



# Quality of Benchmarking Data

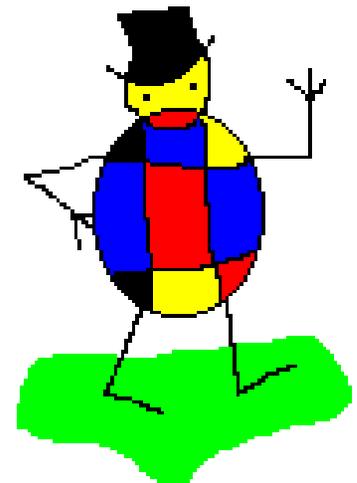
\*\*\* *FiSMA approach* \*\*\*

Pekka Forselius, STTF Oy,  
Managing Director, Project Expert



# Contents

- Benchmarking questions
- Other good questions
- Available benchmarking databases
- Data quality viewpoints and definitions
- Rating rules <sup>1)</sup> for internal data quality (IDQ)
- Rating process <sup>1)</sup> for evaluating IDQ
- Experiences of evaluating IDQ
- Rating rules <sup>1)</sup> for data quality in use
- Data quality assurance concept <sup>1)</sup>
- Conclusions (expected benchmarking results)

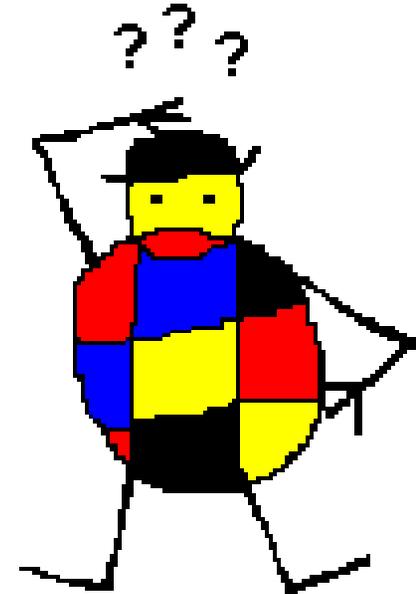


1) FiSMA 2005



# Benchmarking questions

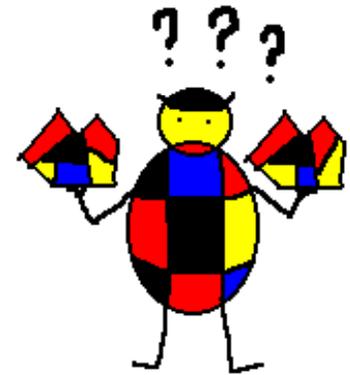
- Are we productive?
  - Are we somehow special, or strange?
  - What should we improve?
  - How much could/should we improve?
- 
- Typically companies at CMMI levels 2+ and 3 start to ask these questions.





# Other relevant questions

- What variables, attributes and other pieces of information are most useful to collect?
- Are there differences in productivity between business sectors?
- Or between companies?
- Or between project/development types?
- Or between development tools?
- Or between hardware platform types?
- Are there differences in productivity between application types?





# Potential sources of answers

- Compass Analysis Development
- FiSMA Experience® database, 700 + projects
- Gartner Group
- ISBSG repository, 3000 + projects in rel9, 2004
- ESA/RISE dataset
- MeLLOW project, 30 + datasets
- ...





# But... the Quality of data?

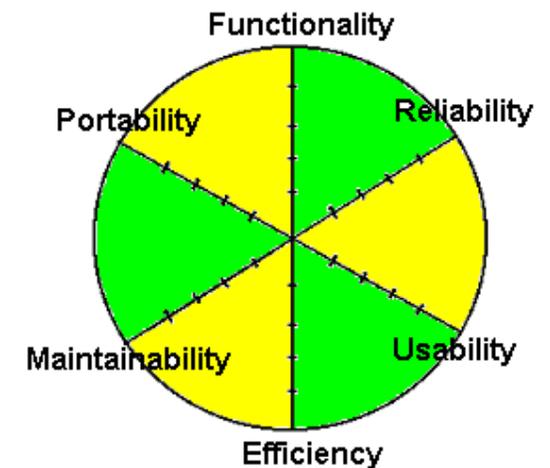
- Missing information
- Missing definitions
- Strange values
- Outliers
- Technical data entry errors
- ...





# Data Quality Issues

- **Internal Data Quality:** capability of a set of static attributes of data to satisfy stated and implied needs when the data are used under specified conditions.
- **External Data Quality:** capability of data to enable the behaviour of a system to satisfy stated and implied needs when the system is used under specified conditions.
- **Data Quality in Use:** capability of the data to enable specific users to achieve specific goals with timeless, amount of information, relevancy, credibility and understandability in specific contexts of use.



