Methodology of Measurement

Dario Petri
Department of Industrial Engineering
University of Trento – Trento – Italy

dario.petri@unitn.it

Outline

UNIVERSITY OF TRENTO - Italy

INTRODUCTION

- ✓ Why do we measure?
- ✓ Measurement: societal impact

FUNDAMENTALS

- ✓ Measurement: what is it?
- ✓ What can be measured
- ✓ Measurement: characteristics
- ✓ Measurement: a definition
- ✓ Types of measurement scales

A FRAMEWORK FOR MEASUREMENT DEVELOPMENT (FraMeD)

- ✓ Planning; Execution; Interpretation

DATA DRIVEN DECISION MAKING

FINAL REMARKS

INTRODUCTION

WHY DO WE MEASURE?

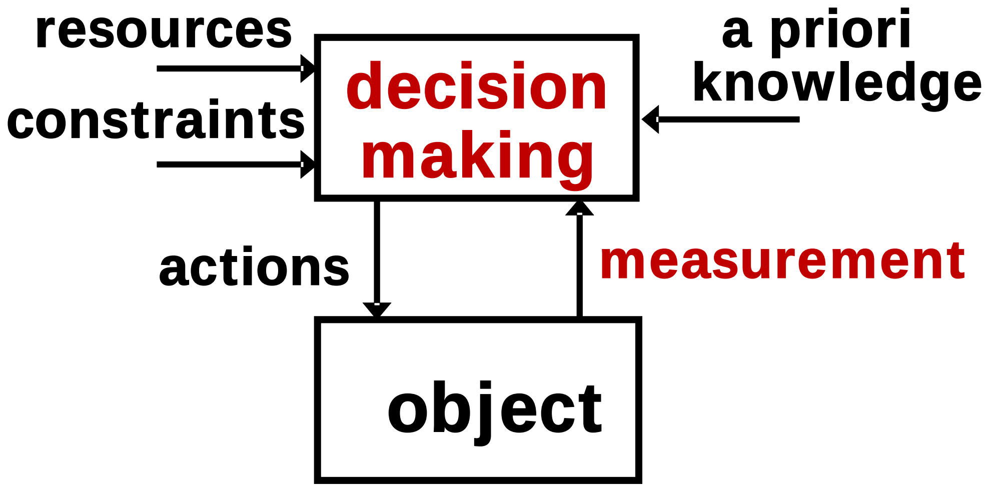
Why do we measure?

UNIVERSITY OF TRENTO - Italy

decisions permeate every moment of our life

information returned by **measurement** is supposed to positively affect the selection of the **best alternative** to achieve established **goals**

decision making support



MEASUREMENT: SOCIETAL IMPACT

Measurement in History

history of measurement is a history of humans, human intelligence and human civilization

- **ANCIENT WORLD**: measurement as a **technical activity** used to **support everyday activities** (trading, building, craftsmanship, …)

- **LAST CENTURIES**: measurement as a **key enabler** for:
  - knowledge advancement (scientific and geographical discoveries)
  - modern civilization (industrialized society, communication, medicine…)
  - social evolution and prosperity

- **NOWADAYS**: **in the era of Big Data**
  measurement is **ubiquitous** and **permeates society**

Every -thing -where -body

one of the most **disruptive technologies**:
it is **changing everything**, transforming **how we live, work, think**

massive amounts of data from:

- the real world (**world datafication**)
- internet, mobile phones and ICT devices

raw material of digital economy, transformed in a major **trade asset** by big players

main drivers of BD phenomenon:

- data driven **decision making** (**D³M**)
- **insight discovery**, by identifying patterns in the datasets

A new ideology: Dataism

Dataist:
- perceives the universe as a **flow of data**
- makes decisions using **only data**
- believes that **AI** can **outperforms human intelligence**
- considers **living organisms** as **biochemical processing systems**

Data Science: a single overarching theory that **unifies all the scientific disciplines** (the **Holy Grail**)

Ex: Beethoven’s Fifth Symphony, a stock-exchange bubble and the flu virus seen as three patterns of dataflow that can be analyzed using the same basic concepts and tools

dark side: **a "quick and easy” (seductive) path to reach the goal**

key question: How to be **aware of the limits of acquired information?**

Big data

volume

amount of data

variety

different shapes and forms

velocity

speed of generated data

variability

various data context, meaning

4Vs characteristics of big data

related to data **quantity** and **technology**

size is celebrated instead of **effectiveness** in support decision

tons of **noise** are useless

data quality: a **huge problem** (given the BD characteristics)

smart data: **useful** part of big data must be **filtered out**

Smart data

VALIDITY

usefulness of data for its intended use

VERACITY

accuracy + contextual data needed to check it

2Vs characteristics of smart data

can be **effectively managed** using principles, methods and tools of:

metrology, a science of **data quality**

- make **aware** of **uncertainty** sources
- ensure **trustworthy data**
- assess and manage the **effects of uncertainty** on the **risk** of wrong decisions

A dichotomy: Post-Truth

denial of scientific facts

facts shaded, selected, and presented within a context that favors interpretation of a truth over another

e.g., flat earth, climate change, evolution, vaccines, smoking, …

emergence of **fake news**

post-truth era

in shaping public opinion, **facts are subordinate to emotions and beliefs**

e.g.: US Presidential election, Brexit referendum, …

FUNDAMENTALS

MEASUREMENT:
WHAT IS IT ?

Value-assignment

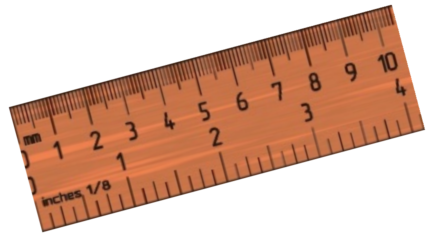
UNIVERSITY OF TRENTO - Italy

EMPIRICAL WORLD

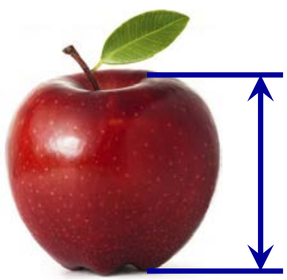
domain of investigation:
universe
(physical, economical, social, …)

OBJECTS

empirical entities
phenomena, bodies, substances, pieces of software, individuals, processes, organizations, …



r **reference**



q **evaluated property (measurand)**
individual property **intended** to be measured

(INDIVIDUAL) PROPERTIES

instance of a general property in an object

inherent aspects of objects

they are the subject to v-assignment (not the objects)

(GENERAL) PROPERTIES

or "**kinds of properties**"
such as length, loudness, extroversion, ...

ABSTRACT WORLD

MODEL

description of a set of **objects**

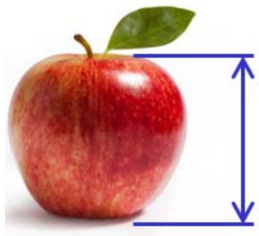
x **property value**

symbol + reference

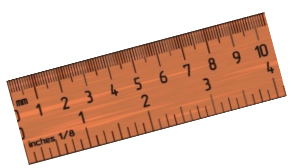
Value-assignment

EMPIRICAL WORLD

q **evaluated property**



r **reference**



v-assignment

ABSTRACT WORLD

x **property value**

symbol + reference

attribution of **values** to **properties** by means of **experimental comparison** with a **reference**

direct/indirect
explicit/implicit

Ex.: subjective assessment, output of a non-calibrated instrument, measurement

MEASUREMENT is **a particular v-assignment**

Kinds of references

UNIVERSITY OF TRENTO - Italy

According to VIM, references that can be used in **measurement** are:

- **measurement unit**
  property value: **product of number and measurement unit**
  also for dimension one (dimensionless) properties

  Ex: length of a rod: 5.34 m; mass fraction of cadmium of a copper sample: 3 μg/kg or 3×10^{-9}

- **measurement procedure**
  Ex: Rockwell C hardness of a given sample (150 kg load): 43.5 HRC(150 kg)

- **reference material**
  Ex: arbitrary amount-of-substance concentration of lutropin in a given sample of plasma (WHO international standard 80/552): 5.0 International Unit/l

- **combination of above**

the choice of reference establishes a **MEASUREMENT SCALE**

Axiomatic definition

UNIVERSITY OF TRENTO - Italy

v-assignament: **experimental comparison** that **maps** an individual property *q* (the measurand) of an object onto a symbol *x*:

$$\mu: Q \rightarrow X$$

set of all possible individual properties q_n
$\{q_1, q_2, \ldots, q, \ldots\}$

a class of symbols
$\{x\}$

or:

$$x = \mu(q)$$

- ✓ **not a one-to-one mapping**: different but indistinguishable individual properties are mapped onto the same symbol
- ✓ issues related to the amount of information are **neglected**
- ✓ mapping execution requires **empirical operations**
  no arbitrary assignment of values
  Ex.: assigning a price to an item, assigning a name to a person

WHAT CAN BE MEASURED?

A “simple” example

At a reception, **numbered tickets** were allotted at the door as people entered so that a **raffle** could be held.

When the winning number, **“97”**, was announced one participant:

- compared it to her ticket to see if she had **won**
  classification, Q = {attendees}
- remarked that there were not 97 **people in the room**
  counting, Q = {total n. of attendees}
  assume tickets assigned consecutively starting at “1”
- compared it to his ticket (“44”) and realized that he **arrived too soon**
  ordering, Q = {attendee’s arrival order}
  assume tickets assigned in increasing order

property of interest depends on **purpose** (information needs)

What can be evaluated?

UNIVERSITY OF TRENTO - Italy

EVALUABILITY (and **MEASURABILITY**) of a **property** requires:

- **empirical existence** (= a cause of an observable phenomenon)
  Ex: hage = height x age of persons - **not a property**, just a mathematical variable which does not provide any new information on persons
- a **model** in which the **object** of interest is acknowledged to exhibit an **instance** of the general property
- an **empirical comparisons** with a reference

meaning of comparison result depends on:

- **purpose** (information needs) motivating the v-assignment
- **a priori knowledge**

What can be evaluated?

Ex: can we **obtain information** about:
organization efficiency, software complexity, quality of life, people satisfaction, people happiness, ...?

