Important concepts in CMMI and what is difficult to understand

Dr. Katalin Balla

Zádor Dániel Kelemen

Budapest University of Technology and Economics

SQI Hungarian Software Quality Consulting Institute Ltd.

JUNE 7-9, 2011

Dublin, Ireland

SEI Partner

Acknowledgement

- Preparing and delivering this presentation was sponsored by
  - BUTE Centre of Information Technology
    - BME Információtechnológiai Innovációs és Tudásközpont, http://www.ik.bme.hu/index_eng.php

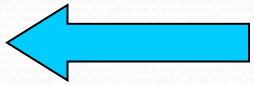
within project SZOMINo8

TECH_08_A2-SZOMINo8 - Software quality assurance service-package for open document format applications

AND

within project „Q-Works" - GOP-1.1.1-08/1-2008-0023 – Automated workflow system for quality assurance, based on CMMI

Contents

- Introduction 
  - Ways of understanding CMMI
- Text mining in CMMI
  - What we did
  - Results of text mining
- Difficulties in applying CMMI concepts
  - What people frequently misunderstand
    - Based on work with companies implementing CMMI

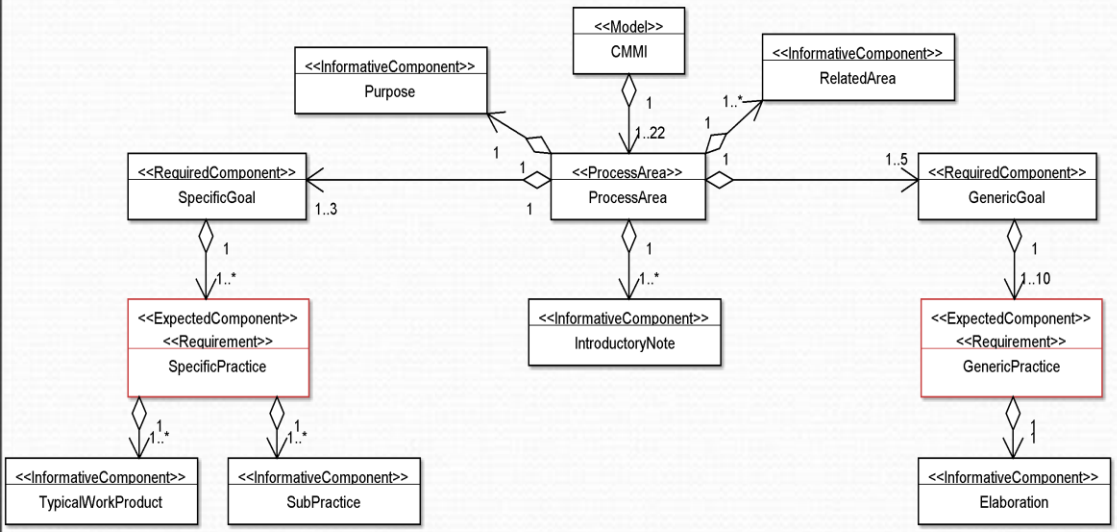
Introduction -1

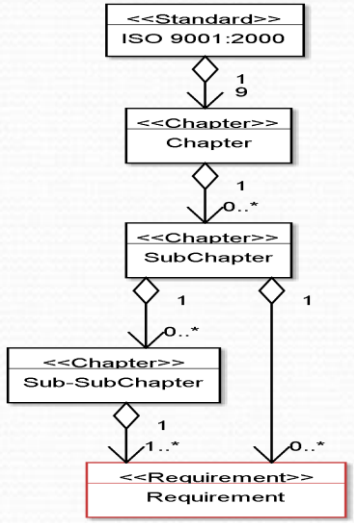
- CMMI: huge amount of text
  - Understandable by
    - Reading it / interpreting it , using experience and knowledge related to SPI, concepts of quality, other models and standards
    - Text mining
    - …

