Metrics or Measures?
Which one would you better trust?

Alain Abran,
École de technologie supérieure,
Université du Québec, Montreal, Canada

Universidad del Pais Vasco, San Sebastian,
June 3, 2005
Objectives of the Presentation

- What do we know about measures?
- What do we know about Software Metrics?
- Software engineering and Measurements?
- Identification of the gaps where further research on software measurement is required
List of topics

- Motivations and Objectives
- Software Metrics
- Models for Designing Measures
- What is Generally Accepted?
- International Standards
- Conclusions and Future Work
Motivations

- Measurement is fundamental in:
  - In day to day life
  - In business
  - In sciences and engineering

- Measurement **instruments** are key to all

- What is the status (maturity) of measurement in software?
Why do you measure?

- To understand:
  - To know - to learn

- To plan:
  - To set targets

- To control:
  - To monitor - To compare
  - To make adjustments
Measurements

Everywhere in sciences and engineering

- Where do they come from?
- How do you know that our measures are OK?
List of topics

- Motivations and Objectives
- Software Metrics
- Models for Designing Measures
- What is Generally Accepted?
- International Standards
- Conclusions and Future Work
Software Metrics

The dominant view in software measurement =

« Software Metrics »
Software Metrics

- Lines of code
- McCabe
- Function Points
- Halstead
- COCOMO
- Software quality metrics: Hundreds +
- Software complexity metrics: Hundreds +
- OO metrics: Hundred ++
- Estimates and Estimations models….
Quality of ‘metrics’

Do you get:
- Reproducible results?
- Repeatable results?
- Accurate results?
- Results that you can trust?

Who design these metrics?
- How qualified are they?
- Who verifies their metrics proposals?
Software Metrics

How are they designed?
- Anything that can be ‘counted’

- How are they defined:
  - Often labels
  - An algorithm

- How do we know if they are valid?
  - Sometimes validation based on measurement theory
  - Sometimes on claims of ‘usefullness’
List of topics

- Motivations and Objectives
- Software Metrics
- Models for Designing Measures
- What is Generally Accepted?
- International Standards
- Conclusions and Future Work
Analytical Tools of Measurement

- **Measurement Process model**
  - Abran & Jacquet

- **Metrology:**
  - ISO International Vocabulary in Metrology
Measurement Process Model

High-level measurement process model

- **Step 1**: Design of the measurement method
- **Step 2**: Application of measurement method rules
- **Step 3**: Measurement result analysis
- **Step 4**: Application of measurement method rules

© 2005 Alain Abran - Laboratoire de Recherche en Génie Logiciel
The long-standing international consensus on measurement terminology

The basis of the International System (IS) of measurements

National Metrology Agencies
  - The legal framework for weights and measures in industrialized countries
Metrology Vocabulary

- Six categories of terms
  - + 120 terms
    - In increasing order of complexity!

- Most challenging to grasp relationships across terms, understand, and remember!
Metrology Model

Measurement foundations

- **Metrology**
  - Science of Measurement
- **Principle of Measurement**
  - Scientific Basis of a Measurement
- **Method of Measurement**
  - Logical Sequence of Operations
- **Measurement**
  - Set of Operations

© 2005 Alain Abran - Laboratoire de Recherche en Génie Logiciel
Metrology Model

High-level model of the ISO Vocabulary

Measuring instruments

Input

Measurements

Measurement Results

Quantities and units

Etalons

Characteristics of measuring instruments
Metrology Model

Measurement Procedure

Measurand \rightarrow Measurement Signal \rightarrow Transformed Value \rightarrow Result of a Measurement

Operator \rightarrow Measurement Method \rightarrow Influence Quantity

© 2005 Alain Abran - Laboratoire de Recherche en Génie Logiciel
## Metrology Model

### Measurement results

<table>
<thead>
<tr>
<th>Types of measurement results</th>
<th>Modes of verification of measurement results</th>
<th>Uncertainty of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indication (of a measuring instrument)</td>
<td>Accuracy of measurement</td>
<td>Experimental standard deviation</td>
</tr>
<tr>
<td>Uncorrected result</td>
<td>Repeatability (of results of measurements)</td>
<td>Error (of measurement)</td>
</tr>
<tr>
<td>Corrected result</td>
<td>Reproducibility (of results of measurements)</td>
<td>Deviation</td>
</tr>
</tbody>
</table>

- Random error
- Systematic error
- Correction
- Correction factor
Mapping between models

Alignment of metrology concepts with the measurement process model

<table>
<thead>
<tr>
<th>Measurement process model</th>
<th>Design of measurement methods</th>
<th>Application of measurement method rules</th>
<th>Measurement results analysis</th>
<th>Exploitation of measurement results</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO metrology model</td>
<td>Quantities and units</td>
<td>Measuring instruments</td>
<td>Measurement results</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Characteristics of measuring instruments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2005 Alain Abran - Laboratoire de Recherche en Génie Logiciel
List of topics

- Motivations and Objectives
- Software Metrics
- Models for Designing Measures
- What is Generally Accepted?
- International Standards
- Conclusions and Future Work
Generally Accepted:

What applies most of the time, to most projects, and which value has been validated by the community of peers

Software Engineering Body of Knowledge - SWEBOK
# Measurement within SWEBOK

<table>
<thead>
<tr>
<th>SWEBOK KA</th>
<th>Topics</th>
<th>Step 1 Design</th>
<th>Step 2 Measuring</th>
<th>Step 3 Results</th>
<th>Step 4 Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software engineering requirements</td>
<td>Process support and management</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Requirements negotiation</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Document quality</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Acceptance tests</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Requirements tracing</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Software engineering design</td>
<td>Measures</td>
<td></td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Software engineering testing</td>
<td>Evaluation of the program under test</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Evaluation of the tests performed</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Software engineering maintenance</td>
<td>Software Maintenance Measurement</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

