# Metadata for Machine-Actionable Documents

Mark Kuster, mjk@ieee.org

NCSL International 141 MII and Automation Committee

*Measurement-Information Infrastructure*

Section 1

# Introduction

# Today's Topics

1. Introduction

# Today's Topics

# Today's Topics

# Today's Topics

# Acronyms

- SoA—scope of accreditation
- DCC—digital calibration certificate (PTB)
- CMC—calibration and measurement capability
- KCDB—key comparison database
- NCSLI—NCSL International
- FAIR—findable, accessible, interoperable, reusable
- PID—persistent identifier
- M-Layer—metrology information layer to support measurement systems

**Definition**

MII (measurement information infrastructure)
—set of normative standards that unambiguously define data structures, taxonomies, service protocols and security for locating, communicating and sharing measurement information

# Incremental Progress toward an Ideal DCC

DCC development strategy advantages:

- Target an easy transition—increases adoption rate
- Quick results and incremental progress—no inordinate delays for the "perfect" DCC
- Pareto coverage—widest applicability in the shortest time

Some ideal DCC elements for fully realizing its potential:

- Metadata with PIDs, e.g., for instrument types
- A universal measurement-results model (Brown, Fluke)
- Rich and rigorous traceability (Hall, MSL—GTC, VNA Tools; White, NRC; Kuster)
- Machine-actionable measurement data

# Toward Machine-Actionable Measurement Results

Distinct quantity concepts and representation levels

1. Quantity kind: most general, associated with measurement units, e.g. length
   - Applies to all quantity values, including uncertainties
   - Wanted: unique quantity IDs for unambiguous quantities
   - Wanted: unambiguous scales & units for any measurement (chemistry, biotech, materials...)
2. Measurand: most specific, the desired measured quantity, e.g., inner diameter
   - Metadata for measurement results in DCCs, functions in instrument specs, CMCs in accreditation statements and the KCDB
   - Wanted: unique measurand IDs
3. A foundation for all measurement data, simple and complex:
   - Measurement models and correction functions
   - Traceability detail
   - Interlaboratory comparisons
   - ...

Section 2

# Quantity Kinds and the M-Layer

# Quantity Kinds

- Associated with quantity values and measurement units: $L = 1\,\mathrm{m}$
- Group references and expression conversions: $1\,\mathrm{m} = 100\,\mathrm{cm}$
- Implicitly understood in the sciences and their applications
- Problems for machines:
  - Units do not always define the quantity kind, e.g.:
    - $10^\circ$: optical incidence or phase angle?
    - Dimensionless quantities with unit 1?
    - $12\,\mathrm{N\,m}$: torque or energy?
    - Free-form quantity descriptions do not solve the problem.
  - Scales not explicitly distinguished ($5\,^\circ\mathrm{C}$: absolute temperature or temperature interval?)
  - Uncontrolled scale operations ($2\,^\circ\mathrm{C} + 1\,^\circ\mathrm{C} = ?$)
  - Confusion from non-SI measurement units (e.g., “g” for standard gravity)

# M-Layer Proposal for Quantity Kinds & Units

Remove all ambiguity and generalize to all measurement scenarios.

- Explicitly augment a quantity value $q\ [Q]$ with its quantity kind (aspect) $q\ [Q]\ \langle Q \rangle$.
- Relate aspects to compatible scales and units
- Handle all scales: ratio, interval, cyclic, logarithmic, ordinal, nominal, ...
- Handle any unit system.
- Uniquely identify aspects, scales and units in FAIR registries.
- Provide conversions as symbolic equations $(x\pi/180)$

Section 3

# A Taxonomy for Measurands

# Measurands

- The precise quantity intended for measurement (radius or diameter?)
- Laboratories require accreditation for competency in specific measurement classes.
- The measurement classes should have unambiguous meaning and correspond to instrument functions (calibration requirements) and calibration results.
- Problems for machines:
  - The KCDB and accreditation statements present CMCs in a variety of ad hoc and unharmonized classification schemes.
  - Free-form text descriptions do not suffice.
  - Current descriptions correlate neither between document types nor organizations.

# Example: Flow-Rate Descriptions in Use

Names from recognized bodies:

- DANAK: “massestrøm” and “volumenstrøm” under “flow” and by instrument type: anemometre, energimålere (varme), massestrømsmålere, or volumenstrømsmålere
- KCDB: “fluid flow” (with “gas flow” or “liquid flow” and species)
- *ISO-IEC 80000*: “mass flow rate” and “volume flow rate”
- NVLAP: “[20/M05] Flow Rate”, plus “liquid flow” and “gas flow”
- DAkkS: ‘Gas flow rate”, “Volume of flowing gases”, and “Mass of flowing gases”

Neither names nor quantities consistent!