Introduction -2

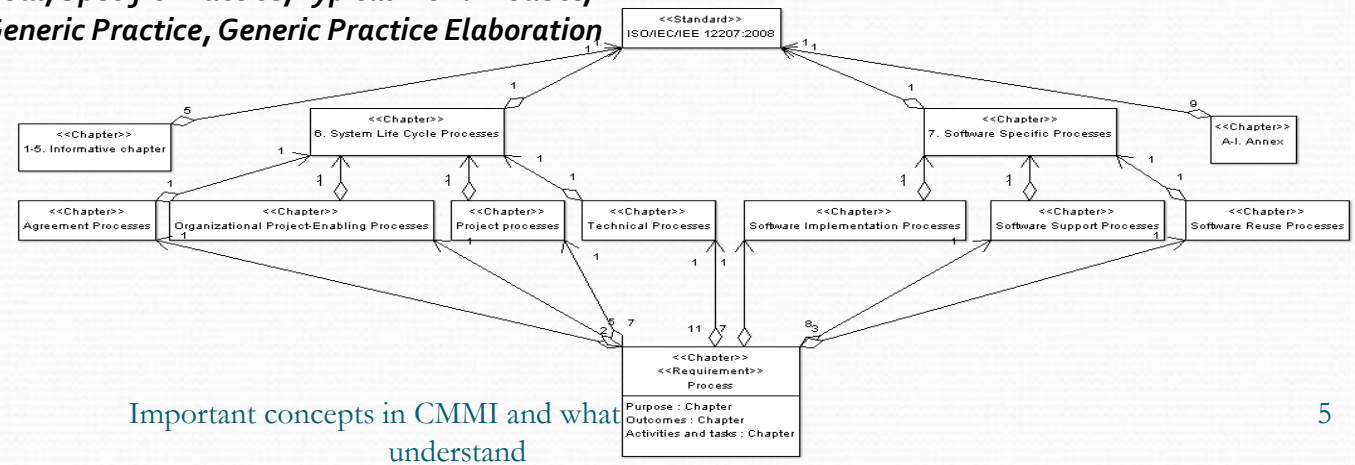
Understanding CMMI

- By reading it / interpreting it , using experience and ~~knowledge~~ related to SPI, concepts of quality, other models and standards
  - PhD research of ***Zádor Dániel Kelemen.*** Poster presentation on SEPG 2009, Prague.

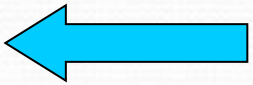




Elements: ***Process area, Specific Goal, Specific Practice, Typical Work Product, Subpractice, Generic Goal, Generic Practice, Generic Practice Elaboration***



Contents

- Introduction
  - Ways of understanding CMMI
- Text mining in CMMI 
  - What we did
  - Results of text mining
- Difficulties in applying CMMI concepts
  - What people frequently misunderstand
    - Based on work with companies implementing CMMI

