





Big Data White Paper Release Breakfast

May 25, 2016

ILNAS / ANEC



PROGRAM

- Introduction and welcome words
- Presentation of the National Standards Body
- Presentation of the White Paper on Big Data
- Big Data standardization and participation in standardization
- Discussions







Introduction

Dr. Jean-Philippe HUMBERT - ILNAS







Presentation of the National Standards Body

Mr. Jérôme HOEROLD - ILNAS





ILNAS, Institut Luxembourgeois de la Normalisation, de l'Accréditation, de la Sécurité et qualité des produits et services

- Creation: Law dated July 14, 2014 (repealing the amended Law of May 20, 2008)
- <u>Status</u>: Public administration under the authority of the Minister of the Economy
- Total staff: 36 civil servants (May 2016)







Luxembourg's Standardization Strategy 2014-2020

PILLAR 1 Information and communication technologies (ICT)

- Support and constant development of the standardization field dedicated to ICT
- Implementation of the Luxembourg's Policy on ICT standardization (2015-2020)
 - Developing the interest and the involvement of the market
 - Promoting and reinforcing the participation of the market
 - Supporting and strengthening the education about standardization and related research activities
- Detection of niche opportunities for economic developments

PILLAR 2 National influence and compliance with legal attributions

PILLAR 3 Products and services





ANEC, Agence pour la Normalisation et l'Économie de la Connaissance

Agency for Standardization and knowledge-based Economy

- Creation: October 4, 2010
- Status: Economic Interest Grouping (EIG)



- Object:
 - Promotion, awareness raising and training, applied research in the field of standardization and metrology in order to support companies' competitiveness in Luxembourg
- Total staff: 15 employees (May 2016)
- Partners:









Position





OLN development plan for 2016

Communication

- Lecture stations
- First national standard "living space"

Research & development and Education

- STAIR
- Awareness raising in high schools

Sectorial approach: Construction domain

- Awareness raising among local actors
- Organization of seminaries/conference



ILNAS Standardization activities in Luxembourg

Create a normative culture in Luxembourg

- University Certificate "Smart ICT for Business Innovation" at the University of Luxembourg
- Bachelor Courses in Informatics and Engineering at the University of Luxembourg
- Promotion in the field of standardization (Newsletter, <u>portail-qualite.lu</u>, LinkedIn, events, ...)
- Trainings and research in the field of standardization

Creation of national standards

- National Annexes of the Eurocodes
- National Annex concerning the Winter Diesel
- National standard about the living surface
- Creation of a national standards office in the field of construction



Availability of standards *Standardization catalogue*

- 59 national standards
- 47.000 European standards from CEN and CENELEC
- 55.000 international standards from ISO and IEC
- 4.100 ETSI standards (free)



New

- 44.000 DIN standards
- More than 150.000 normative documents at your disposal













Availability of standards ILNAS e-shop

- Format: electronic
- Language: French, German and English
- Competitive prices
- Free access to documents in public enquiry









Availability of standards *Free access on lecture stations*

 Availability of all EN (CEN,CENELEC et ETSI), ISO, IEC and ILNAS standards (despite DIN)

Location of the reading stations:

- 1. Université du Luxembourg
 - Campus Kirchberg
- 2. Chambre de Commerce
 - Kirchberg (Espace Entreprises)
- 3. Bibliothèque nationale de Luxembourg
 - Luxembourg centre-ville
- 4. ILNAS
 - Esch-Belval
- 5. LIST
 - Esch-Belval (Maison de l'innovation)
 - Belvaux







Participation in standardization *Different possibilities*

- How to participate in the development of national, European and international standards ?
 - 1. Comment of draft standards in public enquiry
 - 2. Active participation in a technical committee





Participation in standardization *Public enquiry*

 Navigate to l'ILNAS e-shop in order to comment a draft standard which is in the stage of public enquiry











Participation in standardization

- 2. National delegate in standardization
- Who can participate ?
 - Every socio-economic actor with a certain expertise

Cost of participation ?