Ex: **total weight** of a set of people is a clear concept.
What about their "**total happiness**"?

the **background theoretical infrastructure**
needed to measure is
not so simple as generally assumed

MEASUREMENT:

CHARACTERISTICS OF TRUSTWORTHINESS

MEASUREMENT:

v-assignment expected to return **trustworthy** information

How ensuring **trustworthiness**?

reading an instrument is not enough:

- **trust** cannot be ensured
- returned information can be **misinterpreted / misused**

Measurement characteristics

requirement:

ensure **trustworthiness** of returned information

characteristics:

- **Representativeness** (when mirror the empirical world)
- **Objectivity** (object relatedness)
- **Inter-subjectivity** (subject independence)

C.1) Representativeness

measurement **always** involves **two fundamentally different aspects**:

A) CONVENTIONAL

ensure unique interpretation

arbitrary, but **conscious and shared choices** needed to **operatively perform** v-assignment

examples:

- choice of **unit of measurement**
- Mohs and Brinell scales of **hardness**:
  definition of the concept (measurand) lies in the measurement procedure
- (health related) **quality of life**:
  a specified combination of selected properties:
  extent of constant pain, feeling of worthlessness, feeling of no improvement, …

C.1) Representativeness

B) REPRESENTATIONAL

goal: **understand the behavior** of the considered empirical property **by handling symbols** (how the empirical world works)

TWIN UNIVERSES:

the **abstract world** must **mirror** the **empirical world**

constraints on map μ to ensure that the **behavior of symbols mirrors** the **behavior of properties**

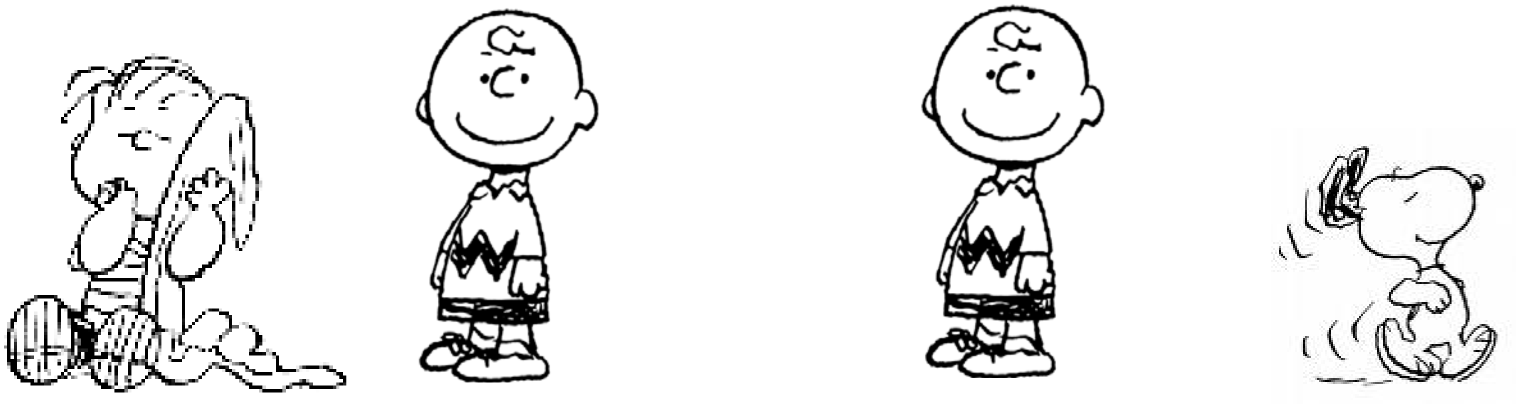
ASSUMPTIONS:

- empirical **properties** are **tangible** (height, length, weight, …)
- empirical **relations** are **observable** (longer than, heavier than, …)

$\Rightarrow$ **univocally perceived** (and definable)

Empirical relations

Ex. of (observable) **EMPIRICAL RELATIONS** for the property **height** defined on the set of **people**:



defined on the set of pairs of people

binary relations: Charlie is ***taller*** than Linus

Charlie is ***much taller*** than Snoopy

unary relations: Charlie is ***tall***, Snoopy is ***not tall***

ternary relation: Snoopy is ***higher*** than Charlie ***if*** sitting on Linus' shoulders

Representation condition

UNIVERSITY OF TRENTO - Italy

set of empirical relations existing on the set of **individual properties** Q
$\{R_{Q1}, R_{Q2}, \ldots, R_{Qn}, \ldots\}$

set of abstract relations existing on the set of **symbols** X
$\{R_{X1}, R_{X2}, \ldots, R_{Xn}, \ldots\}$

$$\Re_Q \leftrightarrow \Re_X$$

must be a **one-to-one mapping** $\Rightarrow$
the two sets have exactly the **same structure**:
behavior of properties **reflected** on behavior of symbols

abstract relations **REPRESENT** empirical relations

Representation condition

empirical world **abstract world**

μ μ

1.20 1.50

Charlie is **taller than** Snoopy

(empirical relation)

μ (Charlie) **>** μ (Snoopy)

(abstract relation)

q_1 is taller than q_2 **<u>if and only if</u>** $\mu(q_1) > \mu(q_2)$

for any admissible μ

the **empirical relations** between properties **are preserved** by **abstract relations**

Intangible properties

UNIVERSITY OF TRENTO - Italy

empirical objects may exhibit:

intangible properties
intelligence, satisfaction, quality of life, attitude, …

subjective relations
“more satisfied than”, “worthier than”, “preferred to”, …

perceivable in different ways by different subjects

not linked trough universally accepted theories (**soft properties**)

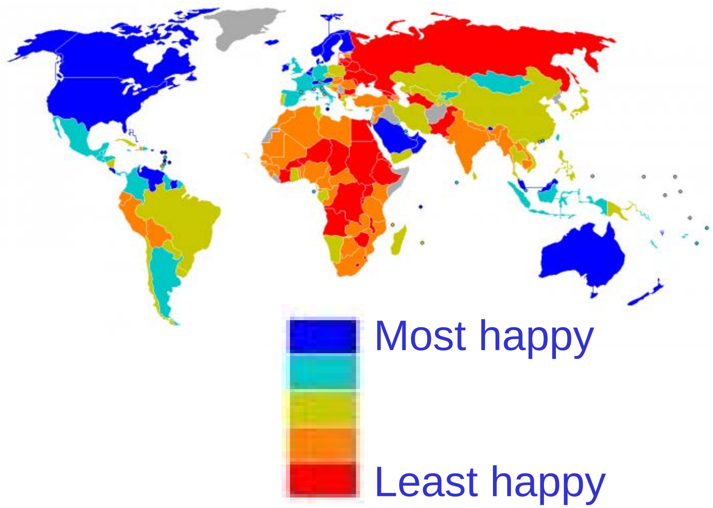
according to representation theory,
in the **abstract world**:

symbols assume a **clear value**

relations have a **unique interpretation**

⇓

properties behavior
can’t be perfectly mirrored



Proxies

UNIVERSITY OF TRENTO - Italy

v-assignment provides **only approximated information** (**proxy**) on the property of interest

proxies for the state of a region economy:
unemployment rate, GDP, …

proxies for quality of research:
n° of citations, h-index, …

quantity of conveyed information not specified
(assessable only through experience and intellectual honesty)

if proxies are used in automatic decisions making,
neglected information can have
dramatic negative impacts

C.2) Objectivity

result of **comparison** with **reference depends only on the measurand** (and the reference), not on other properties

for **physical properties**, objectivity is ensured by:

transducer output,
(mainly) dependent on measurand
(asynchronous direct method of measurement)

functional relationship
linking to other properties
(indirect method of measurement)

$$R = V / I$$

objectivity is a **characteristic of v-assignment**,
not the measurand

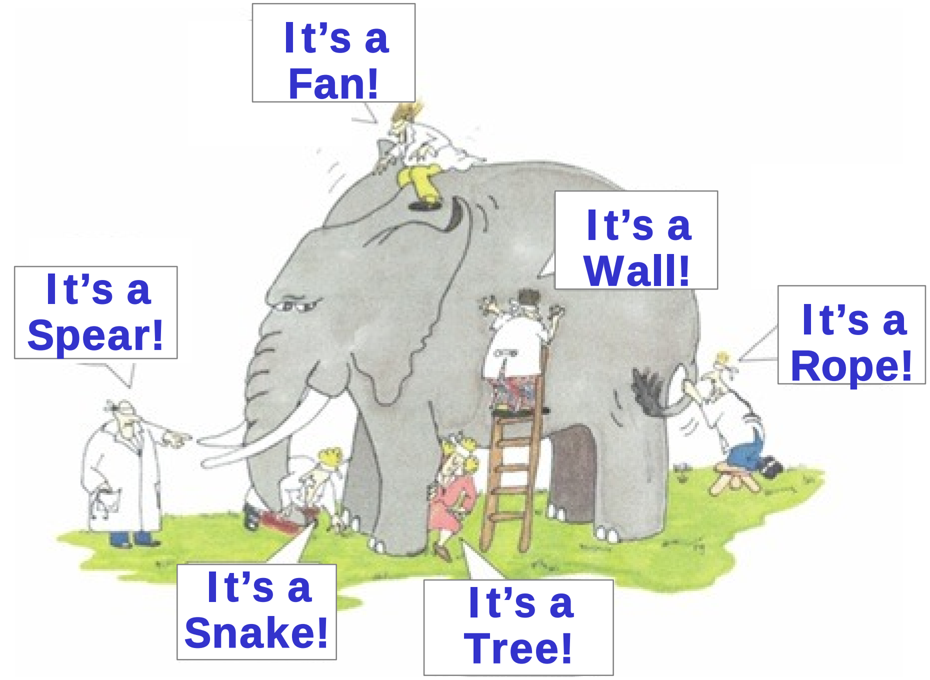
Ex.: for measurands “enjoyment of musical performance” , “happiness”
subjectivity is an integral part of the measurand

C.3) Inter-subjectivity

UNIVERSITY OF TRENTO - Italy

1+1=10

information is **non-ambiguous**:
univocally interpretable by different users in different places and times



inter-subjectivity ensured by:

clear and explicit **shared choices** about **conventional aspects** of measurement

e.g., choice of reference

calibration
of measuring system
(for **physical properties**)

uncalibrated instruments return objective, but non-intersubjective results

Measurement uncertainty

UNIVERSITY OF TRENTO - Italy

complete objectivity and **inter-subjectivity** never achievable

returned information always:

depends also on **other properties** of the empirical environment

can be partially **misunderstood**

e.g., due to limited stability over time of instruments

amount of conveyed information is always finite

knowledge of uncertainty enables the evaluation of the **risk of wrong decision**

measurement uncertainty:

summarizing parameter characterizing the dispersion of the **values** being attributed to the **measurand** for a given **level of confidence**

MEASUREMENT: A DEFINITION

Measurement

UNIVERSITY OF TRENTO - Italy

a **model-based goal-driven process**:

- ✓ providing **information** about a predefined individual property (the **measurand**) of empirical objects
- ✓ based on **empirical comparison** w.r.t. a reference (**v-assignment**)
- ✓ ensuring **known** and **provable level** of **objectivity** and **inter-subjectivity**, specified through measurement **uncertainty**
- ✓ fulfilling **representation condition** (preserves empirical relations among properties), when aimed at **understanding the behavior** of objects

information that does not fulfill this condition must be **used with care** and it normally requires further **knowledge on the context** in which decision making is performed

TYPES OF MEASUREMENT SCALES

Scale types

scales **classified** according to the **set of admissible transformations** on the set of symbols ensuring that the **information** they convey **is preserved** (i.e. representation condition is satisfied)

five main scale types can be defined:

- **NOMINAL**
- **ORDINAL**
- **INTERVAL**
- **RATIO**
- **ABSOLUTE**

increasing level of richness in the conveyed information

(all operations at level *n* allowed at level *n*+1)