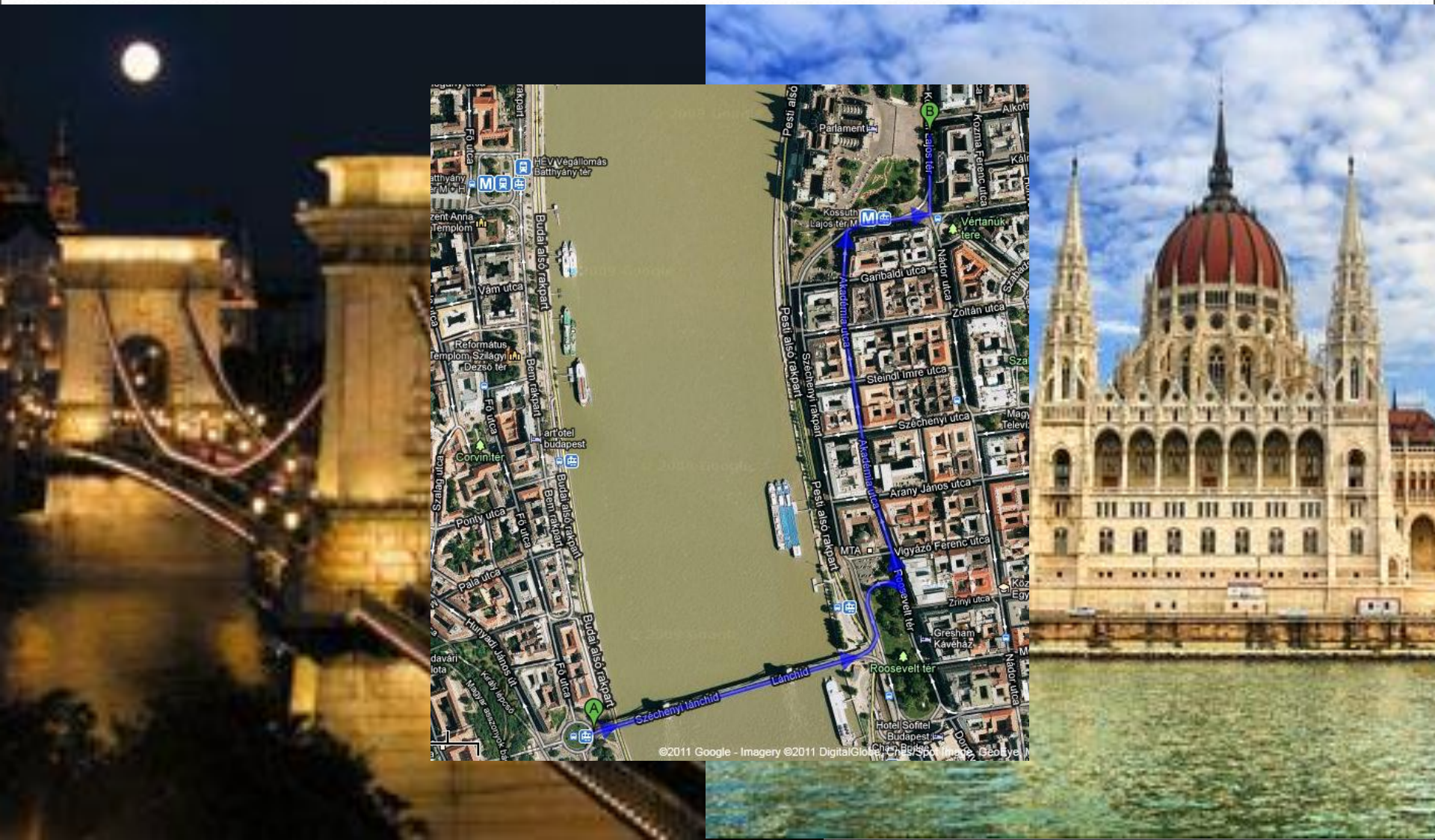
Understanding the most important concepts of CMMI

- "In God we trust; all others must bring data." – W. Edwards Deming.
  - We got data from SEI: ~3*500 pages
  - Somehow we wanted to process it
  - Tools used:
    - open source text mining tool: RapidMiner
    - Excel,
    - MySQL,
    - …and a friend, Imteaz for manual verification

BME IK Informatikai Központ

SQI HUNGARIAN SOFTWARE QUALITY CONSULTING INSTITUTE LTD.

