaining two 15-minute rotations)*	
SESSION IV: Plenary Presentations		**Chair: Andreas Friese, Head, R&D IT Research, *Bayer AG***
14:00	**RWD with A Focus for R&D Use - Challenges and Opportunities in Research and Early Development**	**Thomas Abbott,** Head, RWI Strategic Capabilities & External Alliances, *Astellas*
14:30	**The Roche Data Commons**	**Jan Küntzer,** Principal Scientist, *Roche Diagnostics*
SESSION V: Summing-Up Session		**Chair: John Apathy, VP, R&D Informatics, *Celgene***
15:00	**Breakout Session –** Members and meeting guests will be divided into four groups led by co-captains	**Facilitators:** Group 1: **Andreas Friese & Karsten Wicke** Group 2: **Ashok Upadhyay & Edith Scheringer** Group 3: **Martin Leach & DJ Kleinbaum** Group 4: **Simon Roach & Eduardo Gonzalez-Couto**
16:00	**Plenary Session –** Readouts from breakout groups	
SESSION VI: Meeting Summary & Awards & Reception		**Chair: Dan Chapman, Director, Discovery Research IM, *UCB***
16:30	Meeting Summary	
16:45	*Awards & Networking Reception*	
18:00	*Return to the hotel*	

BIOS AND ABSTRACTS

PRISME Forum Chair: Olivier Gien

Global Head, Clinical IT, *Sanofi*

Olivier Gien, PhD, was elected as the Chairman of the PRISME Forum at the November 2014 PRISME Forum Business Meeting.

Dr. Gien is the Global Head of Clinical IT at Sanofi. He is a Chemical Engineer by training and holds a PhD in Organic Chemistry. His PhD work focused on leveraging Artificial Intelligence technologies and retrosynthetic analysis to build a system helping chemists in the design of synthetic routes.

Dr. Gien started his career in the Exploratory Unit of Sanofi's Hungarian affiliate in Budapest then took charge of Information Systems for Industrial Chemical development at Sanofi's Sisteron site. He led then Global Discovery Research Information Systems at Sanofi-Synthelabo, then Sanofi-Aventis in Montpellier, before taking on his new role in Paris area in 2010.

Welcome Notes: Errol Sandler

Treasurer/Board Secretary, *PRISME Forum*

Errol Sandler, PhD, worked in the information technology industry for 30 years. His career focused on R&D computing problems in the life sciences as well as leadership and technical expertise for R&D computing in the pharmaceutical industry.

He led teams to provide information technology support at several Pfizer Global research and development sites in the United States and the United Kingdom.

Dr. Sandler received his PhD in Physics and Astronomy from the University of Missouri-Columbia.

PRISME Forum Technical Meeting Chair: Dan Chapman

Director, Discovery Research Information Management, *UCB*

Dan Chapman, PhD, is part of the leadership team within Informatics at UCB with responsibility for Software Development and Architecture and Therapeutic Informatics (UK).

Dr. Chapman has 15 years' experience working within the Pharmaceutical industry in a variety of roles.

After completing a PhD in Chemistry at Warwick University, he transitioned to informatics during post-doctoral research at Cambridge University as part of the CLIC consortium. Dr. Chapman joined AstraZeneca in 1997 and worked on a variety of global projects before joining UCB in his present role in 2005.

Since then, he has driven several projects to revolutionize the informatics platform within UCB and is currently actively involved in promoting Data Science across UCB.

SESSION I
KEYNOTE PRESENTATIONS

Session Chair: Dan Chapman
Director, Discovery Research Information Management, *UCB*

April Reeve
Director, Enterprise Information Strategy & Architecture, *Celgene*

April Reeve has spent the last 25 years working as an enterprise architect, program manager, and management consultant. She is now incorporating her enterprise data management knowledge working at Celgene Corporation as the Director of Enterprise Information Strategy & Architecture.

April is an expert in multiple Data disciplines including Data Integration, Big Data, Data Conversion, Data Warehousing, Business Intelligence, Master Data Management, and Data Governance. She has authored a book on Data Integration: "Managing Data in Motion" and the chapter on Data Integration and Interoperability in the Data Management Association DMBoK (Data Management Body of Knowledge) to be published this year.

The Data Catalog: The Key to Managing Data Big and Small

Analysts and industry experts agree that a huge risk to Data Lake and other Big Data projects is the lack of metadata management for physical inventory, business meaning, access security, audit trail, and general asset management. The Data Catalog is a critical factor to Big Data program success that must be implemented as additional pieces to the distributed data environment, since it is not inherent in Apache Hadoop.

The Enterprise Data Catalog is now an essential component of modern enterprise architecture and especially critical as part of any Big Data project. It is the solution that data users in our organizations can use to identify what data is available to them and how they can quickly request and get appropriate access.

