CRI 19 Presents

BUSINESS INTELLIGENCE & SEMANTIC WEB

COLLOQUIUM

2-DECEMBER 2019
Department of Computer Science, Faculty of Science, Yaoundé I
CAMEROON

MONDECA

Ghislain ATEMEZING, PhD
Directeur R&D, Mondeca
Paris, France

35 boulevard Strasbourg
75010 Paris, France
+33 1 4111 3034

Ghislain.Atemezing@mondeca.com
www.mondeca.com
Why This Colloquium?

Search for “Web Semantique” - Google Trends (21/11/2019)
Why This Colloquium?

Search for “RDF” - Google Trends (21/11/2019)
Why This Colloquium?

Search for “Business Intelligence” - Google Trends (21/11/2019)
Why This Colloquium?

Complete trend, starting with January 2013

Graph Databases are trending, and there is an ongoing if subtle war for domination going on. (Image: DB engines)
Data (R)evolution!

- Industrial Revolution (1780 - 1840) - First revolution
- Technical Revolution (1870 - 1920) - 2nd revolution
- Scientific Technical Revolution (1940 - 1970)
- Digital Revolution (1975 - 2011) - 3rd revolution
- (Big) Data Revolution (2013 - today) - 4th revolution
Internet of Things (IoT) refers to billions of “things” connected to the Internet.

2 billion in 2006
Expected: 200 billion objects connected by 2020
L’agriculture numérique

Les capteurs au cœur de l’agriculture numérique

- Second marché de la robotique en France

Source: Slides Frederique Laforest, LIRIS
L’Industrie 4.0

- Le jumeau numérique
  - Maintenance prédictive
  - Conception / évolutivité

© Christoph Roser at AllAboutLean.com

© cetim
La ville intelligente

Business des objets connectés (IoT) :
- Marché estimé à 800B$
- Croissance mondiale accélérée
- Gros potentiel pour les villes, la santé, l’énergie, le commerce, l’environnement...

Bénéfices potentiels
- Offrir de nouveaux services
- Améliorer la qualité des services
- Réduire les coûts
- Renforcer la sécurité

Source: Slides Frederique Laforest, LIRIS
1. Data is a key asset of any organization.

2. The current enterprise software paradigm is “Application-Centric.”

3. Most of the excessive cost and complexity in Enterprise Apps stems from the relationship of the apps to the data.

4. We are committed to reversing this trend.

5. There is even more money to be saved in the data-centric paradigm.

Link: http://datacentricmanifesto.org/principles/
CRI 19 Presents

BUSINESS INTELLIGENCE & SEMANTIC WEB

COLLOQUIUM

2-DECEMBER 2019
Department of Computer Science, Faculty of Science, Yaoundé I
CAMEROON

SPEAKERS

Prof. Dr. Axel-Cyrille Ngonga Ngomo
Dr. Maria Keet, PhD
Dr. Michel Kana, PhD
Dr. Ghislain Atemezing
Dr. Gaoussou Cemara
M. Elie Noël Malouang
Dr. Gayo Diallo
Dr. Paulin Melatagia
M. Dieudonné Djofang
M. Louis Ekani
M. Pierre Lotis Nankep

Web: https://sites.google.com/view/bismwebcri19/
Follow us: @BiSemWeb (#bismweb19) | https://web.facebook.com/BiSemWebColloquium/

Sponsored by AIJ (Artificial Intelligence Journal)
DIK(W) vs DIK(I) model(s)

- Data
  - Raw: Red, 192.234.235.245.678, v2.0
- Information
  - Meaning: South facing traffic light on corner of Pitt and George Streets has turned red
- Knowledge
  - Context: The traffic light I am driving towards has turned red
- Wisdom
  - Applied: I better stop the car!
DIK(W) vs DIK(I) model(s)

- Data = discrete (ex: 37)
- Information = links on data / data with context or data with context
  - body temperature)
- Knowledge = rules (eg: if temp_body(?x) > 37C → Has_fever(?x) or information + meaning
- Wisdom = why? and what for? - insight from knowledge
  - e.g: I go to my doctor
- Intelligence = smart use of wisdom in contextualized environment
DIKW model to solve real world problems
BI Process and Architecture

- **Data acquisition layer**: Consists of components to get data from all the source systems, such as human resources, finance, and billing.
- **Data integration layer**: Consists of integration components for the data flow from the sources to the data repository layer in the architecture.
- **Data repository layer**: Stores data in a relational model to improve query performance and extensibility.
- **Analytics layer**: Stores data in cube format to make it easier for users to perform what-if analysis.
- **Presentation layer**: Applications or portals that give access to different set of users.
BI products - IBM

By Sandip Chowdhury, IBM - 2014
Publish Five-star data on the Web

Source: https://5stardata.info/en/ / By Sir Tim Berners-Lee
Semantic Tools by Google - Schema.org

- Web Developers: [https://search.google.com/structured-data/testing-tool/u/0/](https://search.google.com/structured-data/testing-tool/u/0/)
- Google Dataset search: [https://toolbox.google.com/datasetsearch](https://toolbox.google.com/datasetsearch)

[Diagram of dataset search process]

Some open research challenges in SemWeb

- Entity resolution
- Very large scale ontology / instance data visualization
- Declarative UI generation from Shapes
- Low / no code semantic model driven development
- IoT integration with semantics
- Hybrid AI (combining ML + Cognitive AI)
- Performant real-time reasoning
- Effective graph partitioning
- ....
Housekeeping

- Main page: https://sites.google.com/view/bisemwebcri19/
- Link session: http://linkedvocabs.org/biswa19/session.html
- 2 coffee breaks + lunch
- 2 remote presentations
- Collaborative report: https://docs.google.com/document/d/17s1nrEnQ9DfmFZ1vagTOqI3vSGUXPii50T2yXaaxoIs/edit
- Ask questions as much as you can (interaction): https://app.sli.do/event/benwskp5 (sli.do, code: #BISW19)
- Tweet, retweet about the event: #bisemweb19
- We will share all the slides in the web page
Links

1- Fabien Gandon: “A Survey of the First 20 Years of Research on Semantic Web and Linked Data”: https://hal.inria.fr/hal-01935898/

2- Fabien Gandon: "For everything: Tim Berners-Lee, winner of the 2016 Turing award for having invented… the Web" :https://hal.inria.fr/hal-01843967

3- Pizza tutorial with Protégé: http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf
From Application Thinking to Knowledge Revolution

Application Centric Thinking

Relational DB
- Problems
  - SQL interoperability
  - Costly
  - Scaling hits a wall
  - Data silos
  - Limited semantics
  - Data complexity causes data proliferation & code proliferation

Enterprise Resource Planning
- Problems
  - Ever increasing complexity and feature creep
  - Costly
  - Massive underlying databases (>10,000 tables)
  - Implementation failures
  - Integration complexity

Enterprise Data Modeling
- Problems
  - Skills mismatch
  - Long implementation timeframes
  - "Ivory tower" syndrome
  - Round trip modelling - Conceptual/Logical/Physical
  - Massively complex models

SOA
- Problems
  - Process centric
  - Complex technology stacks
  - Shared message culture
  - ESB’s were very complex
  - Still lots of point to point integration
  - Information was missing!!!

Data Warehousing
- Problems
  - ETL slow
  - Data completedness
  - Resource intensive
  - Complexity
  - Schema agreement

Source: Keynote Dougal Watt / ISWC 2019