Nominal scales

each measured object is placed in a **class**,
according to the value of the considered individual property

symbols are only **class tags** (**no ordering** among classes)

admissible transformations: **one-to-one mapping**

classification is preserved

statistics: **mode**

value of the most commonly occurring item

measurement **confidence**: probability of correct classification

software fault location

specification

design

code

electronic circuits

analog

digital

mixed

Ordinal scales

UNIVERSITY OF TRENTO - Italy

classes **ordered** with respect to the property (called a **quantity**)

admissible transformations: **increasing mapping**

order is preserved

statistics: **median**, **percentiles**

value of the middle-ranked item

software module complexity

trivial | simple | moderate | complex | incomprehensible

Mohs scale of mineral hardness

ability of a harder material to scratch a softer material

1 talc

2

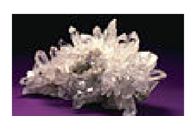
3 calcite

4

5 apatite

6

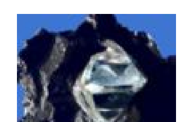
7 quartz



8

9

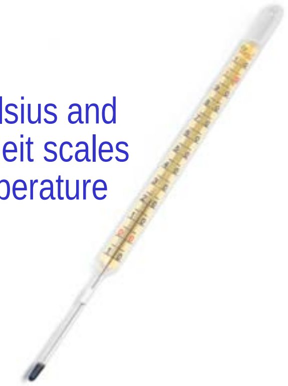
10 diamond



Interval scales

ordinal scales that convey additional information about the **size of the interval** between adjacent classes

ex.: Celsius and Fahrenheit scales of temperature



addition and subtraction are allowed
(not multiplication and division)

saying "q_1 is twice q_2" is meaningless

order and size of intervals are preserved

admissible transformations: **affine mapping** ($y = ax + b$, $a > 0$)

statistics: **mean**, **standard deviation**

uncertainty: dispersion of measured values (at a given **confidence**)

seismic magnitude scale (Richter)

$\propto \log_{10}$ amplitude of waves recorded by seismographs

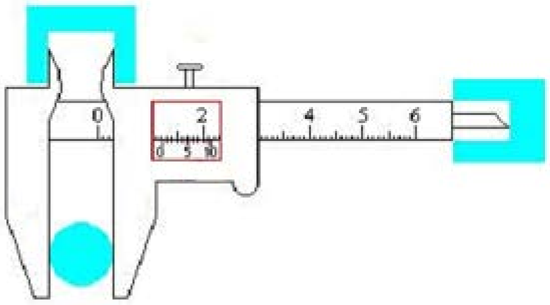
1	2	3	4	5	6	7	8	9	10
micro	minor	light		moderate	strong	major		great	epic

Ratio scales

interval scale with an **absolute** (not conventional) **zero** (**lack** of property)

ordering, **size** of intervals, **ratio** between properties are preserved

ex: length of physical objects



all arithmetic operations are allowed

admissible transformations: **proportional mapping** ($y = ax$, $a > 0$)

ratio is preserved

statistics: geometric mean, coefficient of variation

relative uncertainty (at a given confidence) can be defined

Absolute scale

UNIVERSITY OF TRENTO - Italy

counting the n° of elements in the object

admissible transformations: identity

only one possible mapping = the count

uncertainty can rise also because:

- objects to be counted are **not well defined**
- counting can be difficult

e.g., when dealing with **big data**

Ex.: Counting words in a document (plain-text file)

How many words for: De-noising, sine-wave, yes/no, www.ims.org, EU, USA, …?

selected definition depends on the **information needs** that motivate measurement

A definition: $$\#words = \frac{\#bytes}{< \frac{\#chars}{word} > +1}$$

= 5.5 English dictionary

UNIVERSITY OF TRENTO - Italy

FraMeD

A FRAMEWORK FOR MEASUREMENT DEVELOPMENT

D. Petri, L. Mari, P. Carbone, “A Structured Methodology for Measurement Development",
IEEE Trans. on Instr. and Meas., vol. 64, no. 9, Sept 2015

A conceptual framework

Measurement: a **MODEL-BASED GOAL-DRIVEN PROCESS**

activities performed (many **implicitly**) can be organized in a **CONCEPTUAL FRAMEWORK** inspired to both:

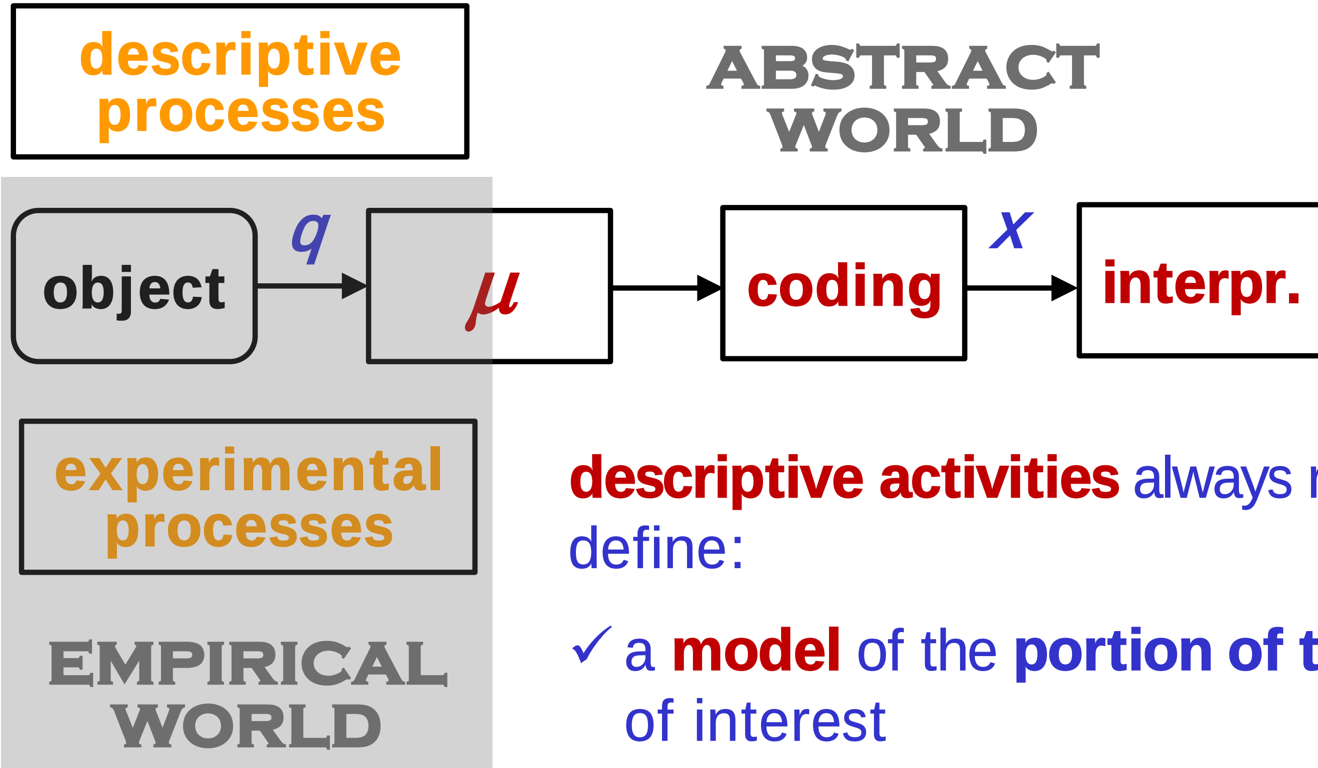
- ✓ **plan-do-check-act** (Deming) **cycle** and
- ✓ models for **product development processes**

STRUCTURAL COMPLEXITY

- ✓ tasks executed according to a loose **ordered sequence**, but
- ✓ many **feedbacks** needed to refine activities until requirements are met

Framework basics

UNIVERSITY OF TRENTO - Italy



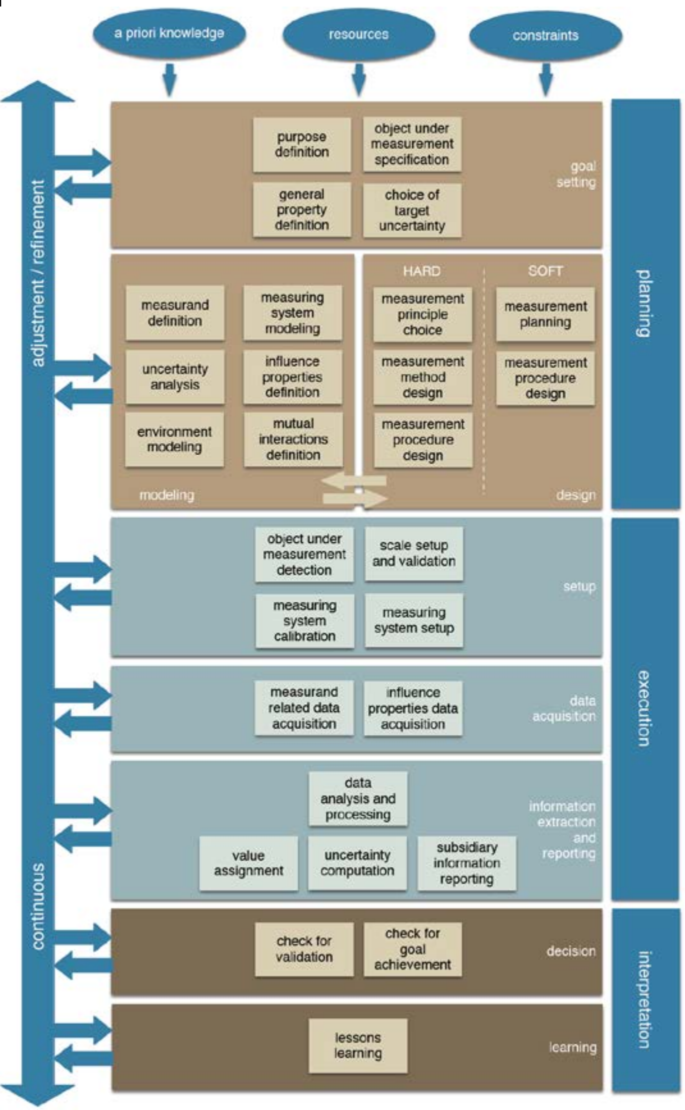
descriptive activities always required to operatively define:

- ✓ a **model** of the **portion of the empirical world** of interest
- ✓ the **measurand** q
- ✓ the **map** μ (instruments, procedures, …)

Ex. "width of a table": Is the table rectangular? A single/more value in a single/more positions? Which instruments can be used? How do we use them?