Martin Hofmann-Apitius

Head, Department of Bioinformatics, *Fraunhofer SCAI;* Professor, *University of Bonn*

Martin Hofmann-Apitius holds a PhD in Molecular Biology and worked for more than 10 years in experimental molecular biology.

The screening for novel genes involved in tumor metastasis lead him into the area of functional genomics and subsequently to applied bioinformatics. Martin Hofmann-Apitius has experience in both, academic (University of Heidelberg (ZMBH), Forschungszentrum Karlsruhe (ITG), German Cancer Research Center (DKFZ)) and industrial (BASF, Boehringer Ingelheim, LION bioscience AG) research. Since 2002 he is leading the Department of Bioinformatics at the Fraunhofer Institute for Algorithms and Scientific Computing (SCAI) in Sankt Augustin (Germany), a governmental non-profit research institute. In July 2006 he has been appointed as a Professor for Applied Life Science Informatics at Bonn-Aachen International Center for Information Technology (B-IT).

Martin Hofmann-Apitius is co-author of more than 130 scientific publications. Major scientific contributions were the cloning and identification of CD44v, the first gene that mediates metastatic potential to tumour cells, the functional annotation of the mouse transriptome, and information extraction methodology used for the semi-automated generation of the first comprehensive, computable model for Alzheimer´s Disease.

Current research activities at the Department of Bioinformatics at Fraunhofer SCAI focus on:

- Automated methods for the extraction of relevant information from unstructured information sources such as journal publications, patents and web-based sources
- Knowledge-based, mechanistic modelling of neurodegenerative diseases
- Mining in real-world data (social networks, patient forums, electronic patient records)
- Scalable solution for unstructured information mining: HPC & cloud computing

Martin Hofmann-Apitius is initiator and academic coordinator of IMI project AETIONOMY (www.aetionomy.org) (industrial coordinator is Prof. Duncan McHale, UCB Pharma).

AETIONOMY - Data, Disease Modeling and Reasoning

Biomedical Research seems to run into an existential crisis. Neither Knowledge-driven, nor Data-driven approaches seem to hold the promise of "quantitative natural science"; namely reproducibility of results and insights. We always claim that it is possible to link the description of a scientific challenge or phenomenon to the design of dedicated experiments suited to enable insights such as correlation or causal links between variables and, most important - that the entire workflow from experiment to analysis and conclusion can be reproduced.
Recent reports on the very limited reproducibility of published data (and insights derived from the analysis of these data) in the area of drug target validation raises massive concerns on the validity of the current research paradigm and the research result communication process (including the established peer review process).

The 4th "V" in the Big Data concept, the "V" that stands for "veracity", seems to be the weakest and most challenging aspect of Big Data in Biomedicine.

In my talk, I will shed some light on the need for curation of public (and proprietary) data and knowledge. I will discuss concepts like "Fair Data" (which are aimed at improving re-usability of data, an essential component of veracity). Furthermore, I will sketch a possible route from "curated objects" in biomedicine (e.g. SwissProt) to curated, computable representations of knowledge in an indication-specific fashion. I will show, how "computable knowledge" can help to "validate" huge data sets at large scale and how the ability to rapidly compare patterns in data to representations of curated knowledge allows us to define the "ignorome", the whole of what we do not know.

Finally, I will show, how economic parameters ("price tags") can be assigned to representations of pathophysiology mechanisms in an indication-specific fashion.

SESSION II
PERSPECTIVES: Data Quality

Session Chair: Lars Greiffenberg

Director, R&D IT and Translational Informatics, *AbbVie*

Lars Greiffenberg holds a MS in Biology and a PhD in Microbiology and has more than 15 years of experience in the field of integrated R&D IT solutions and translational informatics. Dr. Greiffenberg held different R&D IT management positions at Aventis Pharma and Sanofi-Aventis in Frankfurt before relocating to the Sanofi site in Toulouse, France were he was Global Head of Solution Center Translational Medicine with responsibility to manage and lead a global program to enable translational science at Sanofi.

In 2014 he joined AbbVie in Ludwigshafen (Germany) as director of R&D IT and Translational Informatics. Besides the local role to lead the discovery and development IT department, he plays a global role in enabling Translational and Knowledge Science at AbbVie.

Eduardo Gonzalez-Couto

BioInformatics Products Strategist & Manager, *PerkinElmer*

Eduardo Gonzalez-Couto is currently a bioinformatics products strategist & manager at PerkinElmer Informatics. Previously he was CSO at Integromics and he has 20 years of professional experience in bioinformatics gained at GlaxoWellcome, GeneProt, and Siena Biotech.