# Example: Flow-Rate Descriptions in Use

Looking at approved SoAs from various ABs and Laboratories, we additionally find

> *"Mass Flow", "Liquid Flow", "Gas Flow", "Flow - Air", "Flow - Gas", "Flow - Liquid", "Flow - Gas (Air)", "Air/Nitrogen Flow", "Flow Rate by Volume", "Air Volume Flow", "Flow Hydraulic", "Fuel Flow", "Flow Rate by Volume for Compressible Gas", "Volumetric Flow Rate (Water)", "Liquid Flow Rate Inline", "Liquid Flow Rate Non-Intrusive", "Gas Flow - Leak", "Gas Leak", "Gas Flow Rate Into Vacuum", "mole-flow-rate", "Flow Meter Factor", "Flow Calibration Factor", "Flow Meter", "Determination of Flow Meter" (by gas or liquid species), "Electrical Output of Flow & Pressure Devices",*

and many other categorizations by flow instrumentation.

Further confusion: melt-flow index, flow velocity, air velocity, evaporation, load rate. Wow!

# Solution: Organize a Hierarchical Taxonomy Tree

- Source.MassFlowRate.Gas
- Measure.MassFlowRate.Gas.Inline
- Measure.MassFlowRate.Gas.NonIntrusive
- ...MassFlowRate.Liquid...
- Source.VolumeFlowRate.Gas
- Measure.VolumeFlowRate.Gas.Inline
- Measure.VolumeFlowRate.Gas.NonIntrusive
- ...VolumeFlowRate.Liquid...
- Source.Ratio.MassFlowRate.Gas.MeterFactor
- Source.Coefficient.Voltage.MassFlowRate.DC.Gas

Only the whole taxon has machine meaning; individual tokens (except Measure and Source) only facilitate development.

# Solution: Abstract Taxon Definitions (XML)

**Source.MassFlowRate.Gas**

*This process sources a reference mass flow rate of gas for calibrating gas flow meters. The instantaneous mass flow rate $q_m$ equals $dm/dt$, sometimes estimated as $\Delta m/\Delta t$ using the total mass $\Delta m$ flowing through a defined space in time $\Delta t$.*

- **Required Parameters (with ranges)**
  - Mass flow rate
- **Optional Parameters (with ranges)**
  - Gas Type
  - Gas Temperature
  - Gas Pressure
  - Gas Relative Humidity
  - Gas Compressibility
  - Reference Temperature
  - Reference Pressure
  - Reference Relative Humidity
  - Reference Compressibility
  - Ambient Temperature
  - Ambient Pressure
  - Ambient Relative Humidity
  - Outlet Pressure
  - Reynolds Number
  - Gas Velocity

# Customizable for any CMC, Specification, Measurement Result

| Measurand | Uncertainty | Comments |
|---|---|---|
| Mass flow rate, 1 slpm to 1000 slpm<br>Gas: ambient air, dry nitrogen<br>Gas Temp. (inlet): 23 °C<br>Gas Pressure (inlet): 800 kPa<br>100 kPa to 1000 kPa<br>Gas Relative Humidity: 45 %<br>Ref. Temperature: 20 °C<br>Ref. Pressure: 101.325 kPa<br>Ref. Relative Humidity: 36 %<br>Ref. Compressibility: 0.9997<br>Ambient Temperature: 23 °C<br>Ambient Pressure: 800 kPa<br>. . . | ±0.3 % of reference value* | Source for calibrating inline or non-intrusive flow meters<br>"A Guide to Standardizing Digital Calibration and Measurement Capabilities", *Metrologist*, Jan 2022. |

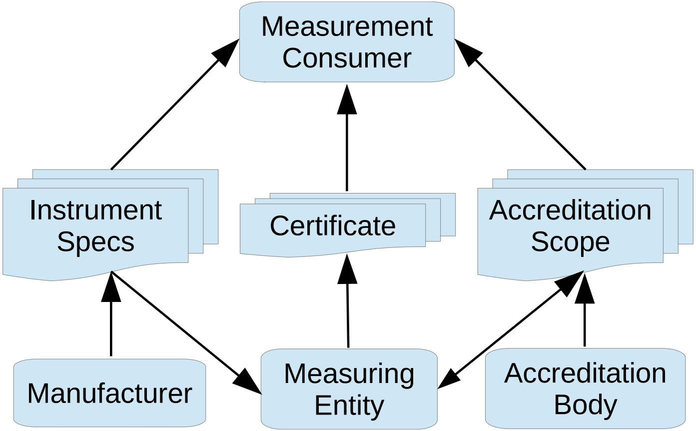
# MII Proposal for Measurands

1. Standardize a taxonomy of measurands.
2. Link taxons to KCDB entry IDs and NMI service codes.
3. Tag digital document data with the taxons as metadata.

