ISO / IEC AI Workshop Series

Inaugural Workshop

Wael William Diab
ISO/IEC JTC 1/SC 42 Chair
ISO/IEC AI Workshop Chair
Welcome
Thank you for participating!

Participation Guidelines

Please use the Q&A function for questions

Please *do not use the chat function* for questions

This webinar is being recorded and will made available

Please keep your video off and audio muted unless speaking
Acknowledgments

Program Committee
- Heather Benko
- Norbert Bensalem
- Andrew Dryden
- Peter Deussen
- Rohit Israni
- Mike Mullane
- Catherine Nelson
- Wael William Diab

ISO
- Elisabeth Gasiorowski
- Catherine Infante
- Vivienne Rojas

IEC
- Stephen Dutnall
- Gabriella Ehrlich
- Ian Gardner
- Giulia Pizzi
- Laura Rocha
Program

SESSION 1
(24th May 13:00 – 16:00 UTC)
- Opening Remarks
- AI Applications
- Novel AI Standardization

SESSION 2
(25th May 05:00 – 08:00 UTC)
- Opening Remarks
- Emerging AI Requirements
- Emerging AI Technology Trends

SESSION 3
(25th May 22:00 – 26th 01:00 UTC)
- AI Applications
- Novel AI Standardization
- Closing Remarks
# Session 1

<table>
<thead>
<tr>
<th>Program</th>
<th>Talk Title</th>
<th>Track Chairs</th>
<th>Aprox Start Time (UTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kickoff</strong></td>
<td></td>
<td>Wael William Diab</td>
<td>13:00</td>
</tr>
<tr>
<td><strong>Opening Remarks</strong></td>
<td></td>
<td>Gilles Thonet</td>
<td>13:05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phil Wennblom</td>
<td></td>
</tr>
<tr>
<td><strong>AI Applications</strong></td>
<td></td>
<td>Rohit Israni</td>
<td>13:15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catherine Nelson</td>
<td></td>
</tr>
<tr>
<td>Mike Glickman</td>
<td>Health informatics and AI, the road to Interoperability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neil Frost</td>
<td>AI potential Applications in Transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christophe Preube</td>
<td>AI finds awareness in standardization work – the White paper of ISO/TMB SMCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Novel AI Standardization Approaches</strong></td>
<td></td>
<td>Peter Deussen</td>
<td>14:25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Norbert Bensalem</td>
<td></td>
</tr>
<tr>
<td>Kimberly Lucy</td>
<td>Creating Trust in AI Through Standards: A Management System Approach</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wo Chang</td>
<td>Data Quality for Analytics and Machine Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tim McGarr / Florian Ostmann</td>
<td>AI Standards Hub (UK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viveka Bonde</td>
<td>Novel Standardization Approach to AI Ethics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Session 2

<table>
<thead>
<tr>
<th>Program</th>
<th>Talk Title</th>
<th>Track Chairs</th>
<th>Aprox Start Time (UTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kickoff</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wael William Diab</td>
<td></td>
<td></td>
<td>5:00</td>
</tr>
<tr>
<td><strong>Opening Remarks</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silvio Dulinsky</td>
<td></td>
<td></td>
<td>5:05</td>
</tr>
<tr>
<td><strong>Emerging AI Requirements</strong></td>
<td></td>
<td>Catherine Nelson and Peter Deussen</td>
<td>5:10</td>
</tr>
<tr>
<td>Elham Tabassi</td>
<td>NIST AI Risk Management Framework</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mariagrazia Squicciarini</td>
<td>The UNESCO Recommendation on the Ethics of AI - Setting the standards for a better and more inclusive future</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liz Coll</td>
<td>Can standards deliver consumer trust and confidence in AI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daniel Loevenich</td>
<td>Evaluation Standards for Conformity Assessment of Trustworthy Cloud-based AI Applications</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Emerging AI Technology Trends</strong></td>
<td></td>
<td>Norbert Bensalem and Rohit Israni</td>
<td>6:30</td>
</tr>
<tr>
<td>Shubhashis Sengupta</td>
<td>Emerging AI Trends – as seen by an Industry Practitioner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tilak Kasturi</td>
<td>How specialized AI drives value for Automotive Service Organizations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>William Uppington</td>
<td>AI Quality: The Next Big Challenge in AI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babak Hodjat</td>
<td>From Data to Decisions, and Back</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Session 3

<table>
<thead>
<tr>
<th>Program</th>
<th>Talk Title</th>
<th>Track Chairs</th>
<th>Aprox Start Time (UTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kickoff</strong></td>
<td>Wael William Diab</td>
<td></td>
<td>22:00</td>
</tr>
<tr>
<td><strong>AI Applications</strong></td>
<td>Charalambos Freed</td>
<td>Safety considerations in autonomous products</td>
<td>Rohit Israni and Catherine Nelson</td>
</tr>
<tr>
<td></td>
<td>Hajime Yamada</td>
<td>Development of SRD 63416 “Ethical Considerations of AI when Applied in the AAL Context”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kenzo Nonami</td>
<td>AI Powered UAV and Future Prospects</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fumihiro Maruyama</td>
<td>Use cases and AI application guidelines in international standardization</td>
<td></td>
</tr>
<tr>
<td><strong>Novel AI Standardization Approaches</strong></td>
<td>Jochen Friedrich</td>
<td>AI Standardisation supporting regulatory needs</td>
<td>Norbert Bensalem and Peter Deussen</td>
</tr>
<tr>
<td></td>
<td>Linzhing Meng / Mike Thieme</td>
<td>Developing a TS on assessment of machine learning classification performance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yonosuke Harada</td>
<td>Introduction of Governance Implications of AI Systems</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paul Cotton / Milan Patel / Wei Wei</td>
<td>The foundational standards for AI – ISO/IEC 22989 and ISO/IEC 23053</td>
<td></td>
</tr>
<tr>
<td><strong>Closing Remarks</strong></td>
<td>Wael William Diab</td>
<td>Overview and introduction of SC 42</td>
<td>0:55</td>
</tr>
<tr>
<td></td>
<td>Program Committee</td>
<td>Insights from the workshop</td>
<td></td>
</tr>
</tbody>
</table>
Thank you

Wael William Diab

wael.diab@gmail.com; Wael’s LinkedIn

Workshop Website

ISO/IEC AI LinkedIn Page
Introducing
ISO/IEC
JTC 1

Phil Wennblom
ISO/IEC JTC 1 Chair
Information Technology

• Utilized by every sector of the economy, enabling gains in productivity and increases in quality

