Delegates Summit:

Best Practice and Definitions of Data-centric and Big Data

– Science, Society, Law, Industry, and Engineering
September 19, 2016

The Sixth Symposium on
Advanced Computation and Information in Natural and Applied Sciences
The International Conference on Numerical Analysis and Applied Mathematics (ICNAAM 2016)
September 19 – 25, 2016, Rhodes, Greece

Dr. rer. nat. Claus-Peter Rückemann[1,2,3]

[1] Westfälische Wilhelms-Universität Münster (WWU), Münster, Germany
[2] Leibniz Universität Hannover, Hannover, Germany
[3] North-German Supercomputing Alliance (HLRN), Germany
ruckema(at)uni-muenster.de

Delegates Summit: Best Practice & Definitions of Data-centric & Big Data

Delegates

- *Claus-Peter Rückemann* (Moderator),
  Westfälische Wilhelms-Universität Münster (WWU) /
  Leibniz Universität Hannover /
  North-German Supercomputing Alliance (HLRN), Germany

- *Zlatinka Kovacheva,*
  Middle East College, Department of Mathematics and Applied
  Sciences, Muscat, Oman

- *Lutz Schubert,*
  University of Ulm, Germany

- *Iryna Lishchuk,*
  Leibniz Universität Hannover, Institut für Rechtsinformatik, Germany

The International Conference on Numerical Analysis and Applied Mathematics (ICNAAM 2016),
The Sixth Symposium on Advanced Computation and Information in Natural and Applied Sciences,
CfP: https://research.cs.wisc.edu/dbworld/messages/2015-10/1446225912.html
Program: http://www.icnaam.org/sites/default/files/Preliminary_Program_of_ICNAAM_2016.pdf

Delegates' Summit: Best Practice and Definitions of Data-centric and Big Data

Recall: Last Years' Post-Summit Results   In 80 Words Around The World.

Recall: Last Years' Post-Summit Results In 80 Words Around The World

Knowledge and Computing (Delegates and other contributors)

- "Knowledge is created from a subjective combination of different attainments as there are intuition, experience, information, education, decision, power of persuasion and so on, which are selected, compared and balanced against each other, which are transformed, interpreted, and used in reasoning, also to infer further knowledge. Therefore, not all the knowledge can be explicitly formalised. Knowledge and content are multi- and inter-disciplinary long-term targets and values. In practice, powerful and secure information technology can support knowledge-based works and values."

- "Computing means methodologies, technological means, and devices applicable for universal automatic manipulation and processing of data and information. Computing is a practical tool and has well defined purposes and goals."

Claus-Peter Rückemann, Friedrich Hülsmann, Birgit Gersbeck-Schierholz, Knowledge in Motion / Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany; Przemysław Skurowski, Michał Staniszewski, Silesian University of Technology, Gliwice, Poland; International EULISP post-graduate participants, ISSC, European Legal Informatics Study Programme, Leibniz Universität Hannover, Germany

Best Practice and Definitions

Statements on Data-centric (1/2) (Delegates and other contributors)

- "Data centric approach considers the data as one connected whole - similar to the sea. Data is all encompassing, connected and fluid, touching everything. As anyone in the water gets wet, similarly, anyone deals with the data. The key to data-centric design is to separate data from behavior and reduce moving of data.

  Data centric application is one where the database plays a key role, where properties in the database may influence the code paths running in the application and where the most business logic is defined through database relations and constraints."

  Zlatinka Kovacheva, Middle East College, Department of Mathematics and Applied Sciences, Muscat, Oman

- "'data-centric': Applications that focus on the analysis of data, rather than on e.g. simulation of physical systems. They do, if you want, consume rather than produce data. The performance of data-centric applications is bound by the storage access speed (RAM and higher), not by the CPU."

  Lutz Schubert, University of Ulm, Germany.

Best Practice and Definitions

Statements on Data-centric (2/2) (Delegates and other contributors)

- "DataCentric is when the focus of data processing is on the data itself and the data is stored, passed and processed in such a way that all elements of a system understand and share value attributed to the data."

  Iryna Lishchuk, Leibniz Universität Hannover, Institut für Rechtsinformatik, Germany.

- "The term data-centric refers to a focus, in which data is most relevant in context with a purpose."

  Claus-Peter Rückemann, Friedrich Hülsmann, Birgit Gersbeck-Schierholz, Knowledge in Motion / Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany

Best Practice and Definitions   In 80 Words Around The World.