Framework structure

UNIVERSITY OF TRENTO - Italy



measurement described as a **three-level hierarchically structured process**

1) **stages**
2) composed of **activities**
3) performed through **tasks**

Stages	Activities		
Planning	Goal setting	Modeling	Design
Execution	Setup	Data acquisition	Information extraction and reporting
Interpretation	Decision	Learning	

MEASUREMENT: PLANNING

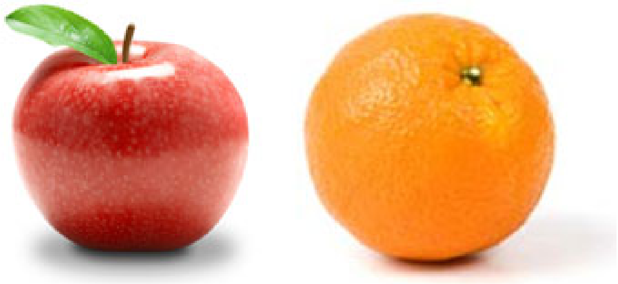
Goal setting

often said: “**you can’t compare apples and oranges**”

basic idea: two objects must be of the **same kind** to compare them and characterize differences, but ...

they have many **obvious similarities** (both are round fruits, have seeds, ...)

to **define the measurand** we need to know the **information needs** motivating the comparison



key question:

why do we want to compare?

rather than: ~~*what*~~ *do we want to compare?*

Goal setting

UNIVERSITY OF TRENTO - Italy

key question: **why do we want to measure?**

MEASUREMENT PURPOSE:

✓ **affect the definition** of the measurand

✓ **depends** on available a-priori knowledge, resources, existing constraints

initial definition of purpose is often **vague** ⇒

refinements needed to achieve a clear and operative purpose definition

often suggested by the following stages

a **minimum quantity of information** about the measurand is **needed** to support decision-making ⇒ fix an **upper limit to measurement uncertainty**:

TARGET UNCERTAINTY

Goal setting

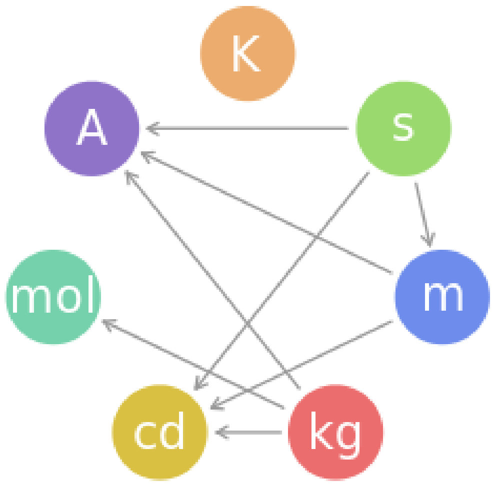
definition of the:

- **CLASS OF OBJECTS** of interest

phenomena, bodies, substances, pieces of software, individuals, research teams, industrial processes, business organizations, …

length, loudness, extroversion, quality, …

- **GENERAL PROPERTY** selected

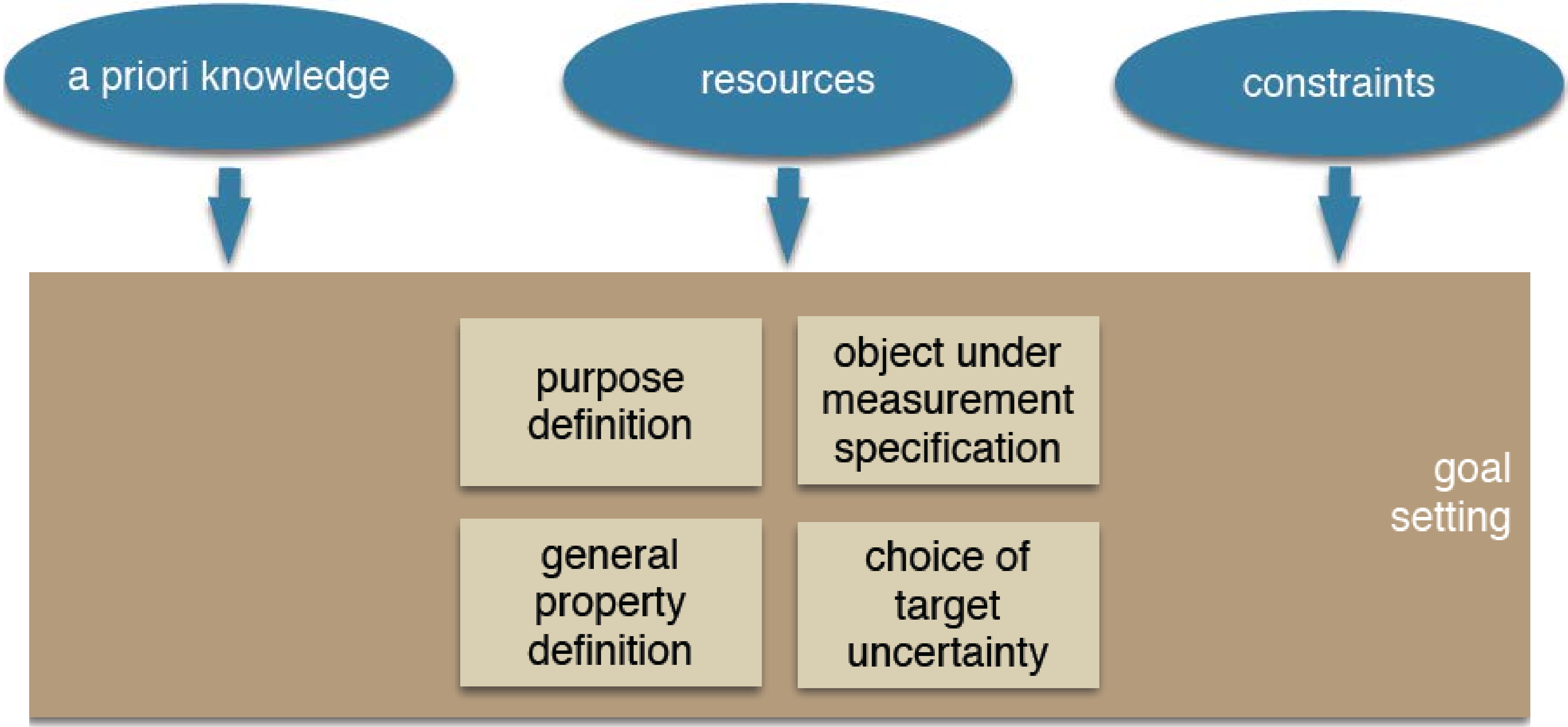


all relevant stakeholders should be involved
in this stage to ensure that
all meaningful underlying aspects are properly considered

Goal setting

UNIVERSITY OF TRENTO - Italy

tasks to be performed:



Modeling

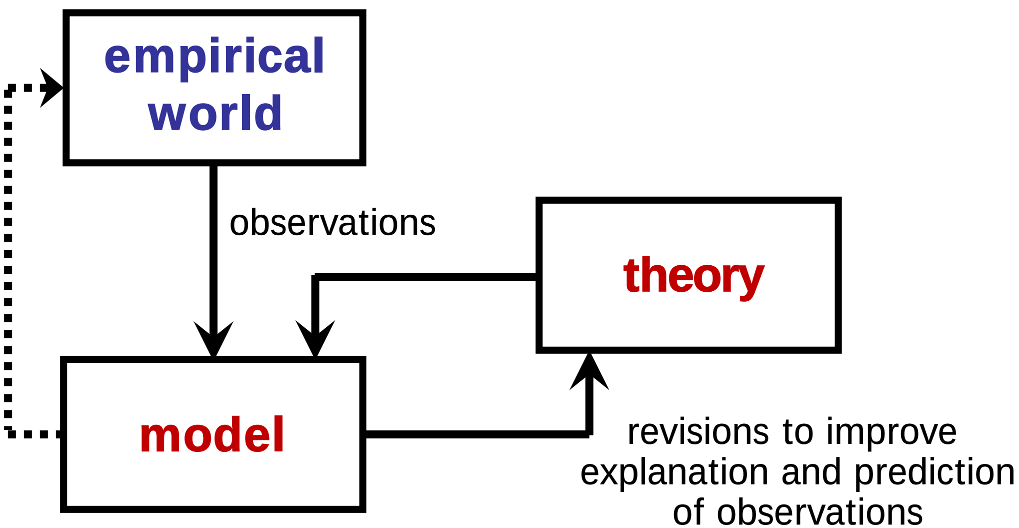
UNIVERSITY OF TRENTO - Italy

abstract process aimed at generating a **conceptual representation** (**model**) of a portion of the empirical world that **depends** on the established **purpose**

using a-priori knowledge, experience, intuition:

1) identify properties of the objects that are relevant for purpose

2) represent them so obtaining a description that is **good enough** to support purpose achievement



different representation forms (mathematical, graphical, mental,...) can be used

Modeling

UNIVERSITY OF TRENTO - Italy

so to avoid possible misuses and misinterpretations

GOOD (trustworthy) models:

using feedbacks with the reality

transparent
based on clear and explicit goals and assumptions

continuously updated
to validate, improve and adapt them when the context change

BAD models
(**dark side** of modeling):

opaque
unquestioned,
unaccountable,
not validated

seem fair,
but may support
biased conclusions

Ex: misuse of research production indicators
racism as a wrong mental model

Model of the context

UNIVERSITY OF TRENTO - Italy

raw measurement result depends on the:

- **measurand** (property intended to be measured)
- employed **references**
- other **empirical properties** pertaining to the experimental setting (e.g.: time, temperature, electric and magnetic fields, operator properties, instruments properties, …)
- **operations** performed during measurement process

(measurement) **CONTEXT**:
the set of **all empirical entities** that "**significantly**" (w.r.t. target uncertainty) **affect** the measurement result

Model of the context

UNIVERSITY OF TRENTO - Italy

to ensure that **objective information** about the measurand can be extracted from raw measurement data,