• Touches nearly every person in the world, enabling a higher quality of life

• Characterized by innovation and a rapid rate of change
About ISO/IEC JTC 1

• Joint Technical Committee of IEC and ISO
  • “Standardization in the field of Information Technology”
  • Consistent IT deliverables across IEC and ISO
  • Building blocks for global markets that respond to the needs of businesses, consumers, governments, other organizations

• 35 P-members and 65 O-members
• Organized in 23 Subcommittees and 4 JTC 1 Working Groups
• More than 4500 active participants currently developing around 600 standards; more than 3200 published
JTC 1 Focus

• Horizontal by design, JTC 1 develops information technology standards that are applicable across domains
  • Enable interoperability and expectations that result in higher quality experiences
  • Encapsulate expertise and best practices
  • Foundation for innovation
• JTC 1 works together with other IEC and ISO TCs and other organizations that are utilizing information technology to develop domain-specific standards
  • More than 400 liaisons (JTC 1 and its SCs)
  • Emphasis on effective cooperation/collaboration
Technology Areas

Coded character sets
Telecommunications and information exchange between systems
Software and systems engineering
Cards and security devices for personal identification
Programming languages
Digitally recorded media
Computer graphics, image processing
Interconnection of IT equipment (home electronics)
IT security techniques
Office equipment (printing)
Coding of audio, picture, multimedia (JPEG, MPEG)
Automatic ID and data capture (RFID)

Data management
Document description, processing
User interfaces
IT for learning, education, training
Biometrics
Cloud computing
IT Sustainability
IT governance
Internet of Things and Digital Twin
Artificial Intelligence
Brain-Computer Interface
Smart cities
3D printing and scanning
Trustworthiness
Quantum Computing
Summary

• JTC 1 welcomes your participation, your collaboration, and your cooperation
• Interested delegates and experts are encouraged to contact their national body or national committee to become engaged
• ISO and IEC Committees are welcome to establish liaisons with JTC 1 and its subcommittees
• Questions are always welcome
Thank you

Phil Wennblom
philip.c.wennblom@intel.com
jtc1info.org
AI potential Applications in Transport

Neil Frost
ISO TC 204 WG 20 Convenor
Date: 25 May 2022
Introduction

• Automated Driving – Multiple AI Applications
• Smart Networks – Possible AI Applications

• AI Applications for Users
• Conceptual Architecture
• The Future
• Q&A
Automated Driving - Multiple AI Applications

- Distance detection and object avoidance
- Lane Guidance
- Origin Destinations
- Optimal Route determination
- Information flow between vehicles in proximity
- Intersection management and guidance
- Vehicle Platooning
Smart Networks

- Traffic signal optimization based on data fusion
- Congestion Traffic management efficiency
- Travel pattern recognition
- Priority Vehicle management
- Kerbside Parking Management
- Infrastructure Management
AI Applications for Users

• Optimized route for travel, private and public transport
• Cost determination and settlement for transport service
• Available services and selection
• Commuter demand and supply determination
Conceptual Architecture

Role model conceptual architecture: Smart city ITS application service ISO/TR4445
An image of role model and functional model of ITS data aggregation servicer
Operational Vehicle in New Zealand and South Korea
Summary

There are a number of focus areas that require AI and BD solutions that make this an extremely interesting space.
Thank you

Neil Frost

nfrost@isaha.co.za
The whitepaper of ISO/TMB SMCC

2022-05-24

Christoph Preuße, Chairman ISO/TMB SMCC
Chairs of:
ISO/IEC JTC 1  Information technology
ISO/IEC JTC 1/SC 7  Software and systems engineering
ISO/IEC JTC 1/SC 17  Cards and personal identification
ISO/IEC JTC 1/SC 27  IT Security techniques
ISO/IEC JTC 1/SC 32  Data management and interchange
ISO/IEC JTC 1/SC 37  Biometrics
ISO/IEC JTC 1/SC 38  Cloud Computing and Distributed Platforms
ISO/IEC JTC 1/SC 40  IT Service Management and IT Governance
ISO/IEC JTC 1/SC 42  Artificial intelligence
ISO/TC 10  Technical product documentation
ISO/TC 10/ SC 10  Process plant documentation
ISO/TC 39  Machine tools
ISO/TC 39/ SC 10  Machine tools - Safety
ISO/TC 184  Automation systems and integration
ISO/TC 184/ SC 1  Physical device control
ISO/TC 184/ SC 4  Industrial data
ISO/TC 184/ SC 5  Interoperability, integration, and architectures for enterprise systems and automation applications
ISO/TC 199  Safety of machinery
ISO/TC 211  Geographic information/Geomatics
ISO/TC 261  Additive manufacturing
ISO/TC 292  Security and resilience
ISO/TC 299  Robotics
plus Chair of IEC SyCSM
The SMCC - whitepaper

The overall goal of the whitepaper is to facilitate for companies, and other stakeholders, to adapt to, and benefit from, the concept of Smart Manufacturing

"Manufacturing that improves its performance aspects with integrated and intelligent use of processes and resources in cyber, physical and human spheres to create and deliver products and services, which also collaborates with other domains within an enterprise’s value chain”

Source: Definition (ISO SMCC resolution 114/2017)
... to explain Smart Manufacturing by utilizing models from innovation.

With certain regularity new disruptive technologies become available and pave the ground for a new wave of innovations.

When the effect of the new innovations is large enough, they will revolutionize the current norm of how things are seen and done.
In the whitepaper SMCC will present new disruptive technology that are matured enough for industry to leverage on, they will be called “Enablers of Smart Manufacturing”.

SMCC will also present
- a set of design principles, referred to as the Enablers for SM, that are of high relevance for achieving a successful implementation of Smart Manufacturing,
- the Effects that are foreseen with Smart Manufacturing
Technologies behind the current revolution are largely connected to computer power and computational capabilities:

- *cheap (hence unlimited) data storage*
- *fast (hence responsive) computer and calculation power*
- *reliant (hence place-independent) communication*
Future effects

What future effects will it lead to?

• Circular manufacturing;
• Model based Manufacturing;
• Lights out factories;
• Product personalization
• Preventive and predictive maintenance;
• Edge automation;
• Servitization;
• Data-driven business models
Needed principles

What **principles** are needed to enhance the development?