Statements on Big Data (1/2) (Delegates and other contributors)

- **"Big Data: A variety of structured and unstructured huge amount of data appearing with high speed from many sources of different types, needing dynamical analysis.**

  An example of Big Data might be petabytes (1,024 terabytes) or exabytes (1,024 petabytes) of data consisting of billions to trillions of records of millions of people. Big Data is particularly a problem in business analytics because standard tools and procedures are not designed to search and analyze massive datasets."

  Zlatinka Kovacheva, Middle East College, Department of Mathematics and Applied Sciences, Muscat, Oman

- **"'big data': ... is in a way a data-centric analysis with data sources from various sites, making the performance highly dependent on network speed. Though big data could simply mean vast amount of data to be processed, it is frequently used to describe the type of analysis performed on this data, namely data mining rather than, e.g., plain search.**

  The problem with data mining consists in the high dependencies between data set and therefore the constant switching between data sources, as well as the constant increase in data faster than processing can be performed."

  Lutz Schubert, University of Ulm, Germany.

Best Practice and Definitions   In 80 Words Around The World.

Statements on Big Data (2/2) (Delegates and other contributors)

- "BigData is a product of digital world when due to the availability of large amounts of data from different sources on the one hand and powerful computing powers on the other it became possible to process such data and derive new knowledge and economic benefit out of it. The essential features of bigdata are: Volume, variety, velocity, veracity."

  Iryna Lishchuk, Leibniz Universität Hannover, Institut für Rechtsinformatik, Germany.

- "The term Big Data is refering to data, which is larger and/or more complex than conventionally handled with storage and computing installations. Data use with associated application scenarios can be categorised by volume, velocity, variability, vitality, veracity, ... associated with the data."

  Claus-Peter Rückemann, Friedrich Hülsmann, Birgit Gersbeck-Schierholz, Knowledge in Motion / Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany.

Conclusions, Discussion, Networking

Data Centricity and Big Data

- The definitions of Big Data are commonly longer than for data-centricity.
- The content / knowledge do have the highest values.
- If data is in the focus then knowledge and value of data can benefit from data-centric models.
- Big Data can be data-centric with the solid situational understanding of data centricity.
- Different data categories, e.g., scientific data, data with capacity computing, social network data, business and industry data, can afford different implementations, engineering, and infrastruture architectures.
- Data-centric models can help to cope with long-term data challenges and Big Data.
- Data centricity is more than data management tools, dynamic table-driven logics, stored procedures, and shared databases and communication in parallel.
- Important aspects are to separate technical implementations from content creation, to create a currency for data value, and to foster knowledge creation.
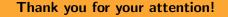
Post-Summit Results In 80 Words Around The World.

Data-centric and Big Data (Delegates and other contributors)

- "The term data-centric refers to a focus, in which data is most relevant in context with a purpose. Data structuring, data shaping, and long-term aspects are important concerns. Data-centricity concentrates on data-based content and is benefitial for information and knowledge and for emphasizing their value. Technical implementations need to consider distributed data, non-distributed data, and data locality and enable advanced data handling and analysis. Implementations should support separating data from technical implementations as far as possible."

- "The term Big Data refers to data of size and/or complexity at the upper limit of what is currently feasible to be handled with storage and computing installations. Big Data can be structured and unstructured. Data use with associated application scenarios can be categorised by volume, velocity, variability, vitality, veracity, value, etc. Driving forces in context with Big Data are advanced data analysis and insight. Disciplines have to define their 'currency' when advancing from Big Data to Value Data."

Citation: Rückemann, C.-P., Kovacheva, Z., Schubert, L., Lishchuk, I., Gersbeck-Schierholz, B., and Hülsmann, F. (2016): Post-Summit Results, Delegates' Summit: Best Practice and Definitions of Data-centric and Big Data – Science, Society, Law, Industry, and Engineering; Sep. 19, 2016, The Sixth Symposium on Advanced Computation and Information in Natural and Applied Sciences (SACINAS), The 14th Internat. Conf. of Numerical Analysis and Applied Mathematics (ICNAAM), Sep. 19–25, 2016, Rhodes, Greece.
Delegates and contributors: Claus-Peter Rückemann, Knowledge in Motion / Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany;Zlatinka Kovacheva, Middle East College, Department of Mathematics and Applied Sciences, Muscat, Oman; Lutz Schubert, University of Ulm, Germany; Iryna Lishchuk, Leibniz Universität Hannover, Institut für Rechtsinformatik, Germany; Birgit Gersbeck-Schierholz, Friedrich Hülsmann, Knowledge in Motion / Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany

Networking and Outlook

Thank you for your attention!

**Wish you an inspiring conference
and a pleasant stay on Rhodos!**

**Looking forward to seeing you again next year for the
Symposium on Advanced Computation and Information!**