the **context** has to be **1) identified** and **2) described**

⇓

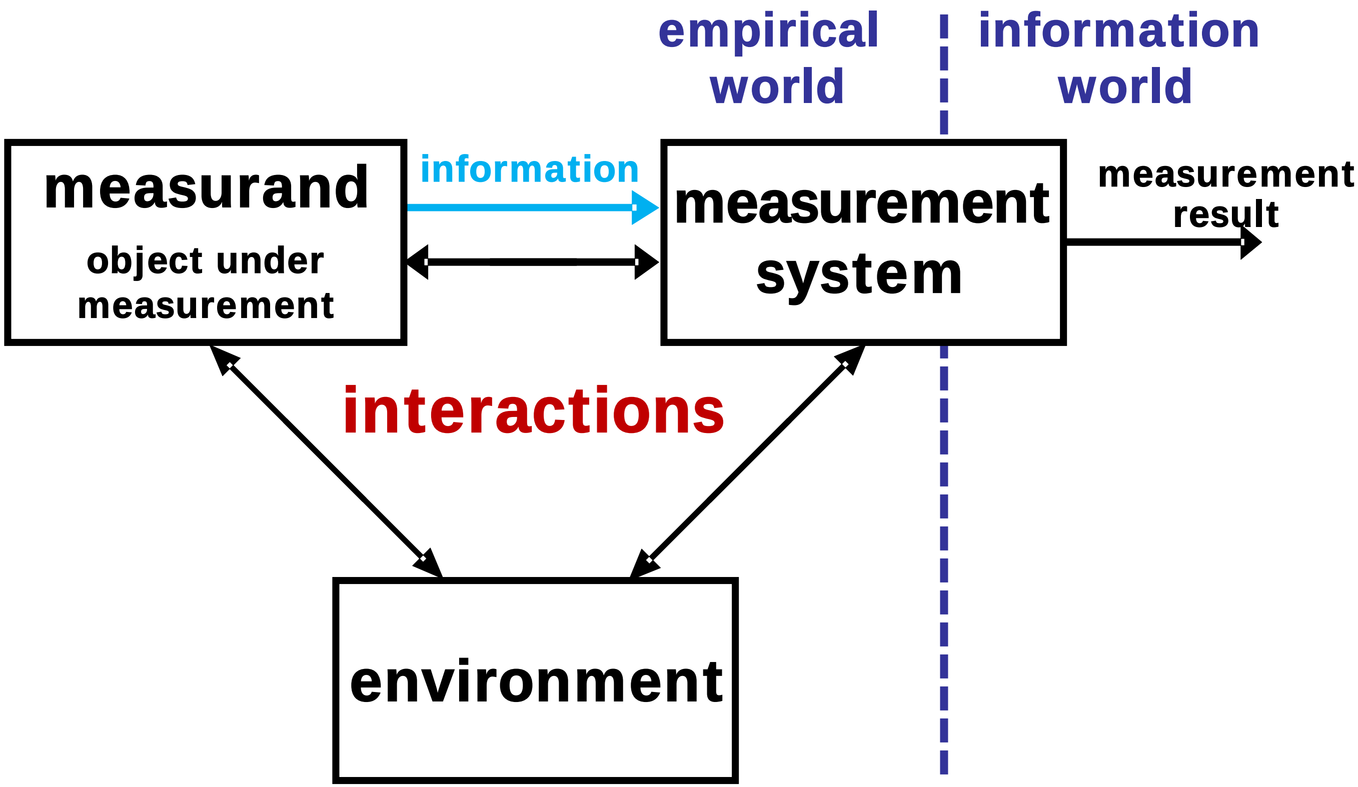
(metrological) **MODEL OF THE CONTEXT**

How can we derive it?

1) Context identification

components of the context:

- **object under measurement** , which exhibits the **measurand**
- **measurement system**, including related operational procedures
- **environment**, all other entities interacting with the above ones



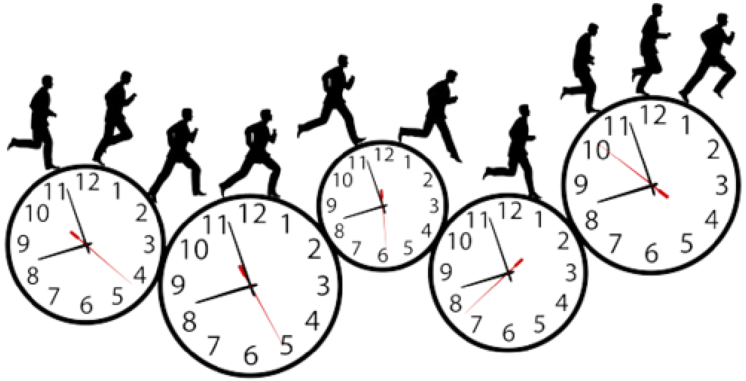
2) Context description

effect of **interactions** and **aging** on raw measurement data described in terms of:

A) INFLUENCE PROPERTIES: properties unlike measurand and standard properties that **significantly** (w.r.t. target uncertainty) **affect** the **measurement result** (these properties can be controlled)

TIME is distinctly considered since it cannot be held constant

B) EMPIRICAL RELATIONS



2) Context description

UNIVERSITY OF TRENTO - Italy

context model contains the **definition of**:

A) SIGNIFICANT PROPERTIES:

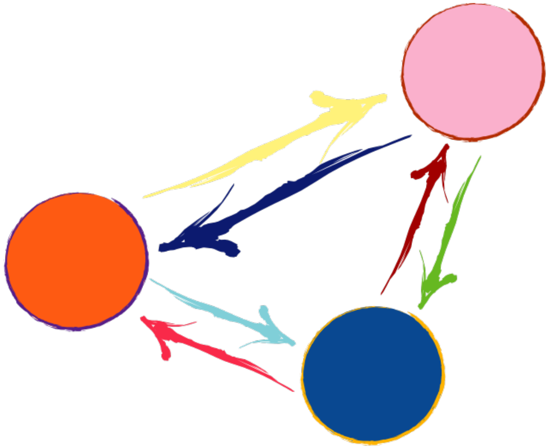
- measurand
- standard properties
- influence properties
- time

each property definition requires: - a *measurement scale*
- a *measurement procedure*

often expressed as **mathematical variables**

B) THEIR MUTUAL INTERACTIONS

often expressed through
mathematical relationships



Context uncertainty

UNIVERSITY OF TRENTO - Italy

models <u>always</u> provide a **partial**, **approximate description**:

properties and interactions are **neglected** because:

- expected to produce negligible effects on raw data
- we are not aware of their effect

$\Downarrow$

UNCERTAINTY OF THE CONTEXT MODEL

$\Downarrow$

it propagates to measurement result, limiting the **amount of conveyed information** about the measurand:

MEASUREMENT UNCERTAINTY

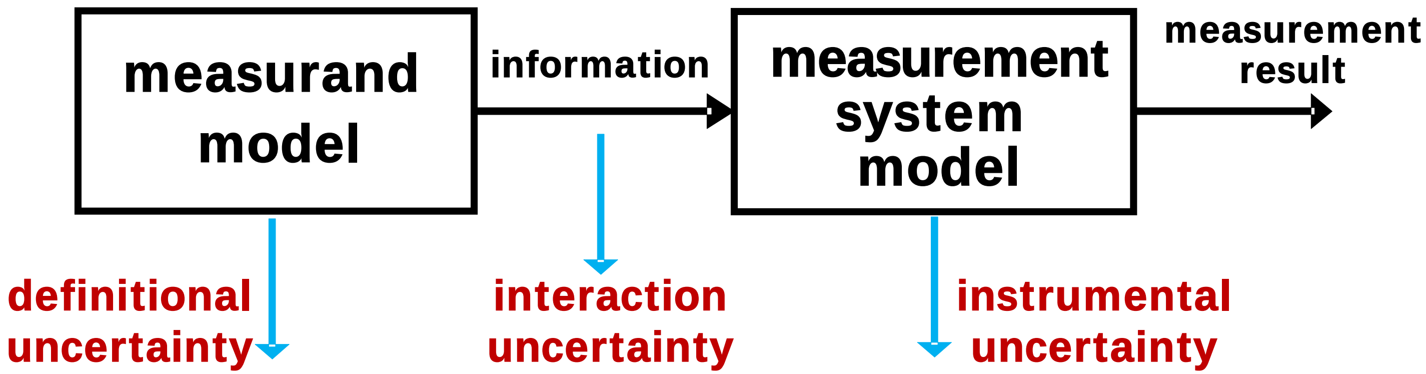
$$\Delta\chi\Delta\rho \geq \frac{\hbar}{2}$$

Measurement uncertaint

UNIVERSITY OF TRENTO - Italy

measurement uncertainty originates during modeling

1) uncertainty sources identification: **three types** are normally present ("**physiological**" to measurement)
2) contributions of uncertainty sources **propagates** to measurement result



"pathological" uncertainty sources can occur due to uncontrolled phenomena; their occurrence must be detected, so that corrective action can be performed

e.g.: reject related data

uncertainty sources must be **controlled during execution**

Context model structure

UNIVERSITY OF TRENTO - Italy

the context model (explicitly or implicitly) contains:

1) the **VARIABLE** that model the **measurand** x
(with definitional uncertainty < target uncertainty $U_{x,UB}$)

2.A) the **VARIABLES** that model the **main influence properties** $\underline{z}$
effect of allowed Δz on the measurement result is comparable or greater than $U_{x,UB}$
need to be compensated

2.B) the **ALLOWED RANGES** $\underline{Z}$ for the main influence properties

2.C) the **RELATIONSHIPS** that describe how $\underline{z}$ affects **raw data** in order to **compensate** their effect on measurement result

3.A) the **VARIABLES** that model the **secondary infl. properties** $\underline{w}$
effect of allowed Δw on the measurement result is sufficiently less than $U_{x,UB}$
not necessarily to compensate

3.B) the **ALLOWED RANGES** $\underline{W}$ for secondary influence properties

Modeling measured result

the context in which measurement is performed is
one of the possible contexts
that can be described by the adopted model

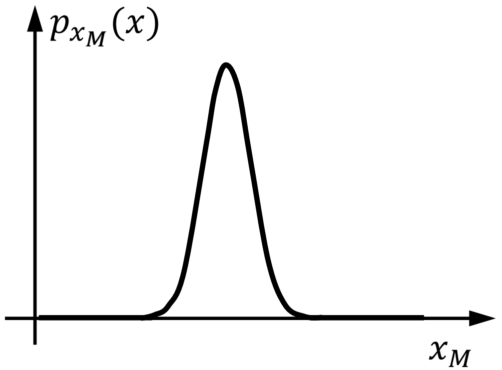
each modeled quantity (i.e.: x, $\underline{z}$, $\underline{w}$)
can assume a **set of values**

quantities are modeled as
random variables
(or **random processes** if depend on time)

described using a
probability density function (pdf)

concept of pdf implies **continuity** of values:

we neglect that measurement values are defined on a **discrete scale** due to **finite resolution** of instruments



possible values of r.v. describing measurement results

Repeated measurement

if **measurement execution** is **repeated over a short-time**, the context can be considered stationary, but actually **influence quantities exhibit little changes**

different measurement results $\{x_{Mn}, n = 1, \dots, N\}$

$$\boldsymbol{x}_M = x_M + \boldsymbol{r}_{x_M}$$

$E[\boldsymbol{x}_M]$

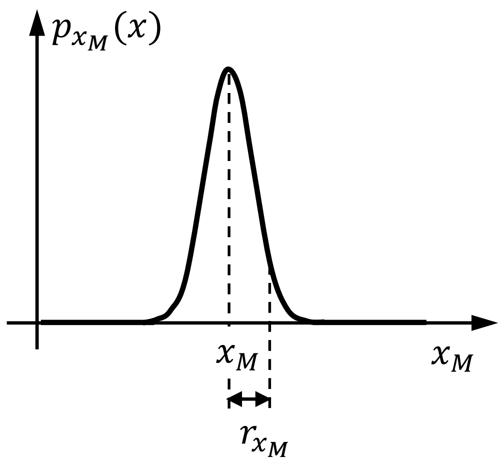
$E[\cdot]$ defined w.r.t. the repeated context

random error

due to the effect of random variations of context quantities

$E[\boldsymbol{r}_{x_M}] = 0$

by definition



Repeated measurement

often **short-time fluctuations of context quantities** $\ll$ magnitude of related **ranges in the context model**

magnitude of measurement variations r_{x_M} $\ll$ **measurement uncertainty**
which is related to context model uncertainty

measurement execution can be ideally
repeated along any context "dimension"

e.g.: simultaneous use of more instruments with equivalent metrological characteristics

often r_{x_M} is only a small fraction of possible deviation from the measurand value