- Terminology and reference models;
- Concepts related to decentralization, modularization and virtualization.
- Integration and interoperability vertically and horizontally;
- Digital twin and digital thread;
- Product transparency;
Enablers – Enhancers – Effects
Be beneficial

Standards: ISO, IEC and ITU for
• the enablers
• treating the enhancers
• the effects

Roadmaps for Enablers – Enhancers – Effects to get a „SMCC-matrix“

Appendix with definitions of used terminology
Thank you very much for your attention

Christoph Preusse

Chairman SMCC
c.preusse@bghm.de
Health informatics and AI, the road to Interoperability

Michael L. Glickman
Chair, ISO TC/215 Health informatics
President, Computer Network Architects
No Global Consensus on AI definition

- Different definition used by regulators, providers, standards organizations, developers
- International Medical Device Regulators Forum (IMDRF) working towards publishing glossary in Q2 2022
  - IMDRF started with 16 different definitions
  - EU proposed 6 different definitions
  - Eventually, IMDRF decide to not define AI at all
What is AI? FDA...

- **Artificial Intelligence** has been broadly defined as the science and engineering of making intelligent machines, especially intelligent computer programs (McCarthy, 2007). Artificial intelligence can use different techniques, including models based on statistical analysis of data, expert systems that primarily rely on if-then statements, and machine learning.

- **Machine Learning** is an artificial intelligence technique that can be used to design and train software algorithms to learn from and act on data. Software developers can use machine learning to create an algorithm that is ‘locked’ so that its function does not change, or ‘adaptive’ so its behavior can change over time based on new data. Some real-world examples of artificial intelligence and machine learning technologies include:
  - An imaging system that uses algorithms to give diagnostic information for skin cancer in patients.
  - A smart sensor device that estimates the probability of a heart attack.

**NOTE:** AI encompasses wide spectrum of software tech ranging from simpler “if -> then” statements, look up tables to Machine Learning to Deep Learning.
Medical devices based on "medical device data" and using ARTIFICIAL intelligence technology to achieve their intended use (i.e., medical use)

**NMPA AI Definition**

- **AI MD has a clear definition** from “AI Medical Device Registration Review Guideline” (No. 8, 2022) on March 9, 22
- **AI Medical Device** is the medical devices based on “medical device data” and using artificial intelligence technology to achieve their medical purposes (i.e., medical use)
- Medical device data: 1) **medical images** (X-Ray, CT, MRI, U/S, etc.); 2) physiological signals, e.g. ECG, EEG; 3) IVD- pathological images, etc.; 4) data generated by non-MD with medical use, e.g. digital health product generated ECG data, etc.

**NMPA AI Classification**

- AI MD will be **Class II** or **Class III** (SaMD or medical device with AI SW component belongs to Class III);
- Classify as **CAD** (Computer-aided detection/diagnosis/triage) or **non-CAD** (e.g. process optimization, image quality/speed improvement, automatic measurement/segmentation, 3D reconstruction) by clinical unities.

**NMPA AI Class III practice in China**

- Only 1 MNC got NMPA Class III clearance—**CAD from Siemens** with intl clinical trial data & FDA clearance;[3]
- GE & Siemens has not cleared AI features with system[3]
- **Philips NMPA submissions got rejected by NMPA mainly due to** 1) Lacking of China-GCP compliant clinical trial evidence for CAD; 2) No Chinese population training dataset

<table>
<thead>
<tr>
<th>BU</th>
<th>AI feature</th>
<th>Remark for NMPA rejection</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR</td>
<td>MRCAT-Brain</td>
<td>× No Chinese population in training dataset; china clinical study required.</td>
</tr>
<tr>
<td>EDI-ISP-12</td>
<td>ClearRead</td>
<td>× Requested to submit as SaMD</td>
</tr>
<tr>
<td></td>
<td>LNA CAD</td>
<td>× No GCP compliant clinical study data</td>
</tr>
<tr>
<td></td>
<td>CT Pulmo</td>
<td>× No GCP compliant clinical study data</td>
</tr>
</tbody>
</table>

Definitions in Standards

ISO/IEC JTC1 SC42 is developing a series of standards for AI, across all industries. It is hoped that these standards can be used as the basis for sector-specific standards such as healthcare. Our details might be different, but the general concepts are (hopefully) the same (an earlier draft of this definition was considered by IMDRF.)

AI System: Engineering system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives.
“Artificial intelligence constitutes a host of computational methods that produce systems that perform tasks normally requiring human intelligence. These computational methods include, but are not limited to, machine image recognition, natural language processing, and machine learning. However, in health care a more appropriate term is “augmented intelligence” (AI), reflecting the enhanced capabilities of human clinical decision making when coupled with these computational methods and systems.”

Augmented intelligence in health care (AMA Board of Trustees, 2018)
Terminology challenges: If you can’t agree on what to call it, there’s no way you can agree on what it should do

Artificial Intelligence practitioners have their own set of terminology that sometimes conflicts with what we think of in Health informatics.

• “AI” might mean multiple techniques that lead to ‘intelligent’ behavior, but is often used to talk specifically about “Machine Learning”

• “Validation” for medical devices often refers to meeting user needs; but “validation” in data science is making sure the data and model are valid.

• “Bias” is something that data scientists try to eliminate, but many caregivers want algorithms to be biased towards their particular patient demographics.

• “Supervised” learning sounds as if there is someone that is supervising the software as it is learning (like you supervise a new employee.) But data scientists use this to mean data that is labelled (vs. an unlabeled pile of information)
Why should I care?

Machines will not replace physicians but physicians using AI will soon replace those not using it.”
AI-augmented multidisciplinary teams: hype or hope? (Di leva, 2019) The Lancet

“Interest in artificial intelligence in healthcare soared in 2019 with investors pouring $4 billion into the sector across 367 deals... That's up from nearly $2.7 billion invested in healthcare AI in 2018 across 264 deals.” “Healthcare led AI investment, topping the $2.2 billion raised by financial and insurance AI.”

Fierce Healthcare coverage of CB Insights Report

AI will be critical in meeting the care needs of a growing, aging population facing projected physician shortages. However, concerted effort is needed to assure this tech advances the quintuple aim.
National Academy of Medicine Report on AI (Matheny et al., 2019)
Example of FDA Cleared AI: IDx-DR
Screening retinal images to detect retinopathy

• Diabetic retinopathy is when high blood sugar levels damage blood vessels in retina

• Retinal images are uploaded to a cloud server where IDx-DR software analyzes and tells doctor “more than mild diabetic retinopathy detected; refer to an eye care professional” or “negative for more than mild diabetic retinopathy; rescreen in 12 months”

• Note that this product provides a screening decision without having the clinician also interpret the results, and is therefore usable by healthcare providers not normally involved in eye care.