- Free participation in Luxembourg
- National experts register (May 2016)
 - 223 persons registered
 - 592 registrations in technical committees

| | Nombre d'inscriptions aux comités tech | iques : |
|---------------------------|---|---|
| | ILNAS/OLN | 24 |
| | CEN | 177 |
| | CENELEC | 15 |
| | CEN/CENELEC | 3 |
| | CEN/CENELEC/ETSI | 2 |
| | ECISS | 21 |
| | ISO/IEC | 129 |
| | ISO | 213 |
| | IEC | 8 |
| | Total | 592 |
| | Nombre de personnes inscrites : 223 | |
| | | |
| | ILNA | S |
| | elvaux - Tél. ; (+352) 24 77 43 40 - Fax ; (+352) 24 79 43 40 | - Email : normalisation@ilnas.etat.lu - www.portail-quali |
| 1, av du Swing - L-4367 B | | |



Products and services

- ILNAS, in collaboration with G.I.E. ANEC, offers the following products and services to the national market :
 - Diffusion of normative information
 - Diagnostic on standardization
 - Training and awareness sessions
 - Standards watch
 - Standards analysis (ICT)
- These products and services are provided for free on simple demand









Stay informed about ILNAS activities

Portail qualité: www.portail-qualite.lu



ILNAS e-shop: ilnas.services-publics.lu









White Paper "Big Data"

Dr. Joseph EMERAS - ANEC



Today: a massive increase of Data

- Social network sources;
- Organizations monitoring user activities;
- **IoT**: billions of devices generate real-time **streams** of data;
- Scientific applications and experiments result in an increase of datasets at an exponential rate;
- Storage capacity is **cheap** buying storage easier than deciding what to delete;
- Data analytics techniques have significantly **improved** enabling higher degree of understanding from data.



Source: science-all.com

Big Data Purpose Examples

- Enhanced customer view Extended *customer views* by combining
- internal and external information sources;
- Exploration

Find, visualize and understand to improve *decision making;*

- Security/Intelligence extension Lower risk, detect fraud and monitor cyber security in real-time;
- Operations analysis

Analyze a variety of machine data for improved business results;

Augmentation

Integrate big data and traditional data for *new services*, e.g. augmented reality.

Big Data



From Business

Per second:

- 26,336 GB of Internet traffic
- 47,965 Google searches
- 97,562 YouTube videos viewed
- 8,780 Tweets sent
- 1,860 Instagram photos uploaded
- 1,692 Skype calls

Per minute:

• 500 hours of video uploaded to YouTube (Nov. 2015).

Per day:

- Facebook creates 10 terabytes (10x10^12 bytes) of data;
- Google produces 24 terabytes of data from its search operations.

From Science

- CERN's Large Hadron Collider (particle accelerator) generates
 40 terabytes (4x10^13 bytes) per second. Its Data Centre processes about one petabyte (10^15 bytes) of data every day.
- 32 petabytes of climate observations and simulations on the supercomputing cluster in the NASA Center for Climate Simulation.
- The Large Synoptic Survey Telescope (LSST) will record 30 exabytes (3x10^19 bytes) of image data in a single day.



Big Data



90% of data in the world has been generated in the last 2 years.

Today, more data are generated in 10 minutes than all of humanity has ever created through to the year 2003.

> Cloud, Internet of Things. Business Intelligence, Forecasting, Machine Learning...

Let's try to formalize Big Data...

Big Data refers to technologies that involve data that are too **massive**, **diverse** and **fast-changing** to be processed efficiently with conventional techniques.

White Paper





Big Data Characteristics



Characteristics

3, 4, 5... Vs

- Volume How much data;
- Velocity How fast they are produced;
- Variety Various types of data, sources, formats;
- Veracity How accurate are data;
- Variability Changes in data rate, format/structure, semantics, and/or quality.

- Value Analytical possibilities;
- Visualization Presentation of results for decision making purposes.



Big Data Data Types



• *Structured data* - is part of a formal structure of data models associated with e.g. relational databases.

ILNAS

- It can be generated both by computer software or humans.
- *Semi-structured data* not part of a formal structure of data models.
 - Contains markers to separate semantic elements and enforce hierarchies of records and fields (example: XML).
- Unstructured data does not belong to a pre-defined data model.
 - Includes data from e-mails, video, social media websites and text streams. Accounts for more than 80% of all data in organizations.