Dr. Gonzalez-Couto received his PhD in Molecular Genetics after obtaining a Master in Molecular Biology and graduating in Biochemistry at the University of Geneva, Switzerland.

Acquiring the Data - Instrumentation Perspective

The life sciences domain is swiftly progressing into a data-centric discipline. In parallel, the life sciences biopharmaceutical industry and technology providers are evolving into digital organizations.

The high degree of automation, along with the massive and diverse data types rapidly produced by cutting-edge instruments or sensors represent an important challenge in this transformation. The need for coherent meta-data, defined ontologies, processing reproducibility and streaming data interpretation goes along with the adoption of cloud-based solutions, the usage of machine learning and AI approaches. In this new, and accelerating, context, the lab of the future will most likely give birth to paradigms like “science as a service” and “augmented scientists”.

DJ Kleinbaum

Co-Founder, *Emerald Cloud Lab*

DJ Kleinbaum, PhD, is the Co-founder and Co-CEO of Emerald Cloud Lab. He serves on the Swartz Center for Entrepreneurship Advisory Board at Carnegie Mellon University and advises a number of other start-up companies.

Dr. Kleinbaum holds a doctorate in nucleic acid chemistry from Stanford University and a BS in Biological Sciences with concentrations in Computational Biology and Biochemistry from Carnegie Mellon University.

No Data Left Behind: Using Fixed Instruction Sets to Provide Computational Closure Around Scientific Instrumentation and Data Acquisition

The recording and storage of experimental results is traditionally thought of as a fundamentally distinct activity from documenting the acquisition methods used to conduct the experiments. Recording information regarding the state of samples and their history, as well as the state of the instrumentation and its history, is often thought of as yet another separate and ancillary activity. With focus intensifying on issues of reproducibility and data quality in the industry, we at the Emerald Cloud Lab have designed systems that think of these activities as a holistic exercise -- where information like experimental methods, sample history, instrumentation history, and environmental conditions are woven into a richly interconnected linked data network along with the experimental results.

One the keys to making a system which treats methods like data work in practice has been to not think of this as an exercise in attempting to be as comprehensive as possible in retroactively documenting the variables and methods. Instead, that exercise is inverted to generate a fixed instruction set in advance for describing any modifications made to the state of the samples or instrumentation in the lab. Scientists use this instruction set to conduct all of their experiments and to handle the control and maintenance of instrumentation by entering commands on a computer. Providing this computational closure around experimental methods and the state of the laboratory provides a naturally error-correcting forcing function around the generation of data.

Christoph Schaub

Director, Product Supply IT Innovation, *Bayer AG*

Christoph Schaub, PhD, is responsible for IT innovation for the supply chain area at Bayer. In addition, Dr. Schaub is representing Bayer on the Board of Directors of Allotrope Foundation. Prior to this role, he was responsible for delivering IT solutions for RDIT in the early development field of pharma and consumer health.

Over the past 18 years, Dr. Schaub has held positions in various IT fields, especially in patent & scientific information, document management, health safety & environment, research & development and supply chain IT at Bayer.

Dr. Schaub holds a doctorate in bioorganic chemistry from University of Konstanz.

Edith Scheringer

Head of IT RDM-E Chemistry Infrastructure & Structural Analytics, *Boehringer-Ingelheim*

Edith Scheringer has worked at Boehringer Ingelheim for more than 20 years in R&D related IT units. Her background is in economics and business administration with a focus on process analysis.

Ms. Scheringer started her career with software development in different areas (DMPK, Drug Discovery Support, Analytic, Structure Elucidation) and led a number projects in these areas with a specific focus on the validation of systems.

She currently leads multidisciplinary, globally-focused IT teams with responsibilities for CMC workflows and currently for chemistry platforms, structure elucidation and several aspects of spectroscopy.

As of early 2017, Ms. Scheringer has been serving on Allotrope's Board of Directors.

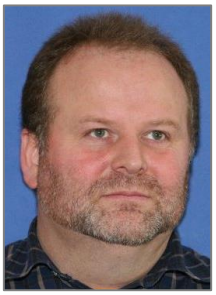
Organization - Standards & Ontologies - The Allotrope Foundation Perspective

Allotrope is a precompetitive collaborative foundation with 13 members from pharmaceutical industry, founded in 2012. The vision of Allotrope is to enable an intelligent, automated and integrated analytical laboratory by unlocking the power of scientific data produced in the labs. Therefore, Allotrope has built a fast growing partner network with more than 40 members, consisting of software and hardware vendors, IT integrators and academic institutions.

This will be done by providing a data standard, which could be seen as a common language for all software and instruments in a lab surrounding. A second key deliverable will be ontologies to provide the necessary context to the data. Thereby Allotrope creates the foundation to reuse the data produced in our laboratories for scientific progress and to implement data integrity by design. The first public release of the data standard is planned to be delivered by end of Q2 2017.