© 2005 Alain Abran - Laboratoire de Recherche en Génie Logiciel
# Measurement within SWEBOK

<table>
<thead>
<tr>
<th>SWEBOK KA</th>
<th>Topics</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Software configuration management</strong></td>
<td>Surveillance of software configuration management</td>
<td></td>
<td></td>
<td></td>
<td>☒</td>
</tr>
<tr>
<td><strong>Software engineering management</strong></td>
<td>Goals</td>
<td></td>
<td></td>
<td>☒</td>
<td>☒</td>
</tr>
<tr>
<td></td>
<td>Measurement selection</td>
<td></td>
<td></td>
<td>☒</td>
<td>☒</td>
</tr>
<tr>
<td></td>
<td>Measuring software and its development</td>
<td></td>
<td></td>
<td>☒</td>
<td>☒</td>
</tr>
<tr>
<td></td>
<td>Collection of data</td>
<td></td>
<td></td>
<td>☒</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Software Measurement Models</td>
<td></td>
<td></td>
<td></td>
<td>☒</td>
</tr>
<tr>
<td><strong>Software engineering process</strong></td>
<td>Methodology in process measurement</td>
<td></td>
<td></td>
<td>☒</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Process Measurement Paradigms</td>
<td></td>
<td></td>
<td></td>
<td>☒</td>
</tr>
</tbody>
</table>

© 2005 Alain Abran - Laboratoire de Recherche en Génie Logiciel
# Measurement within SWEBOK

<table>
<thead>
<tr>
<th>SWEBOK KA</th>
<th>Topics</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software engineering quality</td>
<td>Measuring the value of quality</td>
<td></td>
<td></td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td>Fundamentals of Measurement</td>
<td></td>
<td>✗</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Measures</td>
<td></td>
<td></td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td>Measurement analysis techniques</td>
<td></td>
<td></td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Defect characterization</td>
<td></td>
<td></td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td>Additional Uses of SQA and V&amp;V data</td>
<td></td>
<td></td>
<td></td>
<td>✗</td>
</tr>
</tbody>
</table>
Generally accepted knowledge about software measurement?

Strong recognition of benefits:
- to understand, plan, monitor and control
- Foundations = ??
- And little metrology strengths
List of topics

- Motivations and Objectives
- Software Metrics
- Models for Designing Measures
- What is Generally Accepted?
- International Standards in Software Measurement
- Conclusions and Future Work
International Standards

Currently, only for:

- Software Products Quality
- Software Functional Size
How to Measure Software Quality?

ISO 9126 on Software **Products** Quality

- Part 1: Quality Models and Definitions

- Parts 2 to 4: **Technical Reports**

- Over + 120 Metrics!

- With little about:
  - Measurement method for each (labels & algorithms)
  - Validity & Quality of these ‘metrics’??

- Then (if used in a non consistent manner), how do figure out how measurement results compare across contexts, across time, and across measurers?
Software Functional Size

How do you measure software size?

⊙ The technical size = ?

⊙ The functional size = ?
Functional Size

A unique set of measures in software engineering:

- Designed in the late 1970’s:
  - By Albrecht, from IBM, using 24 MIS projects

- Published in the early 1980’s

- User group in the mid 1980’s
  - Measurement Manual
  - Training & Certification
Functional Size

Does it withstand the test of time?

- Still in use and referenced

But

- The basic method has not evolved significantly since the early 90’s
- Software has changed considerably
- Outside of MIS domain = ?
- In the early 90’s: + 30 variations to tackle weaknesses
Functional Size

Innovation in software measurement: Standardization through ISO

First a meta-standard to layout the ground rules about functional size measurement: ISO 14143

- Part 1 = Definitions of Key Concepts
- Part 2 = Conformity Assessment
- Part 3 = Verification Guide
- Part 4 = Set of References
- Part 5 = Functional Domains
Four specific methods approved by ISO

- ISO 19761: COSMIC-FFP
- ISO 20926: IFPUG
- ISO 20968: MKII
- ISO 24570: NESMA

Will they withstand the test of time as measurement methods?
Will a consensus emerge?
Do they meet ‘metrology’ criteria?
Metrology Model

High-level model of the ISO Vocabulary

- Measuring instruments
- Input
- Measurements
- Quantities and units
- Etalons
- Characteristics of measuring instruments
- Measurement Results
## Metrology Model

### Measurement results

<table>
<thead>
<tr>
<th>Types of measurement results</th>
<th>Modes of verification of measurement results</th>
<th>Uncertainty of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indication (of a measuring instrument)</td>
<td>Accuracy of measurement</td>
<td>Experimental standard deviation</td>
</tr>
<tr>
<td>Uncorrected result</td>
<td>Repeatability (of results of measurements)</td>
<td>Error (of measurement)</td>
</tr>
<tr>
<td>Corrected result</td>
<td>Reproducibility (of results of measurements)</td>
<td>Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relative error</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Random error</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Systematic error</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correction factor</td>
</tr>
</tbody>
</table>
List of topics

- Motivations and Objectives
- Software Metrics
- Models for Analytical Tools of Software Measurement
- What is Generally Accepted?
- International Standards

Conclusions and Future Work
Summary & Conclusion

Generally accepted knowledge about Measurement in software:

- Extensive set of references on the use of measurement results in assessment and predictive models.

- Little discussion on:
  - Quality of measurement results
  - Quality of measuring instruments

- Limited knowledge on the design of measurement methods
Conclusions - Next

- The field of « software metrics » is going where after 30 years of research?
  - Why not learn from the masters in measurement?

- Most majors R & D contributions still waiting for you!

- Standardization is critical for measurement!
Questions