Source: https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm604357.htm
Why now? What is driving this?

There have been several AI Winters in the past – why is AI warming up again? Why do we think this might be the last Winter?

Two + one reasons:
1. IoT / Big Data
2. Video games!
3. The IoMedical Things
Types of AI Applications in Healthcare

Artificial Intelligence will bring significant improvements to healthcare in many ways

- Improve diagnosis and effectiveness of therapy
- Managing population health
- Managing hospital operations
- Managing manufacturing operations
- Managing post-market activities
New Technology = New Risks

Pedestrian fatalities rose 11% in 2016, ‘distraction’ as a contributing factor.

As we gain new skills, what do we give up?

Dance like no one is watching...

Because they are not...

They are checking their phone.
Model Diversification – Refinement & Drift

• Many of us have a long history of working in an environment that is slow to change, and therefore we offer a generic solution to meet common healthcare problems.

• However, with Continuously Learning Systems (CLS), the system wants to change. It wants to be customized for a particular customer. A CLS system can learn about a local population and can optimize for that particular hospital.

• But the manufacturer is responsible for configuration management and change control. The manufacturer is responsible for root cause analysis when something goes wrong and the application doesn’t perform as it should.
Model Diversification – Refinement & Drift continued..

• If every hospital is different, how can you compare performance? How do you handle performance claims that change over time?

• DRIFT: Even if you lock a system and don’t allow changes, patient populations **DO** change over time, and the performance you had 5 years ago might not be the level you are at today. Medical practice also changes over time and that may have an impact to performance.

  • Therefore, having a completely locked system isn’t necessarily the best idea either…

[ISO JTC1 IEC logo]
One challenge is that AI seems mysterious and magical, and people think we need a whole new way of thinking about it.

I propose that we handle data according to these rules:

• Keep records / retain information on the origin of the sample
• Sourcing, processing, preservation, testing and handling should be done in a safe manner
• Protect against contamination, viruses

Note: these concepts are already captured in IMDRF GRRP WGN47 FINAL:2018 document – when talking about tissue samples !!
Levels of Autonomy

<table>
<thead>
<tr>
<th>SAE level</th>
<th>Name</th>
<th>Narrative Definition</th>
<th>Execution of Steering and Acceleration/Deceleration</th>
<th>Monitoring of Driving Environment</th>
<th>Fallback Performance of Dynamic Driving Task</th>
<th>System Capability (Driving Modes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No Automation</td>
<td>the full-time performance by the human driver of all aspects of the dynamic driving task, even when enhanced by warning or intervention systems</td>
<td>Human driver</td>
<td>Human driver</td>
<td>Human driver</td>
<td>n/a</td>
</tr>
<tr>
<td>1</td>
<td>Driver Assistance</td>
<td>the driving mode-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task</td>
<td>Human driver and system</td>
<td>Human driver</td>
<td>Human driver</td>
<td>Some driving modes</td>
</tr>
<tr>
<td>2</td>
<td>Partial Automation</td>
<td>the driving mode-specific execution by one or more driver assistance systems of both steering and acceleration/deceleration using information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task</td>
<td>System</td>
<td>Human driver</td>
<td>Human driver</td>
<td>Some driving modes</td>
</tr>
<tr>
<td>3</td>
<td>Conditional Automation</td>
<td>the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene</td>
<td>System</td>
<td>System</td>
<td>Human driver</td>
<td>Some driving modes</td>
</tr>
<tr>
<td>4</td>
<td>High Automation</td>
<td>the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task, even if a human driver does not respond appropriately to a request to intervene</td>
<td>System</td>
<td>System</td>
<td>System</td>
<td>Some driving modes</td>
</tr>
<tr>
<td>5</td>
<td>Full Automation</td>
<td>the full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver</td>
<td>System</td>
<td>System</td>
<td>System</td>
<td>All driving modes</td>
</tr>
</tbody>
</table>

The automotive industry is also looking at autonomy -- this table is from an automotive standard, SAE J3016.
Loopy Humans - to what degree are “humans in the loop” ??

A simple model of Autonomy

1. **First level**
   - The machine performs a task and then waits for the human user to take an action before continuing.

2. **Second level**
   - The machine can sense, decide, and act on its own. The human user supervises its operation and can intervene, if desired.

3. **Third level**
   - The machine can sense, decide, and act on its own. The human cannot intervene in a timely fashion.
Adaptive artificial intelligence and machine learning technologies differ from other software as a medical device (SaMD) in that they have the potential to adapt and optimize device performance in real-time to continuously improve health care for patients.

- The FDA is considering a total product lifecycle-based regulatory framework for these technologies that would allow for modifications to be made from real-world learning and adaptation, while still ensuring that the safety and effectiveness of the software as a medical device is maintained.
The FDA, Health Canada, and MHRA published a draft set of GMLP Guiding Principles. These principles will help guide the development of future standards and regulations.

| Good Machine Learning Practice for Medical Device Development: Guiding Principles |
|---------------------------------|---------------------------------|
| Multi-Disciplinary Expertise are Leveraged Throughout the Total Product Life Cycle | Good Software Engineering and Security Practices are Implemented |
| Clinical Study Participants and Data Sets are Representative of the Intended Population | Training Data Sets are Independent of Test Sets |
| Selected Reference Datasets are Based Upon Best Available Methods | Model Design is Tailored to the Available Data and Reflects the Intended Use of the Device |
| Focus is Placed on the Performance of the Human-AI Team | Testing Demonstrates Device Performance during Clinically Relevant Conditions |
| Users are Provided Clear, Essential Information | Deployed Models are Monitored for Performance and Re-training Risks are Managed |
In 2019, ISO/TC 215 started a project to assess the impact that AI will have on future TC215 projects, and after publishing the report, established a Task Force to help TC215 develop those standards.

TF5 has been created to (among other things..)