The most important concept



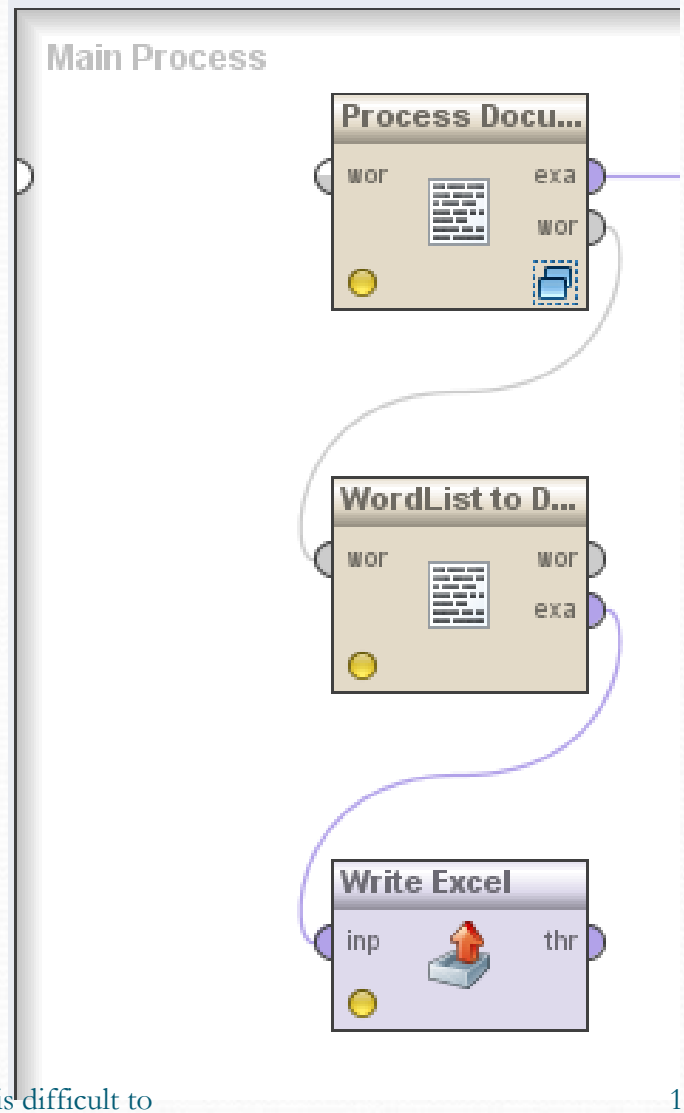
The most important concept

- All occurrences printed would make a 1627 meters long line
- 20 minutes walking distance (Google maps)
- More than
  - walking from Clark Adams square - center of Hungary
  - passing the Danube Europe's second largest river - on Chain bridge
  - walking till the Holy Crown of Hungary - which can be found inside the Hungarian Parliament