ISO/IEC 25000 family



# FiSMA Rationale

- Good internal quality →
- Good external quality →
- Good quality in use →
- Right answers to top management questions →
- Effective process improvement activities





# Internal data quality

## Experience® rating for project data records

- **AAA** Highest quality 90+
- **AA** Excellent 80-89
- **A** Very good 70-79
- **B** Good 60-69
- **C** Satisfactory 50-59
- **D** Acceptable 40-49
- **X** Rejected -39

- New rating scale published in 2005, "top management language"

Requirement:	Max. weight	Source of information	Evaluation rules
Status = actual	2	Cost summary report	Yes or no, -60 if not actual
Existence of basic classifiers	8	Cost summary report	Yes or no, -60 if any missing
Start date	2	Cost summary report	Yes or no, -60 if missing
End date	2	Cost summary report	Yes or no, -60 if missing
Supplier effort	6	Cost summary report	Yes or no, -60 if missing
Size of software (fp)	8	Cost summary report	Yes or no, -60 if missing
Situation coefficient	8	Cost summary report	Existence = 4 points, difference from 1.0 = 8 points, -60 if missing
Completeness of effort data	14	Cost summary report and Final review/Project closure/Sub-windows	Effort and duration of each relevant phase, relevance varying by development type: New development: 2/phase effort and duration, max 10 points. Enhancement, integration or conversion: 4/phase effort and duration, max 10 points. Maintenance rede: 2/phase effort and duration, max 10 points. Start and end dates should be conformant with the corresponding dates of phases. Numbers of staff by phase +0-2 points. Reported accuracy of effort +0-2 points.
Project extra classifiers	6	Classifiers report	Several different classifiers needed to get maximum, if only few, then 1 point per each.
Software extra classifiers	6	Classifiers report	If several applications and several classifiers per each, or only one application with several classifiers for maximum. If only few classifiers reported, then 1 point per each.
Reliability of size measurement	14	Project's functions report	If 10 % of the function points are based on unjustified user requirements => -1 point, if 20 % => -2 points and so on. If all the functions unidentified and belong to only one type => 0 points. If the names of the functions all default => -2, if complexity details all default => -4 p.
Reliability of situation analysis	14	Situation analysis report	If the situation analysis method is not compatible with the development type, - 7 points. Straight row in the middle = 0p, each difference +20, maximum 14 points.
Existence of reuse data	2	Reuse summary report	Yes or no, if only a couple of functions, then 1 point. For maintenance always = 2 points.
Existence of risk factor data	2	Risk analysis report	Yes or no, if only 1 or 2 factors, then 1point.
Precision of basic classifiers	4	Cost summary report	If "multi platform" or "nGL-default development language", -2 points from each.
Number of applications	2	Cost summary report	If = 1, then 2 points, if = 2-4 and unanimous development environment, then 1 point.
Size measurement approach and measurement unit	2	Project's functions report	If backfiling method used, then 0 points, if IFPLG or Mill as the size measurement unit, then 1 point, else 2 points.
Identification of project management	2	Cost summary report	1 point from both the manager and the leader.
<b>SUMMARY:</b>	100		<b>CATEGORY (AAA,AA ...,D tai X)</b>



# The most important data items

- Functional size of software, total size and accurate function mix
- Reliable effort data, covering all relevant phases of development life-cycle
- All 21 productivity factors measured, precise knowledge of development circumstances
- Basic classifiers precisely recorded
- 7 mandatory + 11 complementary attributes or data groups evaluated



# Internal data QA process

- All data collected using Experience® Pro tool
- Standard reports used for quality assurance
- QA made by the Experience® repository manager
- Feedback to data contributors sent
- The QA process was evaluated by 6 FiSMA Scope Managers in May 2005, excellent results





# Internal data quality of current Experience® database

- A sample of 278 projects
- Majority of the project records is good or very high quality data
- No more X-quality since 2001

AAA	13
AA	93
A	109
B	28
C	7
D	1
X	27



# From internal data quality to data quality in use

- Company data vs. common data
- applicability and similarity of sample subset
- # of observations
- age of data in subset
- behavior and internal variation in subset

"This highlights the need for companies to establish their own sw development database!"

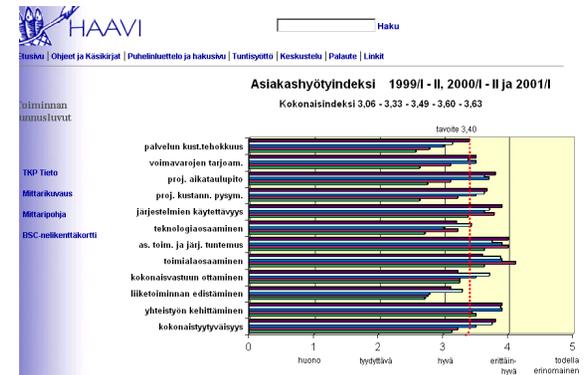


Katrina Maxwell,  
Berlin 27 May 1997



# From excellent data quality in use to excellence in business

- If you want/have to improve from CMMI level 3 to level 4
- Regular measurement and use of high quality benchmarking data are essential to reach level 4
- Quantitatively managed = Predictable → Increased customer satisfaction → Success in business





# Conclusions

- High quality benchmarking data gives right answers to your questions, bad data may lead to wrong decisions.
- Focus on 'good enough quality': don't try to collect everything just in case.
- Remember that you must be better than your competitors, but not too much better
- The best in class must benchmark mainly against themselves, but never stop to benchmark!



# More information

- pekka.forselius@sttf.fi
- www.sttf.fi
- www.fisma.fi
- www.isbsg.org: (NASSCOM is India member)
  - The Software Metrics Compendium, ISBSG, 2003
  - Practical Project Estimation, ISBSG, 2005
  - + other ISBSG products