- Develop Definitions
- Maintain a Standards Landscape of what other committees are developing
- Support use cases
- Consider educational projects
- Liaison and collaborate with other organizations.
Summary

We intend to continue Collaboration & Planning between ISO and IEC for Health informatics and AI Standards reflecting the “vertical domain of Health

We should be a chapter in the book of AI across all industries, not write our own book
Thank you

Michael L. Glickman

MGlickman@CNAlnc.com

Chair, ISO/TC215 Health informatics
Creating Trust in AI Through Standards: A Management System Approach

Kimberly Lucy
Director, GRC Standards
Microsoft
Introduction

Artificial intelligence has the power to affect virtually every aspect of human life, from work to healthcare to leisure. At the same time, the damage that can be created by such powerful systems is immense if left unchecked. **How can AI be developed and used in a way that is responsible and that leads to trust and assurance for consumers and other stakeholders? And what is the role of standards in creating this trust and assurance?**
Two components of this presentation—both critical aspects of management systems standards that create trust in AI

The “traditional” approach

The AI Management System (AIMS, ISO/IEC 42001) as central component of a governance, risk, and compliance ecosystem

The “novel” approach

Conformity assessment/certification of management systems + products/service certification that can lead to a joint certification model for AI
Background: What is a management system standard? What is ISO/IEC 42001 (AIMS)?

- Management System Standard:
  - Based on a common “High-Level Structure” with required management clauses (all MSS have these)
  - Focused/scoped to a particular domain or sector (e.g. information security, privacy, AI)
  - Stresses an iterative process of continuous improvement for an organization
  - Risk-based
  - Organizations can be certified by a third-party to the applicable MSS

### Current AIMS Structure

<table>
<thead>
<tr>
<th>Management Clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Context of the organization</td>
</tr>
<tr>
<td>• Leadership</td>
</tr>
<tr>
<td>• Planning</td>
</tr>
<tr>
<td>• Support</td>
</tr>
<tr>
<td>• Operation</td>
</tr>
<tr>
<td>• Performance Evaluation</td>
</tr>
<tr>
<td>• Improvement</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Annex A Controls + Annex B Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Policies related to AI</td>
</tr>
<tr>
<td>• AI management framework</td>
</tr>
<tr>
<td>• Resources for AI</td>
</tr>
<tr>
<td>• Assessing impacts of AI systems</td>
</tr>
<tr>
<td>• AI system life cycle</td>
</tr>
<tr>
<td>• Data for AI systems</td>
</tr>
<tr>
<td>• Information for interested parties of AI systems</td>
</tr>
<tr>
<td>• Use of AI systems</td>
</tr>
<tr>
<td>• Third-party Relationships</td>
</tr>
</tbody>
</table>
AIMS “fast facts”

- Scope of standard allows organizations to define what “responsible AI” means for them, depending on the context of their organization—want organizations to think about responsible AI, with the flexibility to address this within their own operational environment

- System level controls (Annexes A and B) align well with requirements in draft EU AI Act

- Controls are written at an abstract level to allow organizations the flexibility to add more granular detail depending on their context (particularly regulatory context); they are not meant to cover regulatory requirements 1:1

- Written with a strategic approach to other SC42 standards, particularly governance, risk, and compliance (next slide), which complement AIMS and can be used by organizations to help implement a GRC ecosystem of standards to enable trust and assurance
The governance, risk, and compliance (GRC) ecosystem approach

Hypothetical AI MSS Ecosystem use in an organization
How can AIMS/management systems certification lead to customer/public trust?

- Management systems certifications, particularly those for IT products or “digital services” have a history of being the primary international certification mechanism (e.g. ISO/IEC 27001)

- Customers have come to recognize and rely on management system certification for assurance and trust that organizations are doing the right thing, and that products/services can be trusted

- Supply chains heavily rely on management system certifications as methods of mutual trust and assurance
Will management system certification alone be enough, particularly for higher risk AI systems?

- Potentially not—but it will be a key component!
- AI will likely be highly regulated in many cases (e.g., EU AI Act)
- “Productization” of intangible/digital services like AI—expectations by regulatory authorities and others for certification “equal” to that of tangible manufactured products
  - Accreditation of certification takes place under ISO/IEC 17065 for products/services rather than ISO/IEC 17021 for management systems in this case
- There are several problems to be solved in this space—primarily that digital services/intangible products have not historically been certified under 17065. However, there is a potential solution…
Novel approach to using management system for trust and assurance in a changing regulatory landscape: Joint certification

For more information on the joint certification topic, see the following whitepaper:

A Joint Certification Approach for Digital Services and Regulatory Compliance
Summary

An AI management system will be a key component for providing trust and assurance for users of AI systems. Whether in a more traditional form via an organization’s GRC ecosystem and MSS certification, or as part of the novel approach to joint certification for management system and digital services, AIMS will be the foundation on which trust is created.
Thank you

Kimberly Lucy

kilucy@microsoft.com
ISO/IEC JTC 1/SC 42(AI)/WG 2(Data)
Data Quality for Analytics and Machine Learning (ML)

Wo Chang
Digital Data Advisor
ISO/IEC JTC 1/SC 42/WG 2 Data Working Group, Convenor
National Institute of Standards and Technology, US
wchang@nist.gov

May 24, 2022
Motivations – Why data quality is important?

The success of AI machine learning has led to substantial improvements comparable to human performance and beyond across a variety of application domains. However, data quality for training, validation and verification plays a vital role, especially in the areas of accuracy, precision, credibility, consistency, completeness, safety, security, privacy/societal/ethical concerns, etc.

Sample Domains

Healthcare
- Electronic Health Records
- Precision Medicines
- Drug Discoveries

Autonomous Vehicles
- Drones
- Self-driving Cars
- Vehicle-to-Vehicle Communications

Financial
- Stock Market
- Electronic Commerce
- Online Transaction

Sample Applications

Healthcare
- Care Management
- Lab Testing Services
- Others...

Autonomous Vehicles
- Unmanned Aerial Vehicle
- Others...

Financial
- Banking
- Insurance
- Others...
Data Empower Analytics/AI to Meet Applications Need

Data as oil to analytic/AI engine to solve complex application problems.
Data Quality for Analytics and Machine Learning: Issues

Data (structure, semantic, and metadata) quality impacts (wrong data produce wrong trained models) analytic/AI models results influence applications.
Data Quality for Analytics and Machine Learning: Overall Goal

To provide effective data quality for analytic and machine learning in the areas of measures, process, and management at each stage of the data life cycle across different levels of the organization.

Sample Existing Data Life Cycles Management with Specific Goals

CRISP-DM (Cross Industry Standard Process for Data Mining)

SEMMA (Sample, Explore, Modify, Model, and Assess)

KDD (Knowledge Discovery in Databases)

DSL (Data Science Lifecycle)
ISO/IEC 8183 provides an overarching data life cycle framework that is instantiable for any AI system that is applicable across different levels of the organization from idea conception to system decommissioning stages with common terminologies and processes.
ISO/IEC 5259-x Data Quality for Analytics and Machine Learning

A holistic approach is needed to oversee the implementation and operation of data quality measures, data quality management requirements and guidelines, and data quality process for various types of analytics and machine learnings with adequate controls throughout the ISO/IEC 8183 AI Data Life Cycle Framework.