Karsten Wicke

Associate Director, Neuroscience Research, *AbbVie*



Karsten Wicke, PhD, is currently leading a core technology group in AbbVie's neuroscience discovery. He has over 25 years of experience in the pharmaceutical industry with a focus on neuroscience discovery, starting at Knoll AG, followed by BASF Pharma and Abbott.

Dr. Wicke has led various biological lab units and expertise in multiple technologies ranging from electrophysiology and behavioral science up to histology and CNS drug delivery systems.

He has been working in different disease areas within psychiatry and neurology and led multiple preclinical projects for the treatment of schizophrenia, depression, Parkinson's and Alzheimer's diseases, delivering clinical development candidates.

In his career, Dr. Wicke repeatedly worked in consortia of public-private partnership; his current focus is on an IMI2-call named "EQIPD", dealing with preclinical data quality.

Dr. Wicke holds a degree in Biology with focus on physiology from the Ruhr-University Bochum, Germany and is an inventor with over 15 patents.

IMI 2 Call 9 (EQIPD) Data Quality in Preclinical Research and Development

Data generation and analysis is a common and important topic in the biopharmaceutical industry. Late dropouts in pharmacological development generate massive loss of investment; the later such dropouts happen during development the more expensive they are. Therefore, preclinical studies are used to filter the best development candidates. However, false positive results are common leading to failing compounds in later stages of the clinical development. Over the past years, the reproducibility and quality of preclinical data has been questioned; studies suggest a rate of irreproducible data above 50%.

With a recent IMI2 consortial initiative we want to

a) develop generally applicable, lean and efficacious quality principles for biomedical research,
b) define a quality system fit for purpose, generate metrics to define success, implement and beta-test that system in the PhD program,
c) develop cross site criteria for audit outcomes,
d) review existing guidelines and white paper exercise to develop key principles for guiding the development of standard assays improving robustness, reproducibility and research efficiency.

This approach will help to increase preclinical data quality thereby improving the quality of decisions made based on such datasets.

SESSION III
POSTERS

Session Chair: Simon Roach

SVP, R&D IT, *GlaxoSmithKline*

Simon Roach has been the Senior Vice President of Research and Development IT at GlaxoSmithKline (GSK) since October 2008. Previously he served GSK as Vice President of Development and Commercial Strategy IT.

Mr. Roach joined the group from Coopers & Lybrand where he served as Senior Manager of Computer Assurance Services. Roach began his professional career with NCR as a Consultant from 1986-1989. Roach attended Aston University where he received his BSc in Computing Science and MBA.

Alexandra Grebe de Barron

Senior Information Architect, *Bayer AG*

Alexandra Grebe de Barron, PhD, joined the group of Computational Life Sciences IT at Bayer in 2015 as Senior Information Architect. She enables innovation and collaboration across the R&D functions by driving strategic, cross-functional initiatives. She strives towards making information FAIR (findable, accessible, interoperable, re-usable) by advising on technologies, integrating business processes, and developing and applying information architecture frameworks and practices.

Prior to joining Bayer Dr. Grebe de Barron worked 15 years as consultant in the biotech and pharmaceutical industry leading IT projects along the R&D value chain spanning from early research, early clinical development to clinical operations and regulatory affairs.

Dr. Grebe de Barron is biologist by training and holds a PhD in molecular genetics with research carried out at the universities of Milan, Italy, Boulder, Colorado, Konstanz and Berlin, Germany. After finishing her PhD thesis, she followed her interest in information technologies and qualified in software development to then translate and integrate information and processes at the interface of scientific discovery and IT.

P1: Development and Use of Collaborative Curation Approaches

Disruptive innovations in science, computing and digitalization enable and boost computational life sciences approaches. Even primarily product-centered industries like the pharma industry recognize data as the new currency. Despite the availability of vast amounts of data and advanced analytics tools, the translation of scientific progress into medical products serving patients remains relatively slow. One bottleneck is the provisioning of trustworthy and analysis-ready data. Finding the right data and preparing it for analysis takes a data scientist most of his working time. At Bayer we thus asked “How to make data curation work” using innovative approaches. We found out that data curation is only partially a technology problem but mainly centers around people, communication and collaboration. In the poster we will explain our approach using design thinking to tackle these topics, insights gained, first solutions and next steps.

Peter Smilansky

SVP, Product Strategy, *EDETEK*

Peter Smilansky has more than 25 years of Information Technology experience, all in the Life Sciences industry and R&D systems. Mr. Smilansky has lead development and implementation of clinical, regulatory and safety systems for a major pharmaceutical company, managed application architecture in a global R&D organization, directed clinical and preclinical IT in a biotechnology firm, and worked as lead clinical systems architect for a CRO.