Measurement result

"true" value of the measurand (unknown)

definitional error

$$\boldsymbol{x} = x + \boldsymbol{d}_x$$

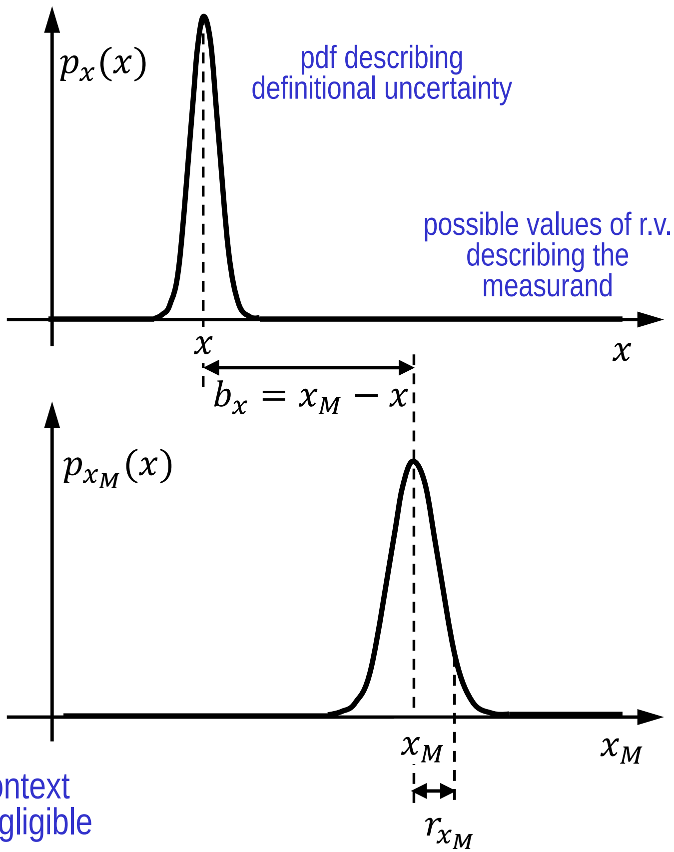
random error

$$\boldsymbol{x}_M = x_M + \boldsymbol{r}_{x_M} = \left(x + b_{x_M}\right) + \boldsymbol{r}_{x_M}$$

systematic error

due to a deviation of measurement context from the reference context

unidentifiable with repeated measurements



$\boldsymbol{p}_x(\cdot)$ **max information** achievable on the measurand for the given measurand model

effect of context assumed negligible

$\boldsymbol{p}_{x_M}(\cdot)$ **information achievable** in the given context

target uncertainty

goal: assure that: $|\boldsymbol{x}_M - x| < U_{x,UB}$ with a given probability

using the **information available** on the context model

Modeling: complexity

context model complexity is related to:

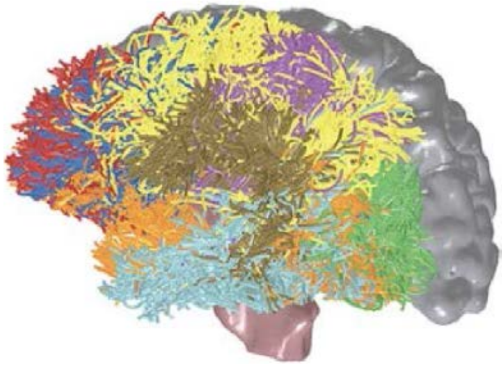
- number of influence properties
- analytical complexity of relationships
- ranges of values allowed for influence properties

if **additional context details** are considered

$\Rightarrow$ model complexity increases

$\Rightarrow$ model uncertainty can decreases

$\Rightarrow$ measurement uncertainty can decreases



Modeling: complexity

measurement cost normally increases with **context model complexity**:

- more sophisticated and expensive instruments
- more complex empirical activities
- higher skilled designers and experimenters

returned **information** can be **too much** to efficiently support decisions

SUITABLE CONTEXT MODEL:

trade-off complexity and resource consumption

its identification can be a very hard task

Design

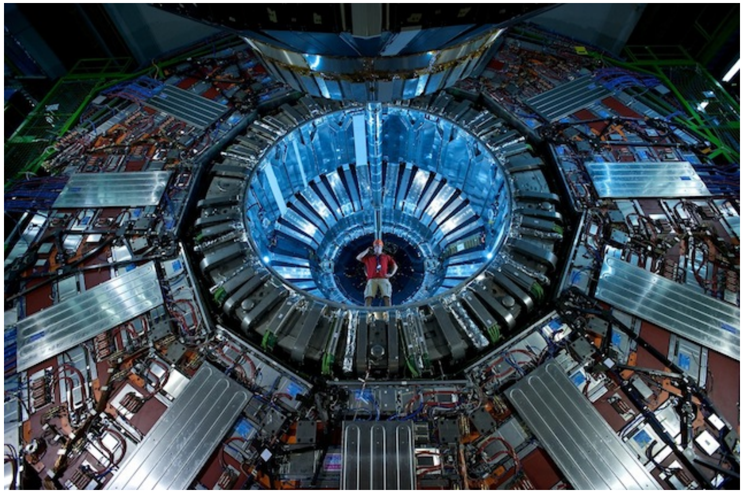
any measurement requires a comparison with a reference (in principle)

often execution is grounded on a **MEASURING SYSTEM** that embeds such a reference

a **set of elements** adapted so as to allow **interaction with the object** under measurement and to **produce**, (as a result of this interaction) a **measurement result** (or an **intermediate output** from which a measurement result can be obtained)

components of measuring systems can be physical sensors, standards, instruments but also software products, questionnaires, ...

a proper **control** of **instrumental uncertainty** ensures measurement **traceability**



the large hadron collider

Design

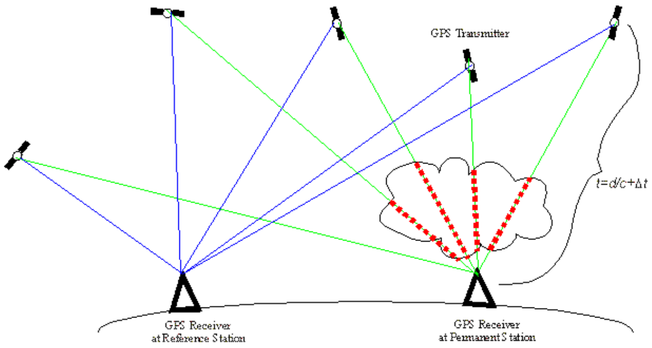
when measuring **PHYSICAL PROPERTIES**, design establishes:

- a **measurement principle** = the adopted physical phenomenon
- a **measurement method** = a generic description of the logical operations required for properly applying the measurement principle
- a **measurement procedure** = a detailed description of all operational steps needed to achieve the measurement result

when dealing with **SOFT MEASUREMENTS**,
no general theory involving the measurand exists

$\Rightarrow$ **no freedom** in choosing among principles, methods, procedures $\Rightarrow$ design establishes:

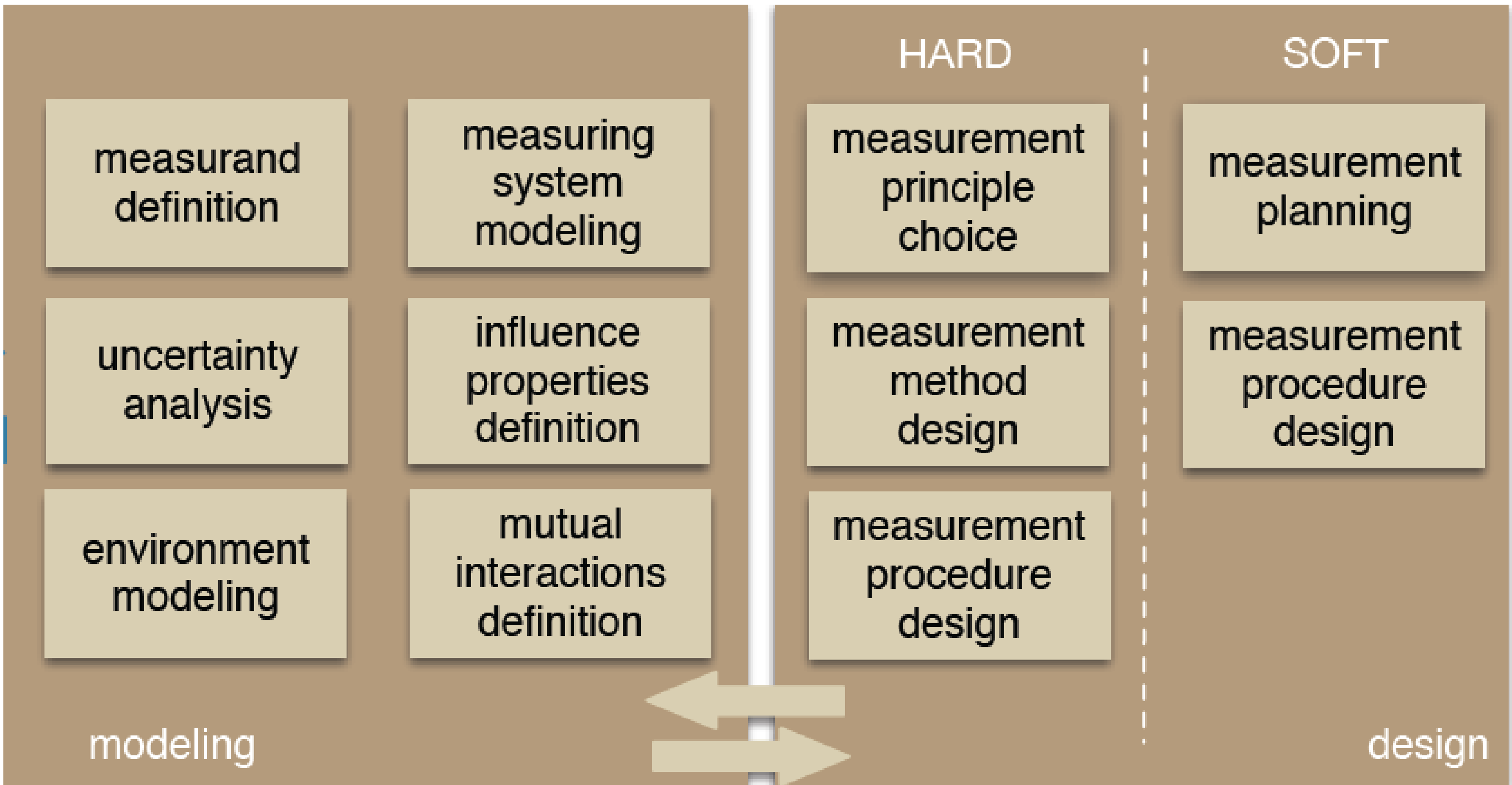
- a **measurement plan** = generic description of the operations performed
- a **measurement procedure**



Modeling - Design

UNIVERSITY OF TRENTO - Italy

summary of the tasks to be performed:

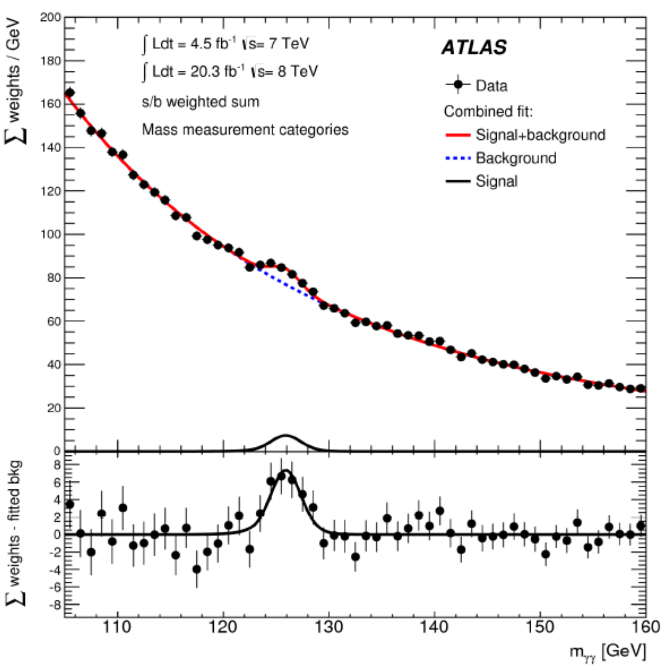


MEASUREMENT: EXECUTION

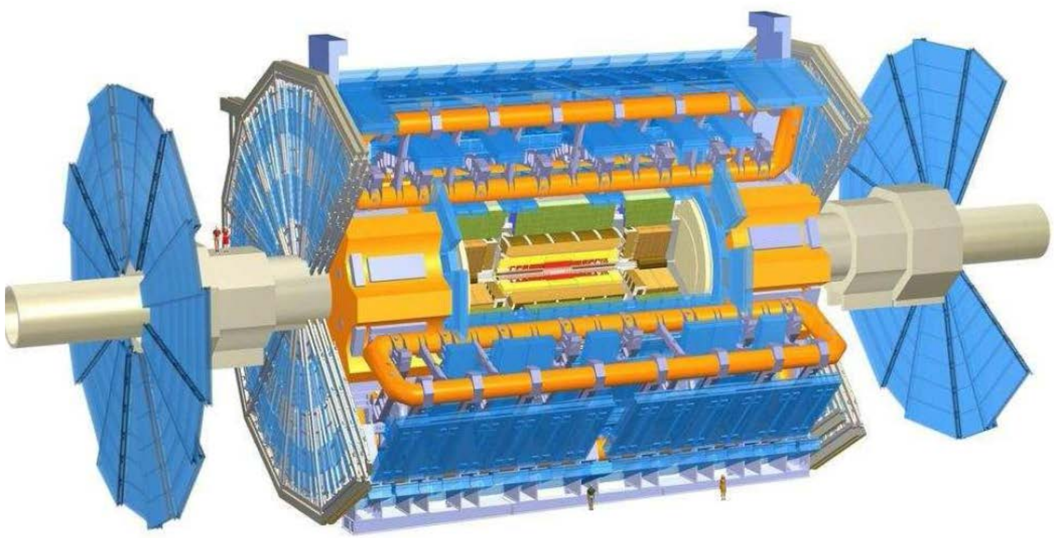
Execution

- the designed **measurement scale** has to be properly constructed
- the **object under measurement** has to be **detected** and sometimes properly **prepared** (e.g. when dealing with chemical quantities)

Higgs detection



ATLAS experiment at LHC



Execution

- designed **measuring system** is properly **assembled**, **calibrated** and **set up** w.r.t. both the object under measurement and environment

  e.g., by suitable connections

- **raw measurement data** about both the measurand and influence properties are **acquired** according to the measurement procedure

- **information** about the measurand is:
  - ✓ **extracted** by **processing raw data** according to the defined context model; measurement **uncertainty** is evaluated
  - ✓ **presented** in a suitable form

Execution: Reporting

UNIVERSITY OF TRENTO - Italy

VIM: **set of values** being attributed to a measurand together with any other **available relevant information**

0) the assumed **measurand definition**

1) a **quantity value** (called the **measured value**)

2) the **amount of information** provided

a single number is not a measurement result!

3) the **measurement scale** used

e.g., a measurement unit

i.e., influence properties and allowed ranges

4) the **limits of validity** of the information provided

5) **subsidiary information** needed/useful for result use and verification

e.g., adopted model for measurement context

Execution: Reporting

UNIVERSITY OF TRENTO - Italy

the **amount of information** about the **measurand** returned by measurement can be reported in **different ways**:

A) explicitly

e.g.: according to GUM, an interval of values with a stated coverage probability

e.g.: $R = 1.234 \pm 0.003\ \Omega$ for $0\ °C \leq T \leq 60\ °C$

B) implicitly, reporting the **significant digits** of measured value

C) not reported (**proxies,** not measurement results!)

it is only ensured that information provided is related to the measurand, without specifying the amount

information should be available to distinguish between B) and C)

Execution

summary of the tasks to be performed:

setup
- object under measurement detection
- scale setup and validation
- measuring system calibration
- measuring system setup

data acquisition
- measurand related data acquisition
- influence properties data acquisition

information extraction and reporting
- data analysis and processing
- value assignment
- uncertainty computation
- subsidiary information reporting

MEASUREMENT: INTERPRETATION

Interpretation

1st aim: establish **measurement validity** through evidences and decision-making

check if the required information needs are satisfied

2nd aim: preserve the acquired knowledge for future use (**lessons learned**)

e.g., new procedures or best practices

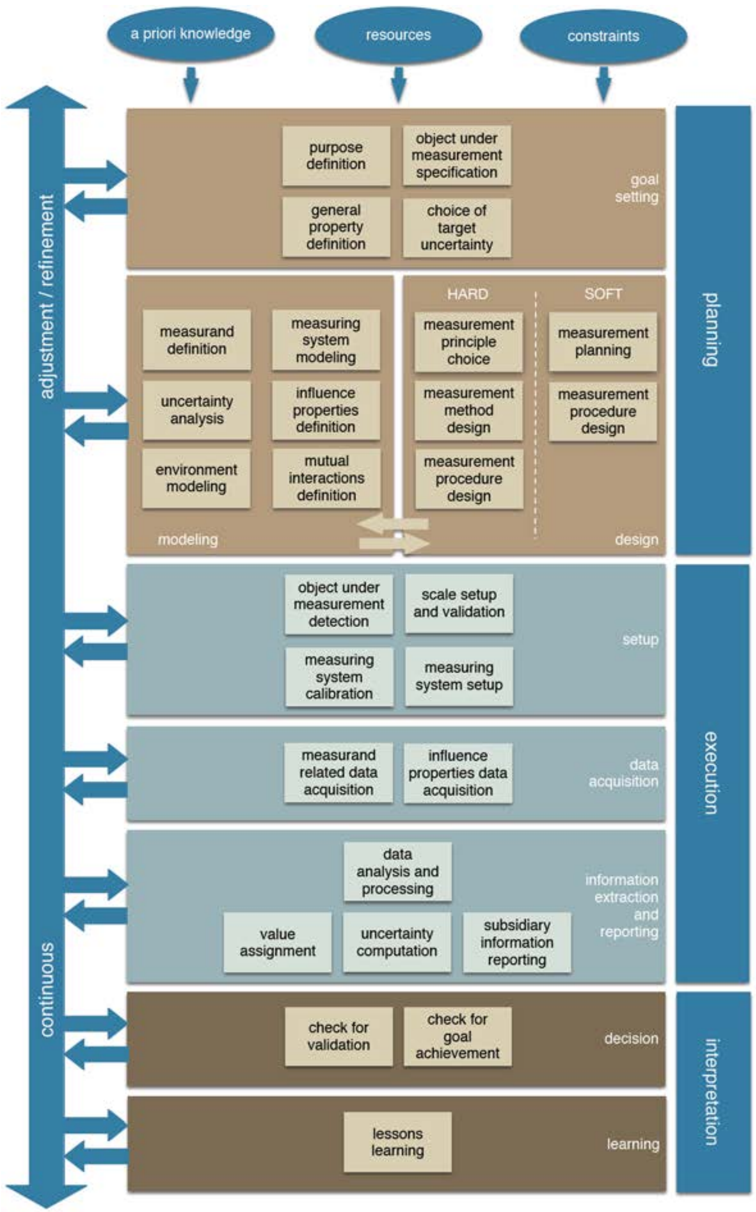
summary of the tasks to be performed:

decision:
- check for validation
- check for goal achievement

learning:
- lessons learning

The whole framework

UNIVERSITY OF TRENTO - Italy



feedbacks aimed to refine previous activities until requirements are met

goal may be reconsidered
(usually as a result of interpretation)

many **activities** can be **omitted**
thanks to
a-priori knowledge and resources

DATA DRIVEN DECISION MAKING (D^3M)

L. Mari, D. Petri, “The Metrological Culture in the Context of Big Data", IEEE Instr. and Meas. Magazine, Oct 2017

D^{3}M structure

according to given criteria

goal: **selection of the best alternative**

Decision | Confidence

Alternatives → **Decision making** ← Procedure

Constraints → ← Acceptable risk level

Information | Quality

guided by a possibly specified **procedure**

subjected to **constraints**

on time, resources, …

supported by **information on the empirical world**

hopefully trustworthy and useful

a lower bound for **decision confidence**

decision confidence is limited by **physiological** and **pathological** **sources of uncertainty**

normally present, hopefully controlled

due to uncontrolled phenomena

D³M: simplest decision

can be completely **automated**

simplest D³M:

single aimed: a single best **choice** exists

fully structured: **procedure** is completely defined

fully informed: information suffices to make **unambiguous decisions** with given level of confidence

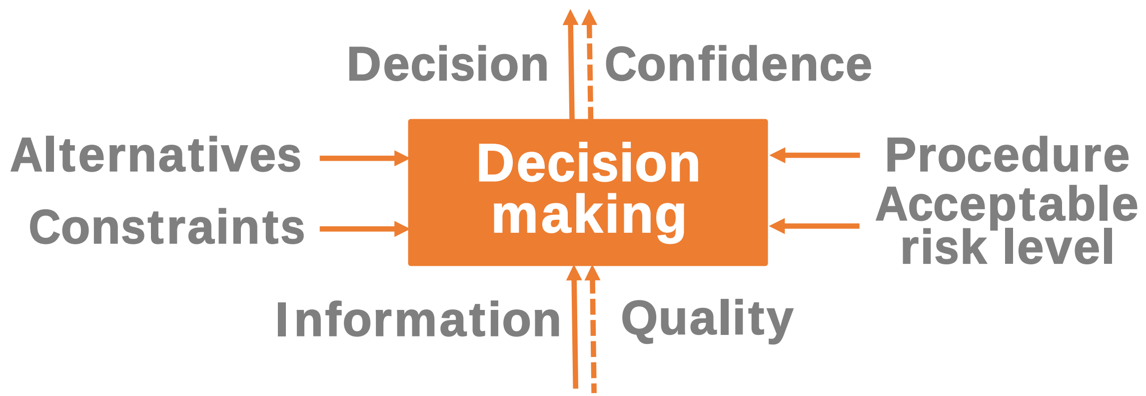
ex.: conformity assessment: measurement result compared with **specification limits** to decide if the measurand fulfills given **requirements**

unsatisfactory | **satisfactory** | rating levels

minimally acceptable | **target range** | **exceeds requirements**

worst case | current level | planned level | measurement value

D^{3}M: formal decision rule



formal (**probabilistic**) approach:

define an **utility function**

derive an **optimal decision rule**

knowledge of information **uncertainty** enables the evaluation of the **risk** (= probability) of wrong decision