ISO/IEC AI Workshop Series, Novel AI Standardization Approaches Track, Wo Chang, NIST/ITL, May 24, 2022

ISO/IEC 8183 AI Data life cycle framework (under balloting)
ISO/IEC 5259-1 Overview, terminology, and examples

**Scope:** This document provides the means for understanding and associating the individual documents of the ISO/IEC “Artificial intelligence — Data quality for analytics and ML” series and is the foundation for conceptual understanding of data quality for analytics and machine learning. It also discusses associated technologies and examples (e.g. use cases and usage scenarios).

![Data life cycle framework](image)

**Key**
- Stage where data are processed
- Iteration
- Primary development pathway
- Feedback pathway

![Instantiations of DLC model for 5259 series](image)
ISO/IEC 5259-1 Overview, terminology, and example

DQ Concept Framework for Analytics and Machine Learning

- **DQ model**: ISO/IEC 5259-1
- **DQ measures**: ISO/IEC 5259-2
- **DQ assessment**: ISO/IEC 5259-3, ISO/IEC 5259-4
- **DQ improvement**: ISO/IEC 5259-2, ISO/IEC 5259-4

**DQ management**

- **DQ reporting**: ISO/IEC 5259-5
- **DQ governance**: ISO 8000-120
- **Data provenance**
ISO/IEC 5259-2 Data Quality Measures

**Scope:** This document provides a data quality model, data quality measures and guidance on reporting data quality in the context of analytics and machine learning (ML). This document builds on ISO 8000 series, ISO/IEC 25012:2008 and ISO/IEC 25024. The aim of this document is to enable organizations to achieve their data quality objectives and is applicable to all types of organizations.
ISO/IEC 5259-2 Data Quality Measures

Data Quality Measurement Framework
(under development)

DQ Characteristics
1. Accuracy
2. Precision
3. Completeness
4. Representativeness
5. Consistency
6. Relevance
7. Data scalability
8. Context coverage
9. Portability
10. Timeliness
11. Currentness
12. Identifiability
13. Auditability
14. Credibility
15. Understandability
16. Balance
17. Effectiveness
18. Similarity
ISO/IEC 5259-3 DQ Management Requirements & Guidelines

**Scope:** This document specifies requirements and provides guidance for establishing, implementing, maintaining and continually improving the quality for data used in the areas of analytics and machine learning. This document does not define a detailed process, methods or metrics. Rather it defines the requirements and guidance for a quality management process along with a reference process and methods that can be tailored to meet the requirements in this document. The requirements and recommendations set out in this document are generic and are intended to be applicable to all organizations, regardless of type, size or nature.
ISO/IEC 5259-4 Data Quality Process Framework

**Scope:** This document provides general common organizational approaches, regardless of type, size or nature of the applying organization, to ensure data quality for training and evaluation in analytics and machine learning. It includes guidelines for (a) supervised machine learning with regard to the labelling of data used for training machine learning systems, including common organizational approaches for training data labelling; (b) unsupervised machine learning; (c) semi-supervised machine learning; and (d) reinforcement machine learning. This document is applicable to training and evaluation data that comes from different sources, including data acquisition and data composition, data pre-processing, data labelling, evaluation, and data use. This document does not define specific services, platforms or tools.
ISO/IEC 5259-4 Data Quality Process Framework

Typical data labelling methods can include object annotation, bounding box, key-point annotation, instance segmentation and semantic segmentation. Data can be labelled manually by a human annotator or automatically by using pseudo-labelling. Pseudo-labelling is the process of using a model trained on labelled data to predict labels for unlabeled data.

---

**Data Labelling Process**

- **Labelling specifications**
- **Participant roles**
- **Labelling tools**

**Labelling execution**
- Labelling task establishment
- Labelling task assignment
- Labelling process control

**Labelling output**
- Labelling result quality checking
- Labelling result revision

---

ISO/IEC AI Workshop Series, Novel AI Standardization Approaches Track, Wo Chang, NIST/ITL, May 24, 2022
ISO/IEC 5259-5 Data Quality Governance Framework

**Scope:** This document provides a data quality governance framework for analytics and machine learning to enable governing bodies of organizations to direct and oversee the implementation and operation of data quality measures, management, and related processes with adequate controls throughout the data life cycle. This document can be applied to any analytics and machine learning. This document does not define specific management requirements or process requirements specified in 5259-3 and 5259-4 respectively.

![Data Quality Governance Framework Diagram](image-url)
ISO/IEC 5259-5 Data Quality Governance Framework

Relationship between Data Quality Governance Framework and 5259 projects

ISO/IEC AI Workshop Series, Novel AI Standardization Approaches Track, Wo Chang, NIST/ITL, May 24, 2022
ISO/IEC 5259-6 Data Quality Visualization

**Scope:** This document is an overview of data visualization within the context of Artificial Intelligence (AI) and Machine Learning (ML) applications. It is intended to provide examples of where data visualization may be employed by various stakeholders at different stages of the AI life cycle.

Concept of Framework for Data-Quality-Driven Visualization for AI

Figure 5 — Relationship among property to quantify, measurement method, QME and QM  
(Source: ISO/IEC 20521:2012)
## ISO/IEC 5259-6 Data Quality Visualization