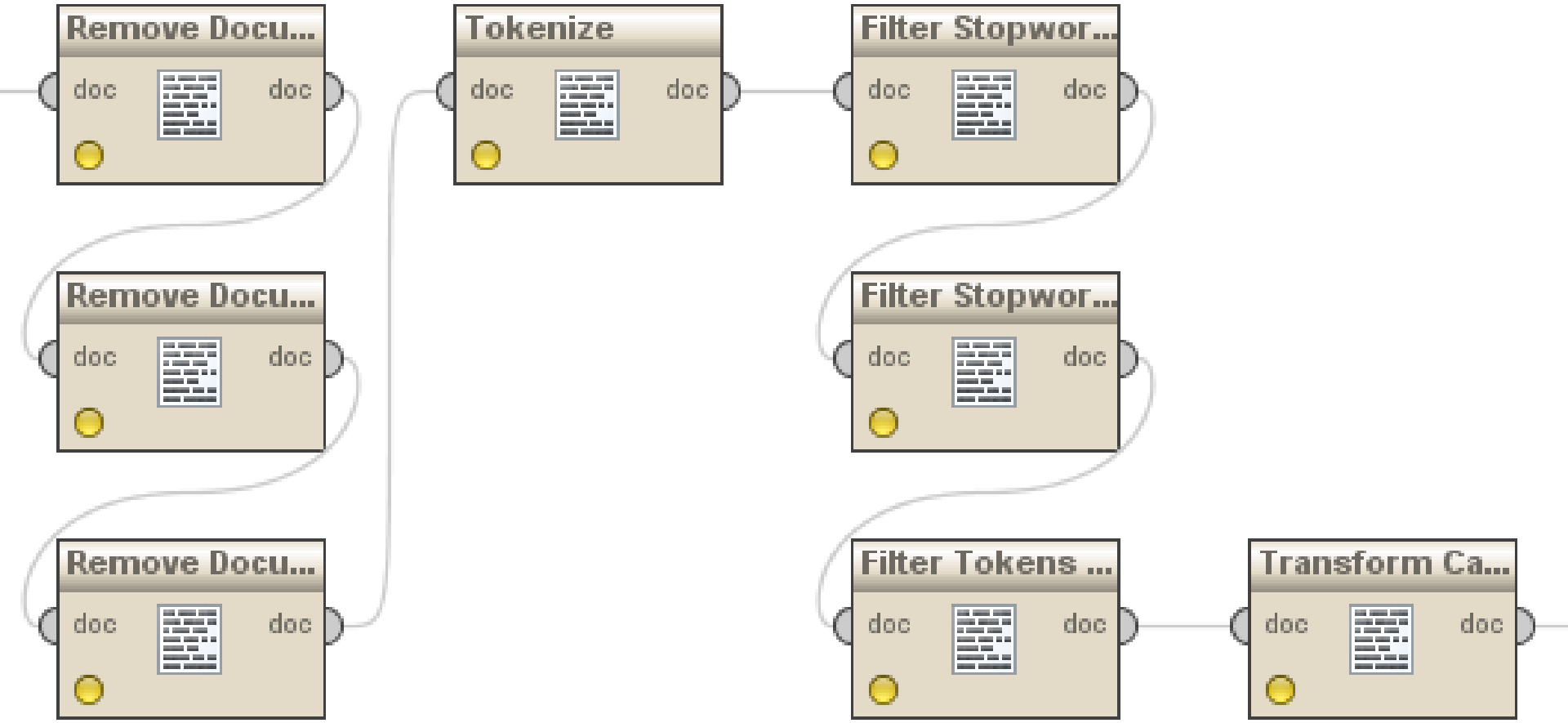
Processing CMMI documents

3 simple steps:

1. Processing documents from pdf/word files
2. Converting wordlist to data
3. Writing results to an output file (e.g. excel)