Difference in format makes Big Data more complex to process.

White Paper





Operational: Non-Relational Databases

Analytical: MapReduce Paradigm

Challenges

ILN4S



MapReduce Paradigm

- Programming Model
- Benefits from **parallelization** of work
- Several implementations Google, Apache, CouchDB...
- Plenty of frameworks for different languages: R (e.g. Rhadoop), Python, Ruby...



The overall MapReduce word count process

Source: R-bloggers

White Paper





Market Readiness and Impact

- Growth doubles between 2014-2017 [1]
- ROI impact up to 20% [2]

Challenges, Implementation

- Roadmap Definition
- Good Practices



[2] Perrey, J., Spillecke, D., Umblijs, A. (2013). Smart analytics: How marketing drives short-term and long-term growth. McKinsey Quarterly.

Big Data Business Prospective



Implementation Strategy

Business impact – assess time and effort required to design Big Data solutions **Projected capacity** – evaluate how much data are required and how fast it needs to be analyzed.

Preferred software development method – agile development process, iterative methodologies. Short time cycles with rapid results and continuous user involvement to incrementally deliver a proper business solution.

Available budgets and skill sets -

understand the expected investments and required knowledge for the Big Data implementation and ensure appropriate budget and sponsorship.

Risk appetite - evaluate the risk depending on the scope and expected benefits.

Challenges



Source: Accenture Survey 2014

Good Practices

Discover available data, what data is missing, and the impact on analysis. Security plan, evaluate what data needs encryption and which does not. Plan data governance strategy, accountability.

Adjusting the impact, adapt business strategy according analytics findings.

White Paper





White Paper





Many **challenges** including analysis, capture, querying, updating, search, sharing, storage, transfer, visualization, data curation, information privacy...

Embark your organization on the Big Data journey with ILNAS White Paper: <u>http://www.portail-qualite.public.lu/fr/publications/normes-</u>normalisation/information-sensibilisation/white-paper-big-data/WP_BigData_v1.pdf







Big Data standardization and participation in standardization

Dr. Sune NIELSEN - ANEC GIE Mr. Nicolas DOMENJOUD - ANEC GIE





Standardization organizations

| | National Level | European International Level Level |
|--------------------------------------|-------------------|---------------------------------------|
| Standardization in general | ILNAS | Vienna Agreements |
| Electotechnical standardization | ILNAS | CENELEC Dresden Agreements |
| Telecommunication standardization | ILNAS | ETSI World Class Standards |
| * ITU-T | | Fora & Consortia |





ISO/IEC JTC 1 structure and national level involvement





ISO/IEC JTC 1 Big Data Report

- Scope:
 - Preliminary Report 2014
 - Survey the existing ICT landscape for key technologies and relevant standards/models/studies/ use cases and scenarios for Big Data from JTC 1, ISO, IEC and other standard setting organizations;
 - Identify key terms and definitions commonly used in the area of Big Data; and
 - Assess the current status of Big Data standardization market requirements, identify standards gaps, and propose standardization priorities to serve as a basis for future JTC 1 work.
- 16 potential standardization gaps:
 - Use cases; metadata and data provenance; Application models;
 - Query languages; Advanced network protocols;
 - Security and privacy access controls;
 - Remote, Distributed, and federated analytics;
 - Data sharing and exchange; Visualization;
 - Energy measurement; Quality and Veracity description and management







ISO/IEC JTC 1/WG 9 – Big Data

• <u>Created:</u> April 2015

Main focus areas:

- Serve as the **focus of and proponent** for JTC 1's Big Data standardization program
- Develop **foundational standards** for Big Data for guiding Big Data efforts throughout JTC 1
- Develop other Big Data standards that build on the foundational standards
- Identify gaps in Big Data standardization
- Identify, develop and maintain liaisons with all relevant entities and investigate ongoing work
- Engage with the **community** outside of JTC 1 to grow the **awareness** of and encourage engagement

Members:

130+ from 21 countries: United States, Australia, Austria, Brazil, Canada, China, Finland, France, Germany, India, Ireland, Italy, Japan, Republic of Korea, Luxembourg, Netherlands, Norway, Russian Federation, Singapore, Spain, Sweden, United Kingdom