Prior to joining EDETEK in 2015, he managed a clinical technology consulting practice for a global systems integrator.

P2: Three Dimensions of Enterprise Clinical Informatics: Metadata, Events, Repository

1) Benefits and use cases of meta-data driven clinical information management
2) Latest generation clinical MDR - a centralized repository of all enterprise clinical research standards (beyond CDISC)
3) Event-driven clinical systems to support Digital Clinical Trials and continuous study quality monitoring
4) Modern cloud-based Clinical Information Repository: A union of Clinical data lake and Clinical data warehouse

Khushi Sharma

Account Executive, *Ataccama*

As an Account Executive, Khushi Sharma is responsible for managing the North American accounts which include Alexion Pharmaceutical Inc.

Ms. Sharma communicates between the team of Ataccama consultants and the clients in order to provide the best solution and services based on the client's use case. She also develops and manages relationships with C-level executives, key partners and customers.

Tomas Bolek

Presales Team Leader, *Ataccama*

As a Presales Team Leader, Tomas Bolek is responsible for Ataccama's global presales activities.

With his team of senior consultants, Mr. Bolek is designing and delivering complex Data Quality and Master Data Management projects and solutions to clients across various verticals. Healthcare & Pharma references include large organizations such as HCA, NHS, GSK, Novozymes, and Alexion.

P3: Data Quality and MDM – From Zero to Production in 9 Months

In order to facilitate enterprise analytics and stewardship, Alexion needed a tool that would simultaneously enhance their Data Quality capabilities and connect with their incumbent MDM system.

Join this session to learn:

- About Alexion's use case details and solution benefits.
- How this leading U.S. pharmaceutical company improves their data quality while enabling business users and data stewards to monitor Data Quality trends, and track and address discrepancies and data quality issues over time.
- How to start small and initiate your own data quality project using Ataccama free data profiling tool.

Robert Greenwood

CEO, *SciBite*

Robert Greenwood is the Chief Executive Officer of SciBite. SciBite is a young, exciting software company head quartered in Cambridge, UK with offices in Tokyo and Boston MA, helping its customers transform the way they unlock value from unstructured text. SciBite brings 'Semantics as a Service' to all science based businesses around the world in any IT environment. Mr. Greenwood oversees all aspects of SciBite putting the right team in place to support the growing use cases for existing and new customers.

Prior to his appointment at SciBite, he was the General Manager for BIOVIA for EMEA for 2 years. Previously to that, over a 10-year period, he held several senior roles within Thomson Reuters, leading to the Global Leadership of Life Science sales team across EMEA, NA and APAC.

Mr. Greenwood has over 20 years' experience managing global teams and supporting customers with world class data and software solutions for bioinformatics, genomics, competitive intelligence, clinical and IT.

P4: Empowering Data Management and Analysis with Semantic Microservices

Life science organizations are awash with data. While text content (such as literature, reports, ELNs and so on) represents the deepest knowledge, it is also often the poor relation with respect to data management and mining when compared to other structured data. SciBite's semantic services are designed to transform biomedical text into a first class citizen in today's data-driven scientific organizations. Using a cutting-edge 'microservices' approach, SciBite builds rich structured data that serves a vast range of use cases in machine learning/AI, drug repurposing, pharmacovigilance, competitor intelligence and beyond. SciBite's tools cast a lightweight semantic net across different data silos, integrating their content through structured semantic meaning. The flexible API is designed to be embedded in 3rd party applications, helping these applications "understand scientific language" and transforming the user experience. By providing these advanced capabilities as simple 'drop-in' functionality, the way internal and commercial applications handle biomedical content can be vastly improved, connecting more data and delivering deeper insights.

Ibrahim Emam

Bioinformatician, Data Science Institute, *Imperial College London*

Ibrahim Emam is a Research Associate at the Data Science Institute at Imperial College leading the development of IMI eTRIKS Harmonization Service Platform (eHS).