Processing documents

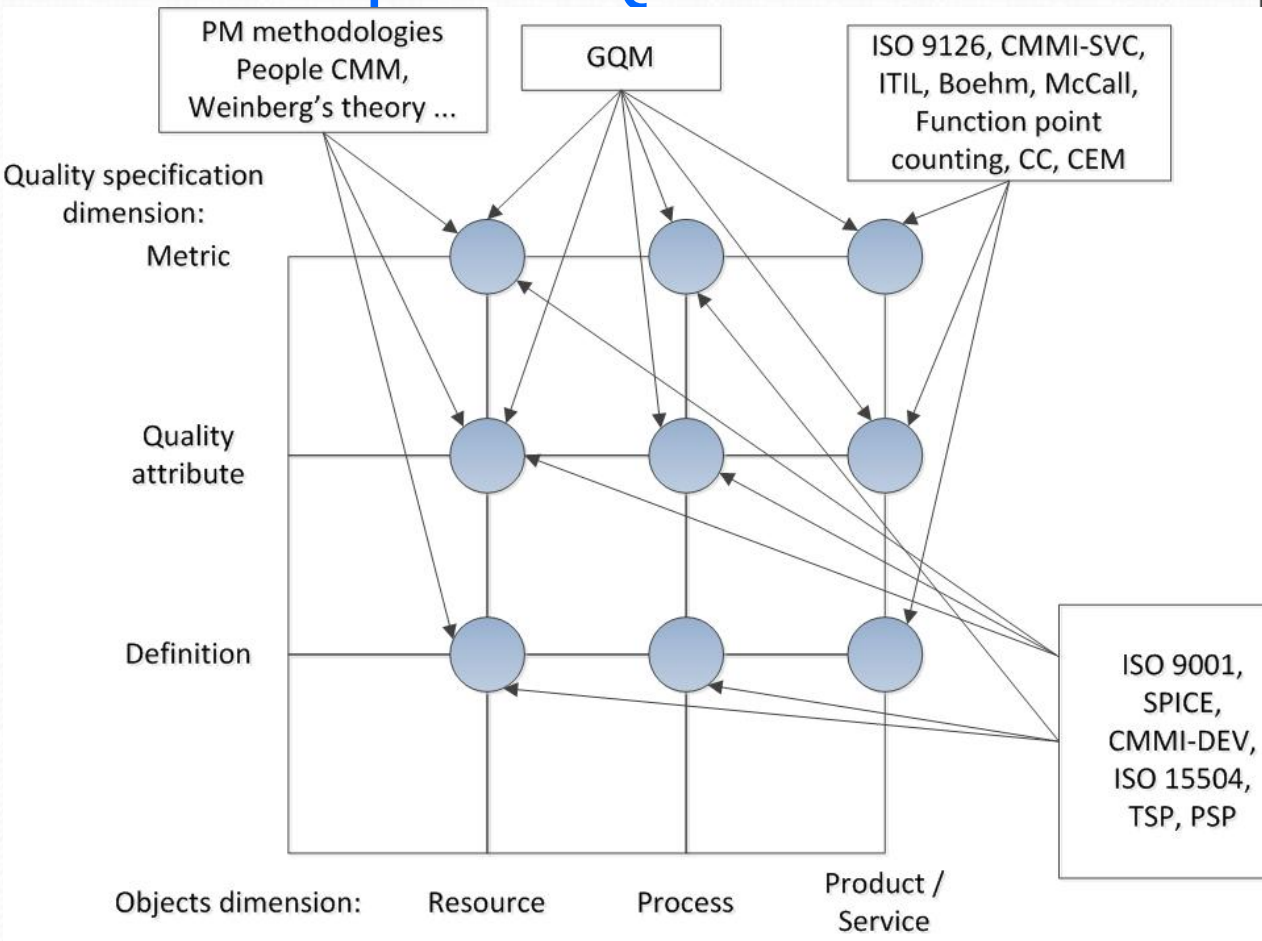


Resu

#	Tokenized wordlist word	in documents	total	Stemmed (Snowball) wordlist word	in documents	total
1	process	3	8946	process	3	10853
2	work	3	3706	product	3	4370
3	project	3	3170	servic	3	4219
4	service	3	2934	work	3	3751
5	cmmi	3	2682	project	3	3556
6	management	3	2532	perform	3	3501
7	performance	3	2437	manag	3	3459
8	requirements	3	2406	requir	3	3022
9	product	3	2338	plan	3	2988
10	organization	3	2194	area	3	2930
11	area	3	2044	cmmi	3	2682
12	products	3	1903	organ	3	2546
13	processes	3	1879	includ	3	2319
14	organizational	3	1641	measur	3	2124
15	information	3	1589	risk	3	2089
16	version	3	1577	develop	3	2017
17	objectives	3	1545	establish	3	1969
18	include	3	1538	improv	3	1924
19	analysis	3	1366	exampl	3	1863
20	supplier	3	1359	object	3	1798
21	data	3	1298	inform	3	1769
22	services	3	1285	supplier	3	1714
23	training	3	1274	organiz	3	1650
24	development	3	1262	level	3	1638
25	quality	3	1261	identifi	3	1636
26	risk	3	1225	use	3	1603
27	plan	3	1215	version	3	1594
28	activities	3	1203	select	3	1567
29	level	3	1113	practic	3	1549

The most important concepts in QMIM

- Process
- Product
- Work/Project
- Service



Finding cross-references in CMMI

Conducting the search

- We were looking for:
  - "Refer to the" … "process area …"

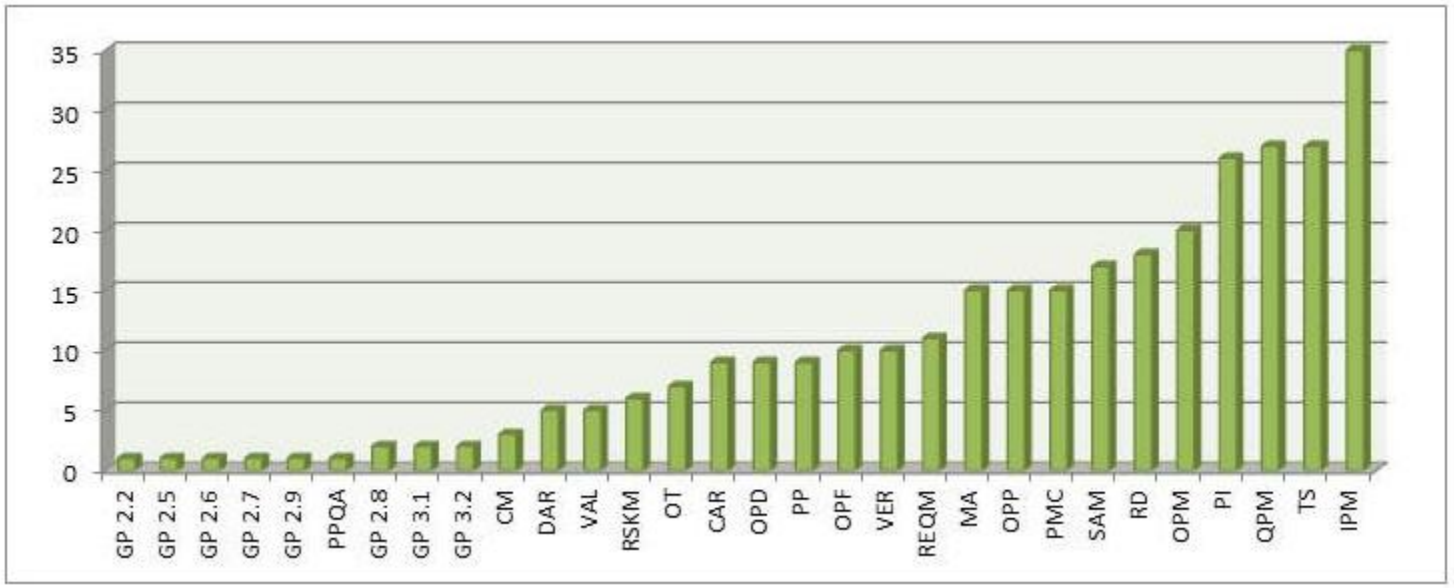
  OR
  - "Refer to the" … "specific practice in" … "process area…"
- Part 2 of all constellations were examined
- Examples were filtered
- Irrelevant references were filtered, e.g.:
  - the organization's set of standard processes can ***refer to the*** standard processes established at the organization level