Luxembourg's involvement:

- Mr. Johnatan PECERO, Mr. Joseph EMERAS, Mr. Sune NIELSEN (ANEC GIE)
- Mrs. Aida HORANIET (Docler Holding)
- Mr. Emmanuel KIEFFER (University of Luxembourg)
- Projects under development: 6 International Standards and Technical Reports
- **<u>Chairperson</u>**: Mr. Wo Chang (United States)





ISO/IEC 20546 - Information Technology - Big Data - Definition and Vocabulary



- This International Standard provides an overview of Big Data along with a set of terms and definitions
- It provides a terminological foundation for Big Data-related standards

Example:

- Definition of Big Data characteristics
- Horizontal scaling / vertical scaling
- And more
- Expected date of publication:
 - March 2017





ISO/IEC 20547 - Information technology - Big Data - Reference Architecture



Scope:

The ISO/IEC 20547 series of standards and technical reports will provide a framework and **reference architecture** which organizations can apply to their problem domain to effectively and consistently describe their architecture and its **implementations** with respect to the **roles/actors** and their concerns as well as the **underlying technology**

Structure:

- Part 1: Framework and Application Process
- Part 2: Use Cases and Derived Requirements
- Part 3: Reference Architecture
- Part 4: Security and Privacy Fabric has been assigned to JTC 1/SC 27 "IT security techniques"
- Part 5: Standards Roadmap
- Expected dates for the publication: December 2017



Selected internal projects in liaison



- ISO/IEC JTC 1/SC 38 Cloud Computing and Distributed Platforms
 - Structure and scope:
 - WG 2 Service Oriented Architecture (SOA)
 - WG 3 Cloud Computing Service Level Agreements (CCSLA)
 - WG 4 Cloud Computing Interoperability and Portability (CCIP)
 - WG 5 Data and their Flow across Devices and Cloud Services (CCDF)
 - ISO/IEC 19941 Information Technology -- Cloud Computing -- Interoperability and Portability
 - ISO/IEC 17826 Information technology -- Cloud Data Management Interface (CDMI)
 - ISO/IEC 17789 Information technology -- Cloud computing -- Reference architecture
 - Many more...

ILN4S



Selected internal projects in liaison



- ISO/IEC JTC 1/SC 32 Data Management and Interchange
 - Standards for data management within and among local and distributed information systems environments.
 - Enabling technologies to promote harmonization of data management facilities across sector-specific areas.
 - ISO/IEC 9075 Information technology -- Database languages SQL
 - Part 9: Management of External Data (SQL/MED)
 - Part 15: Multi dimensional arrays
 - ISO/IEC 11179 Information technology -- Metadata registries (MDR)
 - ISO/IEC 15944 Information technology -- Business Operational View
 - Part 7: eBusiness vocabulary
 - Part 12: Privacy protection requirements on information life cycle management (ILCM) and EDI of personal information



Selected internal projects in liaison



- ISO/IEC JTC 1/SC 29 Coding of audio, picture, multimedia and hypermedia information
 - Standards for coded representation of audio, picture, multimedia and hypermedia information - and sets of compression and control functions for use with such information
 - "Big Media" activity in SC 29/WG 11 MPEG standards related to Big Data
 - Real-time image processing correlated with sensor data
 - For parking occupancy, traffic, pollution, augmented reality, 3D reconstruction etc.
 - Video data used by 34% of survey participants¹
 - ISO/IEC 23000 Information technology Multimedia application format (MPEG-A)
 - Part 13: Augmented reality application format
 - ISO/IEC 23001 Information technology -- MPEG systems technologies
 - Part 9: Common encryption of MPEG-2 transport streams
 - Many more...



Other relevant organizations

External organization in liaison with JTC 1 / WG 9

- ISO/TC 69 Applications of statistical methods
- ISO/TC 204 Intelligent transport systems
- ITU-T SG13 Future networks including cloud computing, mobile and next-gen networks
- Industrial Internet Consortium (IIC)

Additional relevant organizations

- ISO/TC 184/SC 4/WG 13 Industrial Data Quality
- Research Data Alliance's Big Data Interest Group
- NIST Big Data Public Working Group (NBD-WG)
- Cloud Security Alliance (CSA)
 - Big Data Working Group (BDWG)
- Cloud Standards Customer Council (CSCC)
 - Big Data in the Cloud Working Group





ILNAS CANEC



Why we need Standards for Big Data?