<table>
<thead>
<tr>
<th>Sample Data Quality Measure from ISO/IEC 25012</th>
<th>Inherent/System Dependent</th>
<th>Phase in System Life Cycle</th>
<th>Stakeholders</th>
<th>Quality Measurement Requirements</th>
<th>Visualization Requirements</th>
<th>Applicable Visualization Methods</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>I</td>
<td>Inception, Development</td>
<td>Developer, Consumer</td>
<td>Sampling reflects the real world</td>
<td>Index or bar chart</td>
<td>Chart/graph</td>
<td>Multiplot of features</td>
</tr>
<tr>
<td>Completeness</td>
<td>I</td>
<td>Inception</td>
<td>Developer, Regulator</td>
<td>Statistically large data sample</td>
<td>Index</td>
<td>Chart/graph</td>
<td>Multiplot of features</td>
</tr>
<tr>
<td>Consistency</td>
<td>I</td>
<td>Inception, Development, Deployment</td>
<td>Developer</td>
<td>Data samples have same dimensions and features</td>
<td>Index/bar chart</td>
<td>Chart/graph</td>
<td>Group cluster plots</td>
</tr>
<tr>
<td>Credibility</td>
<td>I</td>
<td>Deployment</td>
<td>Consumer</td>
<td>Limited bias and impartial sampling</td>
<td>Index</td>
<td>Chart</td>
<td>Histogram</td>
</tr>
<tr>
<td>Current-ness</td>
<td>I</td>
<td>Inception thru Deployment</td>
<td>Developer, Regulator</td>
<td>Age of the data remains reflective of reality/use case</td>
<td>Time date</td>
<td>Numerical</td>
<td>Numerical</td>
</tr>
<tr>
<td>Accessibility</td>
<td>I/D</td>
<td>Development</td>
<td>Regulator, Consumer</td>
<td>High up time and available to stakeholder</td>
<td>N/A</td>
<td>N/A</td>
<td>Index by access method and stakeholder/privilege</td>
</tr>
<tr>
<td>Compliance</td>
<td>I/D</td>
<td>All</td>
<td>Regulator, Developer</td>
<td>Regulator ensures data gathered per applicable norms</td>
<td>N/A</td>
<td>Index chart of criteria</td>
<td>Glyph of compliance criteria</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>I/D</td>
<td>All</td>
<td>Consumer, Regulator, Developer</td>
<td>Domain dependent</td>
<td>N/A</td>
<td>Numerical/ index</td>
<td>N/A</td>
</tr>
<tr>
<td>Efficiency</td>
<td>I/D</td>
<td>All</td>
<td>Developer</td>
<td>Only what is used to answer a question/problem is in the data</td>
<td>N/A</td>
<td>Index</td>
<td>N/A</td>
</tr>
<tr>
<td>Precision</td>
<td>I/D</td>
<td>Development</td>
<td>Developer</td>
<td>Meets the intended application</td>
<td>Numerical</td>
<td>Numerical</td>
<td>N/A</td>
</tr>
<tr>
<td>Traceability</td>
<td>I/D</td>
<td>All</td>
<td>Regulator, Developer</td>
<td>Lineage is tracked</td>
<td>Index</td>
<td>Chart</td>
<td>Flow chart</td>
</tr>
<tr>
<td>Understandability</td>
<td>I/D</td>
<td>All</td>
<td>Regulator, Consumer</td>
<td>A stake holder can assess the data</td>
<td>Index or binary</td>
<td>Index</td>
<td>Multiplot by stake holder</td>
</tr>
<tr>
<td>Availability</td>
<td>D</td>
<td>All</td>
<td>Consumer</td>
<td>Data exists</td>
<td>Index</td>
<td>Index</td>
<td>Check list</td>
</tr>
<tr>
<td>Portability</td>
<td>D</td>
<td>Development</td>
<td>Developer</td>
<td>Data can be transferred to other systems for other application</td>
<td>Check list</td>
<td>Check list</td>
<td>N/A</td>
</tr>
<tr>
<td>Recoverability</td>
<td>D</td>
<td>Development</td>
<td>Producer, Developer</td>
<td>Data is stored in a manner as to be redundant and errors can be fixed</td>
<td>Check list</td>
<td>Check list</td>
<td>N/A</td>
</tr>
</tbody>
</table>
## ISO/IEC 8183 and ISO/IEC 5259-x Projects Timeline

<table>
<thead>
<tr>
<th>ISO/IEC Standard</th>
<th>Approved</th>
<th>DTR/WD</th>
<th>TR/CD</th>
<th>DIS</th>
<th>FDIS</th>
<th>IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>8183 (IS) AI data life cycle framework</td>
<td>02/07/22</td>
<td>--</td>
<td>02/07/22</td>
<td>05/01/22</td>
<td>12/01/22</td>
<td>05/01/23</td>
</tr>
<tr>
<td>5259-1 (IS) Data quality for analytics and machine learning (ML): Part 1: Overview, terminology, and examples</td>
<td>07/30/20</td>
<td>12/21/20</td>
<td>08/21/22</td>
<td>03/21/23</td>
<td>01/28/24</td>
<td>05/27/24</td>
</tr>
<tr>
<td>5259-2 (IS) Data quality for analytics and machine learning (ML): Part 2: Data quality measures</td>
<td>12/31/20</td>
<td>06/17/21</td>
<td>02/13/23</td>
<td>09/20/23</td>
<td>04/20/24</td>
<td>09/20/24</td>
</tr>
<tr>
<td>5259-3 (IS) Data quality for analytics and machine learning (ML): Part 3: Data quality management requirements and guidelines</td>
<td>07/30/20</td>
<td>12/21/20</td>
<td>08/21/22</td>
<td>03/21/23</td>
<td>01/28/24</td>
<td>05/27/24</td>
</tr>
<tr>
<td>5259-4 (IS) Data quality for analytics and machine learning (ML): Part 4: Data quality process framework</td>
<td>07/30/20</td>
<td>12/21/20</td>
<td>08/21/22</td>
<td>03/21/23</td>
<td>01/28/24</td>
<td>05/27/24</td>
</tr>
<tr>
<td>5259-5 (IS) Data quality for analytics and machine learning (ML): Part 5: Data quality governance framework</td>
<td>02/17/22</td>
<td>05/31/22</td>
<td>03/13/23</td>
<td>03/11/24</td>
<td>10/11/24</td>
<td>03/10/25</td>
</tr>
</tbody>
</table>
SC 42/WG 2 Overview and Officers

Title: Data

Terms of Reference: Standardization in relation to data in the context of artificial intelligence, big data, and data analytics

Officers

- Convenor: Wo Chang
- Secretary: Heather Benko
- Editors: Gautam Banerjee (24668), Colin Crone (8183 PWI), Suwook Ha (5259-1), KyoungSook Kim (5259-2), Martin Saerbeck (5259-3), Wanzhong Ma (5259-4), Gyeung-Min Kim (5259-5)

Membership

310+ Experts from 27 National Bodies: Australia, Austria, Belgium, Canada, China, Denmark, Finland, France, Germany, India, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Netherlands, Norway, Poland, Portugal, Russian Federation, Singapore, Spain, Switzerland, United Arab Emirates, United Kingdom, United States
Questions?