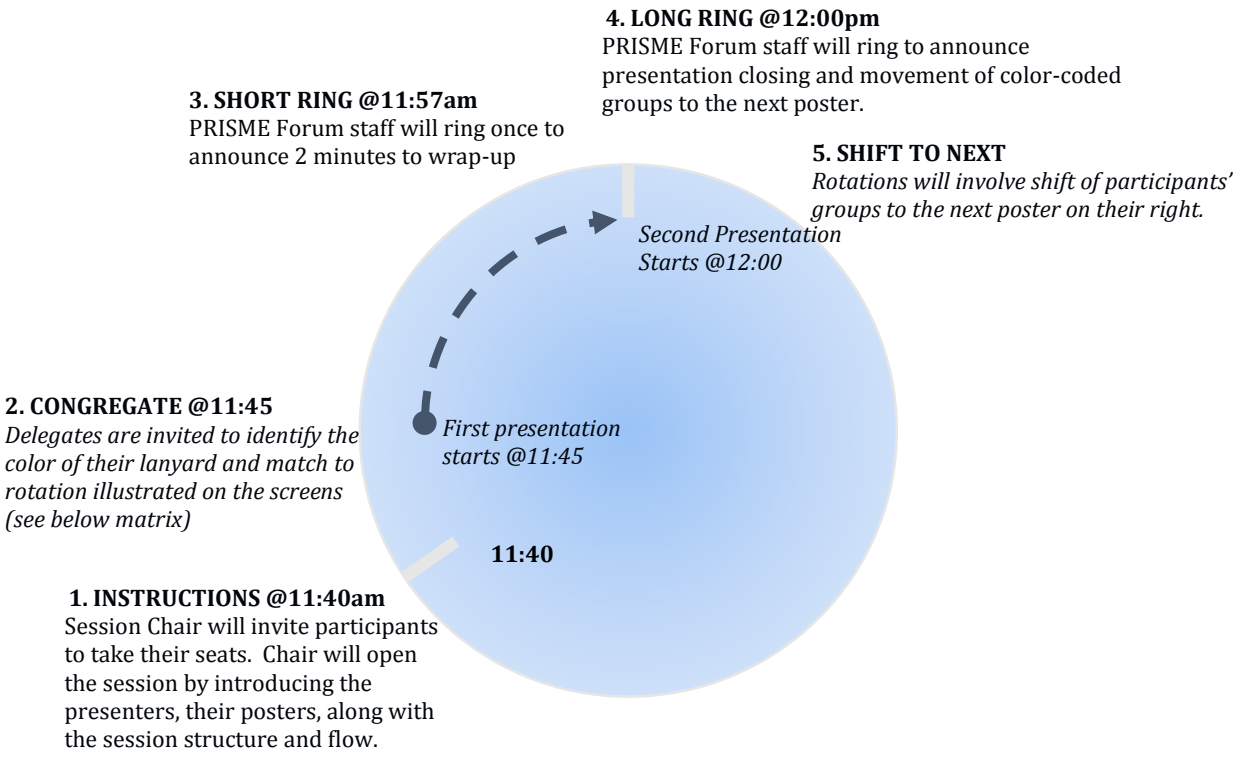
Proir to this, he was worked as a senior software engineer at the Functional Genomics group at the European Bioinformatics Institute developing the Gene Expression Atlas and ArrayExpress2.

Mr. Emam received his MSc in Bioinformatics from the University of Oxford and his dual BSc in Computer Science and Biology from the American University in Cairo in Egypt. His working experience revolves around bioinformatics application design and development, large scale data management and analysis, data standardization, modeling and visualization.

P5: Data Standards – The eTRIKS Perspective

A common feature of translational research studies is their approach to generate diverse types of biological and medical data. To realize the benefits of a translational approach in observing and integrating clinical and biomedical data it is pivotal to develop a unified metadata infrastructure. BioSPEAK-DB is a new standard compliant repository, part of the eTRIKS Harmonization Service platform (eHS) designed and developed at Imperial College Data Science Institute as the first standards compliant general-purpose data management system for translational research data. BioSPEAK-DB is modeled on a newly developed hybrid data model that integrates clinical and biomarker assays (omics) standards for reporting data and meta-data elements. This metadata model builds on already existing standard community initiatives in clinical and molecular (omics) data management. The study design is modeled in compliance with the CDISC Study/Trial Design Model (SDM) and the Protocol Representation Model (PRM), the clinical observations are modeled according to the CDISC general clinical observation model reported in the Study Data Tabulation Model (SDTM), and the assays and samples were modeled according to the ISA-TAB model for multi-omics metadata standards.

POSTER SESSION ROTATIONS



FIRST SET OF ROTATIONS:

ROTATION 1 - 11:45	ROTATION 2 - 12:00	ROTATION 3 - 12:15
P1 – Orange	P1 – Yellow	P1 – Red
P2 – Blue	P2 – Orange	P2 – Yellow
P3 – Green	P3 – Blue	P3 – Orange
P4 – Red	P4 – Green	P4 – Blue
P5 – Yellow	P5 – Red	P5 – Green

BREAK FOR LUNCH

SECOND SET OF ROTATIONS:

ROTATION 4 - 13:30	ROTATION 5 - 13:45
P1 – Green	P1 – Blue
P2 – Red	P2 – Green
P3 – Yellow	P3 – Red
P4 – Orange	P4 – Yellow
P5 – Blue	P5 – Orange