- More clarity for government and industry
- Solve the Vendor Locking problem
 - Need for a Reference Architecture, Open APIs standards
- Ensure commonality between different industry sectors in terms of Data Flow
- Ensure Data Quality / Data Governance to address veracity and value of data
- Semantic is a key to extract Value from data
 - Need Models and Ontologies to enable "Smarter" reasoning

Be part of standardization to participate in the emergence and consolidation of Big Data!





Standardization stakes for Organizations



- Defend the interests of your business
- Spread and promote your innovations

- Learn about your competitors
- Defend common interests
- Promote your business





Organization of participation in standardization



- **TC**: Technical Committee
- **SC**: Sub-committee
- **WG**: Working Group

- **NSB**: National Standardization Body
- NMC: National Mirror Committee

ILN4S

Rights of delegates

Rights

- Access to documents
- Work on standards under development
- Vote during the validation or approval process
- Participate in international meetings
- Give feedbacks to ILNAS, if necessary, on malfunctions

Duties

- Respect of the standardization policy
- Nondisclosure of technical committee's documents to third parties
- Active participation in the standardization process
- Providing a periodic review to ILNAS (personal activities, comments, etc.)







Registration process to a standardization technical committee

- Application form for registration to a standardization technical committee
 - Form ILNAS/OLN/F001
 - Available at <u>http://www.portail-qualite.public.lu/fr/normes-</u> normalisation/developpement-normes/devenir-delegue-national/index.html

Processing of the application by ILNAS

- Evaluation of the application: competencies, relevance, interest for the national economy
- Approval of the registration application
- Entry to the national register of standardization delegates
 - Notification to the candidate
 - Access to the collaborative platform



New services under development - Coaching for national standardization delegates

First step (available now)



- Personalized support for the handling of collaborative work platforms and voting system
 - On demand for the national standardization delegates of the ICT sector
 - Complement the Training session "New delegate in standardization"
- Second step (development during 2016)
 - New tools & services based on the needs and barriers identified in step 1
- Objectives
 - Set up good practices common to all national delegates of the ICT sector
 - Facilitate the standardization work of national delegates
 - Understanding of the standardization environment
 - Organization of the national mirror committees
 - Encourage a stronger involvement of the national standardization community

Contact: anec@ilnas.etat.lu







Discussions







ISO/IEC CD 20546, Information Technology -- Big Data -- Definition and Vocabulary

ISO/IEC AWI TR 20547-1, Information technology -- BDRA -- Part 1: Framework and application process

ISO/IEC AWI TR 20547-2, Information technology -- BDRA -- Part 2: Use cases and derived requirements

ISO/IEC AWI 20547-3, Information technology -- BDRA -- Part 3: Reference architecture

ISO/IEC AWI 20547-4, Information technology -- BDRA -- Part 4: Security and privacy fabric

ISO/IEC AWI 20547-5, Information technology -- BDRA -- Part 5: Standards Roadmap

| | 0 month | | 12 m. | 24 m. | 30 m. | 36 m. |
|----|---------|----|-------|-------|-------|-------|
| NP | | WD | CD | DIS | FDIS | IS |
| | | | PDTR | TR | | |





CONTACT



Institut luxembourgeois de la normalisation, de l'accréditation, de la sécurité et qualité des produits et services Institut luxembourgeois de la normalisation, de l'accréditation, de la sécurité et qualité des produits et services - Organisme luxembourgeois de normalisation

Tél. : (+352) 247 743 – 40 Fax : (+352) 247 943 – 40 E-mail : <u>normalisation@ilnas.etat.lu</u>



Agence pour la Normalisation et l'Économie de la Connaissance GIE

Tél. : (+352) 247 743 – 70 Fax : (+352) 247 943 – 70 E-mail : <u>anec@ilnas.etat.lu</u>

Follow us on Linked in

LinkedIn Group: "ICT Standardization Luxembourg"