Spec: `Source.Voltage.DC`

SoA: `Measure.Voltage.DC`

Cert: `Source.Voltage.DC`



# Example Application

CMC UID `5df0`...
(SoA Data Excerpt)

`Measure.Voltage.AC`, $V$: $1\,\mathrm{V} \pm 50\,\mu\mathrm{V\,V^{-1}}\,V$

- `Frequency`, $f$: $10\,\mathrm{kHz}$ to $100\,\mathrm{kHz}$

Instrument Range UID `78a5`...
(Instrument Specs Excerpt)

`Source.Voltage.AC`, $V$: $(1\,\mathrm{V}$ to $10\,\mathrm{V}) \pm (100\,\mu\mathrm{V\,V^{-1}}\,V + 10\,\mu\mathrm{V})$

- `Frequency`, $f$: $10\,\mathrm{kHz}$ to $20\,\mathrm{kHz}$

Calibration Results
(DCC Data Excerpt)

`Source.Voltage.AC`

- Instrument Range UID `78a5`...
  - $V_{nom} = 1\,\mathrm{V}$, $f_{nom} = 10\,\mathrm{kHz}$
    $V_m = 1.000\,034\,9\,\mathrm{V} \pm 50.001\,745\,\mu\mathrm{V}$
    CMC: `Measure.Voltage.AC` `5df0`...
  - $V_{nom} = 1\,\mathrm{V}$, $f_{nom} = 20\,\mathrm{kHz}$
    $V_m = 1.000\,025\,1\,\mathrm{V} \pm 50.001\,25\,\mu\mathrm{V}$

  ...

No rounding in digital data!

M-Layer encoding not shown

Ultimately, replace these simple uncertainty entries with rigorously encoded traceability.

Contact Diego Coppa, INTI, PTB for exploratory ideas on embedding taxons in DCC XML.

# Draft Taxons

Capacitance
Conductance
Conductivity
Current.AC
Current.AC.Noise.RMS
Current.AC.Sinewave
Current.AC.Sinewave.2Phase
Current.AC.Sinewave.3Phase
Current.AC.Squarewave
Current.AC.Trianglewave
Current.DC
Current.DC.Delta.Current.LoadEffect
Current.DC.Delta.Current.SourceEffect
Current.DC.OutputAndReadback
Density.Mass.Gas
Density.Mass.Liquid
Density.Mass.Solid
Energy.AC.Sinewave
Energy.AC.Sinewave.Simulated
Energy.AC.Sinewave.Simulated.2Phase
Energy.AC.Sinewave.Simulated.3Phase
Energy.DC
Energy.DC.Simulated
Force
Frequency
Frequency.AmplitudeModulation.Rate
Frequency.FrequencyModulation.Deviation
Frequency.FrequencyModulation.Rate
Frequency.PhaseModulation.Rate
Humidity.Absolute
Impedance
Inductance
Length
Length.Circumference
Length.Diameter
Length.Form.Flatness
Length.Form.Parallelism
Length.Form.Perpendicularity
Length.Form.Roughness
Length.Form.Roundness
Length.Form.Sphericity
Length.Form.Straightness.Axis
Length.Form.Straightness.Surface
Length.Radius
Mass.Apparent
Mass.Conventional
Mass.True
Phase.PhaseModulation
Phase.ReflectionFactor.RF
Phase.TransmissionFactor
PhaseNoise.SideBand
Power.AC.Sinewave
Power.AC.Sinewave.Simulated
Power.AC.Sinewave.Simulated.2Phase
Power.AC.Sinewave.Simulated.3Phase
Power.DC
Power.DC.Simulated
Power.RF.Sinewave
Pressure.Hydraulic.Static
Pressure.Pneumatic.Absolute.Static
Pressure.Pneumatic.Differential.Static
Pressure.Pneumatic.Gage.Static
Ratio.AmplitudeModulation
Ratio.AmplitudeModulation.Delta.Rate
Ratio.Density.Mass
Ratio.Distortion
Ratio.Distortion.AmplitudeModulation
Ratio.Distortion.FrequencyModulation
Ratio.Distortion.PhaseModulation
Ratio.DutyCycle
Ratio.FrequencyModulation.Delta.Rate
Ratio.Humidity.Relative
Ratio.Humidity.Specific
Ratio.PhaseModulation.Delta.Rate
Ratio.Power.ReflectionFactor.RF
Ratio.Power.RF.Sinewave.Delta.Frequency
Ratio.Power.RF.Sinewave.Delta.Power
Ratio.Power.RF.Sinewave.Harmonic
Ratio.Power.RF.Sinewave.Spur
Ratio.Power.TransmissionFactor
Ratio.PulseModulation.CWtoPulsedPower
Ratio.PulseModulation.OnOffPower
Ratio.Torque
Ratio.Voltage.AC.Ripple.OnDC