POSTER ROTATIONS *(lanyard colors)*

ORANGE	
Thomas Abbott	***Astellas***
Andrew Allen	***Regeneron***
Monika Arenz	***NIBR***
David Christie	***Amgen***
Devendra Deshmukh	***Merck***
Lars Greiffenberg	***AbbVie***
Alain Nanzer	***Roche***
April Reeve	***Celgene***
Simon Roach	***GlaxoSmithKline***
Jean-Luc Schmidt	***Sanofi***
BLUE	
John Apathy	***Celgene***
Andreas Friese	***Bayer***
Sabine Grote	***AbbVie***
Derek Marren	***Eli Lilly and Company***
Birgitte Mathiesen	***Novo Nordisk***
Keith Murphy	***Pfizer***
Michiel Ringkjøbing-Elema	***Lundbeck***
Ashok Upadhyay	***Otsuka***
Nick Wright	***AstraZeneca***
GREEN	
Dan Chapman	***UCB***
Klaus Hofenbitzer	***Celgene***
Jan Küntzer	***Roche***
Roy Ladd	***AbbVie***
Anders Mortensen	***Novo Nordisk***
Errol Sandler	***PRISME Forum***
Christoph Schaub	***Bayer***
Alex Schuleit	***Lundbeck***
Romain Taillard	***Roche***
Nicole van Poppel	***Accenture***
RED	
Beatrice Chapuzet	***Servier/Praxis***
Mary Donlan	***PerkinElmer***
Joel Ekstrom	***Ionis***
Martin Hofmann-Apitius	***Fraunhofer Society***
DJ Kleinbaum	***Emerald Cloud Lab***
Martin Leach	***Alexion***
Tomoyuki Matsunaga	***Takeda***
Rob Saiter	***Accenture***
YELLOW	
Massimo de Francesco	***UCB***
Eduardo Gonzalez-Couto	***PerkinElmer***
James Hanly	***BMS***
Rainer Kappes	***Sanofi***
Edith Scheringer	***Boehringer Ingelheim***
David Sedlock	***Takeda***
Pascual Starink	***Celgene***
Etzard Stolte	***Roche***
Karsten Wicke	***AbbVie***

SESSION IV
PLENARY PRESENTATIONS

Session Chair: Andreas Friese

Head, R&D IT Research, *Bayer AG*

Andreas Friese started his professional career as a software developer in 1987 at Schering AG, Berlin, Germany. From the beginning he was focused on IT solutions that addressed specific needs of the Research organization. Over the years, he held various positions as system analyst and project manager – all with focus on Research specific systems or projects.

In 1999, Mr. Friese moved to Richmond, CA, USA as an IT Business Partner for the Medicinal Chemistry department of Berlex Biosciences. During the merger of Schering AG with Bayer AG, he returned to Germany.

Based in Wuppertal, Mr. Friese is now Director of Research-IT for Bayer AG.

Thomas Abbott

Head, RWI Strategic Capabilities & External Alliances, *Astellas*

Thomas Abbott recently joined Astellas as Head of Real World Informatics & Analytics Capabilities and Alliances – a central function accountable for building Astellas's ability to leverage the emerging healthcare information ecosystem across its entire value chain.

Prior to joining Astellas, Dr. Abbott held leadership roles at HCL Technologies as Vice President Global Solutions, Life Sciences; Johnson & Johnson, as Head, Healthcare Informatics for Medical Devices and Diagnostics; Optum as SVP Analytics & Data Strategy, SVP Advanced Analytics, President I3-PharmaInformatics; WoltersKluwer (now Symphony Health Analytics) as VP Brand Analytics; Thomson-Medstat (now Truven Health Analytics) as VP Pharmaceutical Information Products & Services; and Merck as Senior Director Outcomes Research & Management, and Financial Evaluations & Analysis. Dr. Abbott began his career as Assistant Professor of Economics and Finance at the Graduate School of Management, Rutgers University.

Dr. Abbott completed his PhD in Economics at Harvard University, post-doctoral work in Health Services Research at Yale School of Medicine and holds an MBA from the Wharton School. Tom has over 40 publications in scholarly journals and books, and over 100 abstracts and conference presentations covering a wide range of topics including research methods, populations screening programs, treatment effectiveness and productivity measurement.

RWD with A Focus for R&D Use - Challenges and Opportunities in Research and Early Development

Research and Development has largely focused on using data from Clinical Trials to assess the safety and efficacy of New Molecular Entities (NMEs). However, recent advances in the availability and accessibility of Real World Data (RWD) sources have led to opportunities to leverage RWD to improve the efficiency of research and development.

This discussion will focus on some of these innovative uses of RWD for improving the targeting and execution of clinical development programs including: target selection and prioritization based on unmet clinical need; on identification of high-risk patient sub-groups for early POC analysis, and on the use of RWD for testing clinical trial design, country and site selection.