Please contact: wchang@nist.gov
Introduction to the AI Standards Hub

Florian Ostmann (The Alan Turing Institute)
Tim McGarr (BSI)
Background and introduction

Press release

**New UK initiative to shape global standards for Artificial Intelligence**

The Alan Turing Institute, supported by the British Standards Institution (BSI) and the National Physical Laboratory (NPL), will pilot a new UK government initiative to lead in shaping global technical standards for Artificial Intelligence.

From: [Department for Digital, Culture, Media & Sport, Office for Artificial Intelligence](https://www.gov.uk/government/organisations/offices) and [Chris Philp MP](https://www.parliament.uk/members Mosul)

Published 12 January 2022
The AI Standards Hub’s mission

– **Advance responsible AI** by unlocking the potential of standards as a governance and innovation mechanism.

– **Empower stakeholders** to navigate the international AI standardisation landscape and become *actively involved* in it.
  
  ▶ Participation in the development of standards

  ▶ Informed use/assessment of published standards
Motivation for the AI Standards Hub

1) AI standards are **increasingly important** as a **governance** and **innovation** mechanism.
Motivation for the AI Standards Hub

1) AI standards are **increasingly important** as a **governance** and **innovation** mechanism.

2) AI standardisation touches on the **interests of many stakeholders**:
   - AI developers
   - Procurers of AI
   - Users of AI
   - Consumers / civil society
   - Regulators & policymakers
   - Academic researchers

3) The AI standards landscape is **increasingly complex** and **rapidly evolving**.
<table>
<thead>
<tr>
<th>1</th>
<th>AI standards observatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Connecting and community building</td>
</tr>
<tr>
<td>3</td>
<td>Education, training, and professional development</td>
</tr>
<tr>
<td>4</td>
<td>Research and analysis</td>
</tr>
</tbody>
</table>
Signing up

Register for updates and provide input to shape the AI Standards Hub:

www.aistandardshub.org

Contact details:
Florian Ostmann  fostmann@turing.ac.uk
Tim McGarr  Tim.McGarr@bsigroup.com
ISO/IEC JTC 1
SC 42
Novel AI Standardisation Approaches

Viveka Bonde
Introduction of SC 42

- Committee on AI standardization under the joint initiative on Information Technology (JTC 1)
- IEC and ISO produce international standards and with participation by country – one country one vote
- Collaboration with long established sister ISO and IEC committees covering a broad range of app domains
- Works with JTC 1 committees producing horizontal stds in key areas such as security, cloud, IoT, governance
Societal Concerns and Ethics

Overview

- Adoption of transformative technologies like AI have impacts that go beyond the technology
- AI-specific trustworthiness issues e.g. reliability, privacy, security, explainability, controllability
- Emerging issues related to the context of use of the technology, e.g. algorithmic bias, safety directives in industrial AI, eavesdropping
- AI ethical considerations not limited to SC 42 but extend to ISO/IEC TCs in their applications
Societal Concerns and Ethics

Motivation and current focus areas

- Standards can mitigate ethical issues allowing for broad **responsible adoption**.

- Stakeholders include:
  - industry,
  - regulatory,
  - technologists,
  - interest groups, app domain, society at large
Societal Concerns and Ethics

Motivation and current focus areas, cont.

- Ethical considerations and societal concerns **considered across entire SC 42 program**
  - **Dedicated** projects
    - **Overview** of AI ethical and societal concerns with tie-in to trustworthiness and exemplary use cases
    - PWI on **AI best practice guidance for mitigating ethical and societal concerns**
    - **Integrated into and enabled by entire SC 42 deliverables portfolio**. For instance,
      - Governance, Management System Standards (MSS) – 42001 (AIMS), use cases, application guidelines, safety etc.
  - **SC 42 collaborating** with other work programs such as OECD, UNESCO, European Commission
Societal Concerns and Ethics

TR 24368 - Overview of ethical and societal concerns

- Links to traditional standardization work by adding the non-technical ethical perspective necessary to address ethical and societal concerns of an AI system and applications:
  - The ethical and societal layer may be emerging as part of “Harmonized Standards” (new legislative framework).
- Addressing that an ethical and societal concern with an AI system and application often depends upon the particular social, economic or physical context of its development, implementation, audience and use.
  - Practices for building and using ethical and socially acceptable AI. Organization’s management process.
  - Stakeholder considerations for building and using ethical and socially acceptable AI
- Visionary element: Enable support and a starting point for other international projects
Key Topics: Societal Concerns and Ethics

Overview of ethical and societal concerns


- 22989; provides standardized AI terminology
- 23894; provides guidelines managing AI-related risks
- 24027; bias in AI-systems
- 24028; introduction AI system transparency and explainability
- 24030; use cases of AI applications
- 38507; guidance on governance implications
Societal Concerns and Ethics

**Overview of ethical and societal concerns**

- Takes into account the context of use of the technology by looking at both technology capability and non-technical requirements such as business requirements, regulatory and policy requirements, application domain needs, and ethical and societal concerns.

- Translating the above into technical requirements.

- Building foundational standards that allow communities to build upon such as terminology, use cases, application guidance and reference architectures.

- Linking technology innovation communities such as proprietary implementations, research and open source communities.
Societal Concerns and Ethics

Overview of ethical and societal concerns

- General all-embracing document
- Has encouraged interdisciplinary expert participation and contribution
- Practical document, organizational focus
- Hands on approach; tangible product - how to address AI themes and principles in practice
- Themes and Principles: Addressing AI systems’ and applications’ potential misuse, abuse and disuse
Societal Concerns and Ethics

**Novel AI Standardization Approaches**

- Align technical experts with non-technical experts
- Necessity to apply expertise inter-disciplinary for creation of ethical AI
- Interplay conformity assessment: The specification of general ethical themes and principles, guidance, considerations - facilitate due diligence of AI systems, on a system/technical level + organizational/ethical level – enhancing public confidence and trust
Societal Concerns and Ethics

Novel AI Standardization Approaches

• International standardization creates a level playing field in markets throughout the world

• New approach legislation feasible

• Importance of conformity assessment

• AI stakeholders: standards define requirements - making compliance with themes and/or principles understandable by concrete guidance
Result: Interplay between legislation, conformity assessment and standardization

Standardization element:
- The link between legislation and conformity assessment stage
- Complement the requirements of legislation
- Facilitates conformity assessment
- Why: Flexible, inclusive and responsive. Expression of the principles of subsidiarity and proportionality
Thank you

Viveka Bonde

viveka@bondebarzey.com

www.bondebarzey.com