# Draft Taxons

Ratio.Voltage.AC.Sinewave.Delta.Frequency
Ratio.Voltage.AC.Sinewave.Delta.Voltage
Resistance
Resistance.Insulation
Temperature
Temperature.Radiometric
Temperature.Simulated.PRT
Temperature.Simulated.RTD
Temperature.Simulated.Thermocouple
Time.Interval

Time.Period
Time.PulseWidth
Time.Transient
Time.Transition
Time.Transition.PulsedRF
Time.UTC
Torque
Torque.HydraulicPressure
Voltage.AC
Voltage.AC.NoisePeakToPeak
Voltage.AC.Ripple.OnDC

Voltage.AC.Sinewave
Voltage.AC.Sinewave.2Phase
Voltage.AC.Sinewave.3Phase
Voltage.AC.Squarewave
Voltage.AC.Trianglewave
Voltage.DC
Voltage.DC.Delta.Voltage.LoadEffect
Voltage.DC.Delta.Voltage.SourceEffect
Voltage.DC.OutputAndReadback
Voltage.DC.Segemented.Delta
Voltage.PeakToPeak

# Separate but Integrated: the MII Taxonomy and the M-Layer

Taxons link to the aspect ID that governs the quantity values.

| Quantity Value | Aspect ID | Taxon (Measure... or Source...) |
|---|---|---|
| | | Length.Circumference |
| | | Length.Diameter |
| | | Length.Form.Flatness |
| 1 ft | | Length.Form.Parallelism |
| 12 in | | Length.Form.Perpendicularity |
| 0.3048 m | ⇔ ⟨length⟩ ⇔ | Length.Form.Roughness |
| 30.48 cm | | Length.Form.Roundness |
| 304.8 mm | | Length.Form.Sphericity |
| | | Length.Form.Straightness.Axis |
| | | Length.Form.Straightness.Surface |
| | | Length.Radius |

Section 4

# Conclusion

# Incremental Steps toward Interoperability

1. M-Layer
   - Applies to all quantity values, including uncertainties
   - Unambiguous quantities to support any measurement software
   - Generalized scales & units for any measurement scenario
2. MII Measurand Taxonomy
   - Unique measurand IDs with unlimited human-readable aliases
   - Fully qualifies measurands for interoperable digital documents
   - (DCCs, instrument specs-DCRs, SoAs)
3. Interoperability Proposal
   1. Tag all measurement data with the measurand taxon metadata.
   2. Map taxons $\Leftrightarrow$ KCDB entries, taxons $\Leftrightarrow$ M-Layer $\Leftrightarrow$ other unit systems.
   3. Encode all quantity values via the M-Layer.

# Collaboration

- Current information
  - Open-source taxonomy and SoA editors: `https://github.com/CalLabSolutions/Metrology.NET_Public`
  - Prototype M-Layer registry and API now up and running
  - Further MII info: http://miiknowledge.wikidot.com/
- Going forward
  - Moving the taxonomy for configuration management: collaboration, submissions, approvals
    - from `https://www.metrology.net/home/metrology-taxonomy/`
    - to `https://github.com/NCSLI-MII` (opening for public participation soon)
  - Open M-Layer and MII Taxonomy specification and governance documents
  - Prototype M-Layer registry to become public in the future

# Acknowledgments

Please see the literature for further details.

Many thanks go to

- DTI and its seminar organizers for the kind invitation
- NCSL International for its MII support
- NCSL International Committee members for their MII development work
- Cherine-Marie Kuster

And Thank You for your time!

Collaboration opportunities? Please bring your expertise!

Questions?