Jan Küntzer

Principal Scientist, *Roche Diagnostics GmbH*

Jan Küntzer, PhD, is a principal scientist in pRED Informatics at the Roche Research Center Penzberg. With a background in computer science and mathematics, he completed his PhD at Saarland University in Bioinformatics with a primary focus on Integrative Systems Biology. He spent a year at Leeds University in the UK and did a research fellowship at Kyoto University in Japan.

Dr. Küntzer joined Roche 2008 with a Roche Postdoc Fellowship working on a biological information system for cancer genome mutation. Within pRED Informatics, he is leading the Data Science Application Service, operating multiple bioinformatics, chemo-informatics, biostatistics, and information science applications used by Data Scientists to support pRED drug projects. He and his team designed and implement the architecture for the Roche Data Commons within pRED.

The Roche Data Commons

The amount of Roche molecular information data grows rapidly, which creates new challenges for both science as well as IT systems. We present the Roche Data Commons: a holistic approach for an agile future-oriented research environment to make molecular information data actionable.

SESSION V
BREAKOUT SESSION

Session Chair: John Apathy

VP, R&D Informatics, *Celgene*

John Apathy is the Vice President of R&D IT at Celgene Corporation. Mr. Apathy joined Celgene in 2013, leading its Global R&D Informatics efforts for the IT organization.

Mr. Apathy brings over 25 years of experience in Pharmaceutical new product and R&D capability development across the Pharmaceutical/Biotech, Life Sciences and Management Consulting industries. He has extensive knowledge and experience based upon leadership positions within both industry and consulting at Celgene, GlaxoSmithKline, Pfizer, Wyeth, Accenture, Eli Lilly & Company, PA Consulting Group, and Solutia Pharmaceutical Advisors.

He is also an experienced leader in strategy, processes, and systems for development and commercialization of new pharmaceutical products—skills and experience include delivery of strategy, large-scale program management, product development processes, life-cycle extension, product launch (particular emphasis on the Early Development phase of Pharmaceutical R&D, combining technical knowledge with change leadership, consulting problem-solving, strategy development, and change management processes).

BREAKOUT GROUPS

GROUP 1

Co-captains: Andreas Friese (Bayer) and Karsten Wicke (AbbVie)

Andrew Allen (Regeneron)
John Apathy (Celgene)
Monika Arenz (NIBR)
David Christie (Amgen)
Mary Donlan (PerkinElmer)
Robert Greenwood (SciBite)
James Hanly (BMS)
Birgitte Mathiesen (Novo-Nordisk)
Alain Nanzer (Roche)
Rob Saiter (Accenture)
Alex Schuleit (Lundbeck)

GROUP 2

Co-captains: Ashok Upadhyay (Otsuka) and Edith Scheringer (B-I)

Joel Ekstrom (Ionis)
Sabine Grote (AbbVie)
Rainer Kappes (Sanofi)
Derek Marren (Lilly)
Keith Murphy (Pfizer)
Christoph Schaub (Bayer)
Pascual Starink (Celgene)
Etzard Stolte (Roche)
Jason Swift (AstraZeneca)
Romain Taillard (Roche)
Nicole van Poppel (Accenture)

GROUP 3

Co-captains: Martin Leach (Alexion) and DJ Kleinbaum (Emerald Cloud Lab)

Tomas Bolek (Ataccama)
Dan Chapman (UCB)
Alexandra Grebe de Barron (Bayer)
Lars Greiffenberg (AbbVie)
Martin Hofmann-Apitius (Fraunhofer Society)
Jan Küntzer (Roche)
Tomoyuki Matsunaga (Takeda)
Anders Mortensen (Novo-Nordisk)
April Reeve (Celgene)
Michiel Ringkjøbing-Elema (Lundbeck)
Errol Sandler (PRISME Forum)
Nick Wright (AstraZeneca)

GROUP 4

Co-captains: Simon Roach (GSK) and Eduardo Gonzalez-Couto (PerkinElmer)

Thomas Abbott (Astellas)
Béatrice Chapuzet (Servier/Praxis)
Massimo de Francesco (UCB)
Devendra Deshmukh (Merck)
Ibrahim Emam (Imperial College London)
Klaus Hofenbitzer (Celgene)
Roy Ladd (AbbVie)
David Sedlock (Takeda)
Jean-Luc Schmidt (Sanofi)
Khushi Sharma (Ataccama)
Peter Smilansky (EDETEK)
John Wise (PRISME Forum)

SESSION VI
MEETING SUMMARY & AWARDS

Session Chair: Dan Chapman
Director, Discovery Research Information Management, *UCB*