D^{3}M: complex decisions

UNIVERSITY OF TRENTO - Italy

information is often:

acquired using **non-validated procedures**

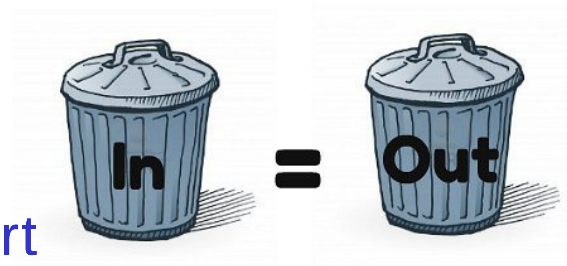
processed using **non-validated models**

possible **unidentified uncertainty** sources (**poor information**)

trustworthiness and usefulness not assessable

GIGO principle

garbage in – garbage out



information expected to support **wrong decisions**

non-single aimed, semi-structured, partially-informed **decisions** performed **automatically** through algorithms

potential **misinterpretation** or **misuse of information**

SUBJECTIVE JUDGMENT

always necessary to manage the **risk of wrong decision**

VIM: body of knowledge aimed at identifying, quantifying, assessing uncertainty sources

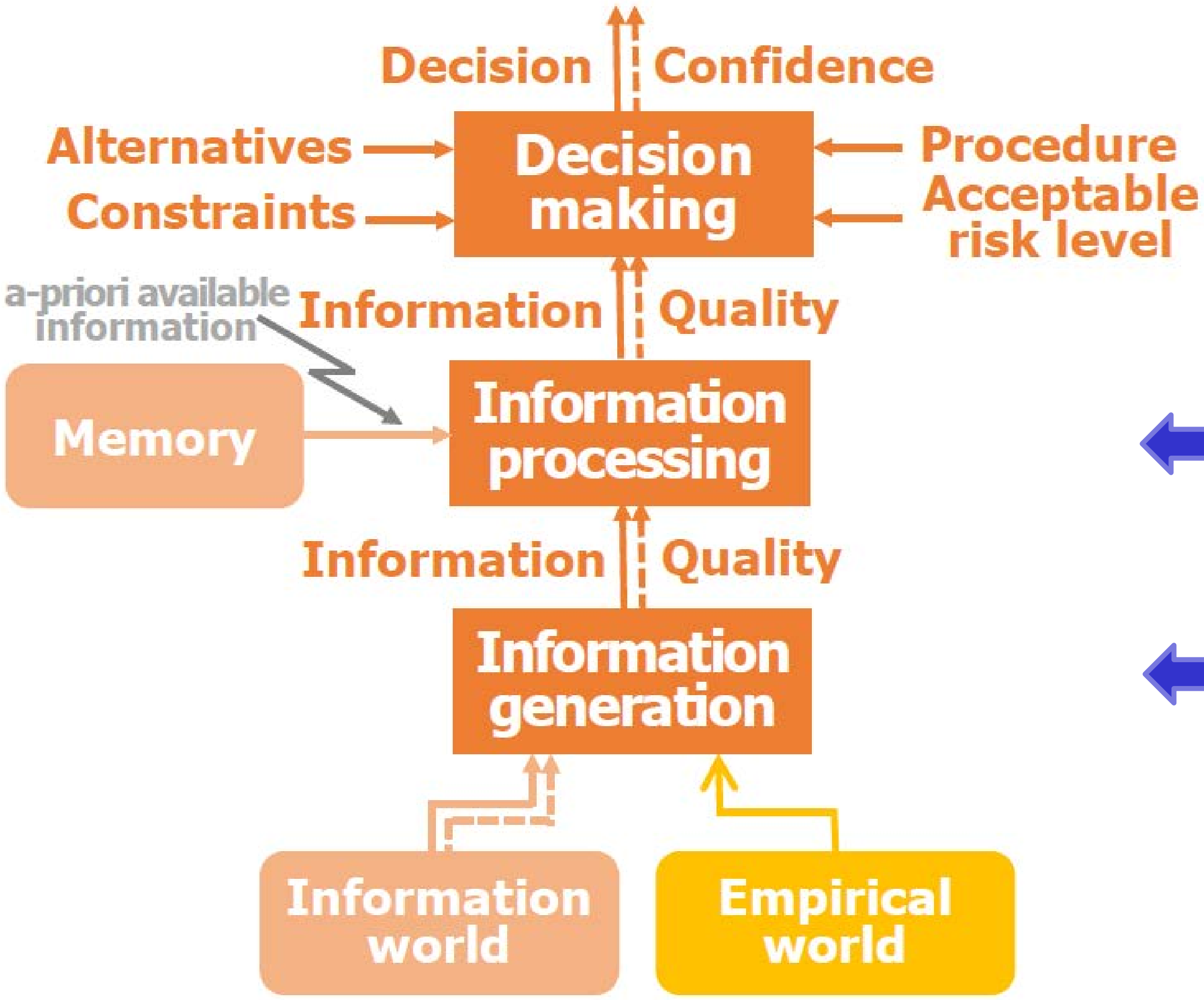
METROLOGY:

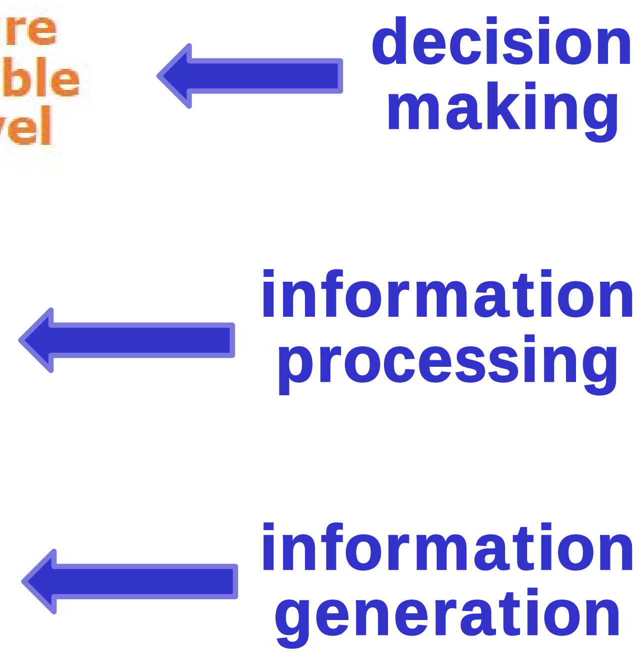
a **science of information quality**

is crucial to support effective **decision confidence** management

D³M: Metrological Model

three-level hierarchically structured process:





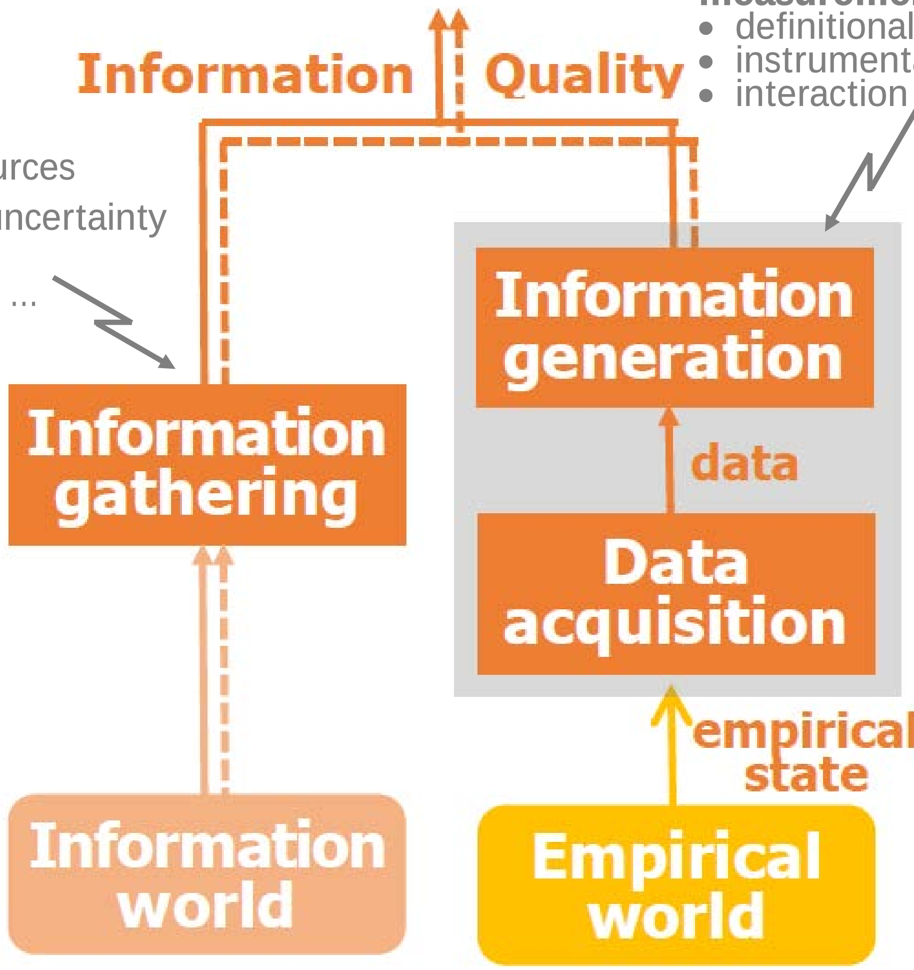
D^{3}M: Information Gener.

physiological sources

- gathering software uncertainty
- repository reliability
- repository coverage, ...

gathering of information from repositories

e.g.: from internet



physiological sources

measurement uncertainty

- definitional
- instrumental
- interaction

pathological sources

- erroneous data
- missing data, ...

acquisition of information from empirical world

e.g.: by sensors

information extraction by processing, using a-priori knowledge

physiological sources
processing uncertainty
- processing modeling uncertainty
- statistical sampling (limited coverage), …

wrong processing model

pathological sources
- statistical correlation misinterpreted as causation
- processing meaningfulness, …

e.g.: non-admissible arithmetic operations
($F = m \cdot a$, $F = m + a$)

D^{3}M: Decision

UNIVERSITY OF TRENTO - Italy

misuse
of information

pathological sources

- wrong decision procedure
- reasoning biases (fallacy)

e.g., prejudices

physiological sources

decision uncertainty

- vague alternatives
- ambiguous procedure
- conflicting constraints
- reasoning biases
- information relevance
- inferential uncertainty

when conclusion achieved
through **logical reasoning**

ex. in forensic:
guilt/innocence based on evidences
(e.g. fingerprint recognition, DNA and blood analysis)

D³M: Subjective judgment

UNIVERSITY OF TRENTO - Italy

to draw meaningful conclusions:

non-single aimed, **semi-structured**, **partially-informed decisions**

always need **subjective judgment** based on professional experience, a priori knowledge, critical analysis of available information, intellectual honesty

metrological culture helps to **avoid wrong decisions** due to misinterpretation or misuse of information available