Cross-references in CMMI

Results:

- Quite complex: 992 (1016) cross-references in total!
  - 311 in CMMI-DEV
  - 388 in CMMI-SVC
  - 293 in CMMI-ACQ
- Referring from different levels:
  - In PAs: Introductory notes, related process areas, specific practice level
  - In GPs
- Referring to different levels:
  - E.g. to PA, SG, SP

Understanding Cross-references in CMMI-DEV (referring from)

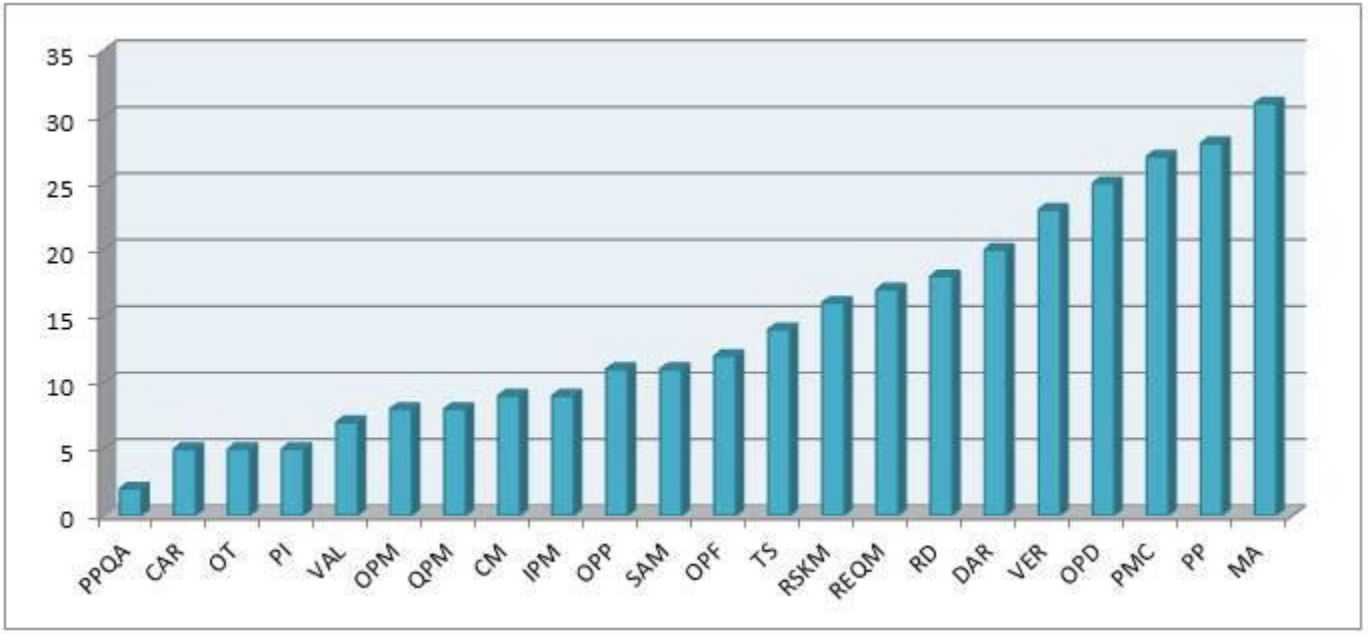


Training: GPs and PPQA are probably easier to understand

IPM implementation can be quite difficult (especially in multimodel environment),

since it relies on many other elements

Understanding Cross-references in CMMI-DEV (referring to)

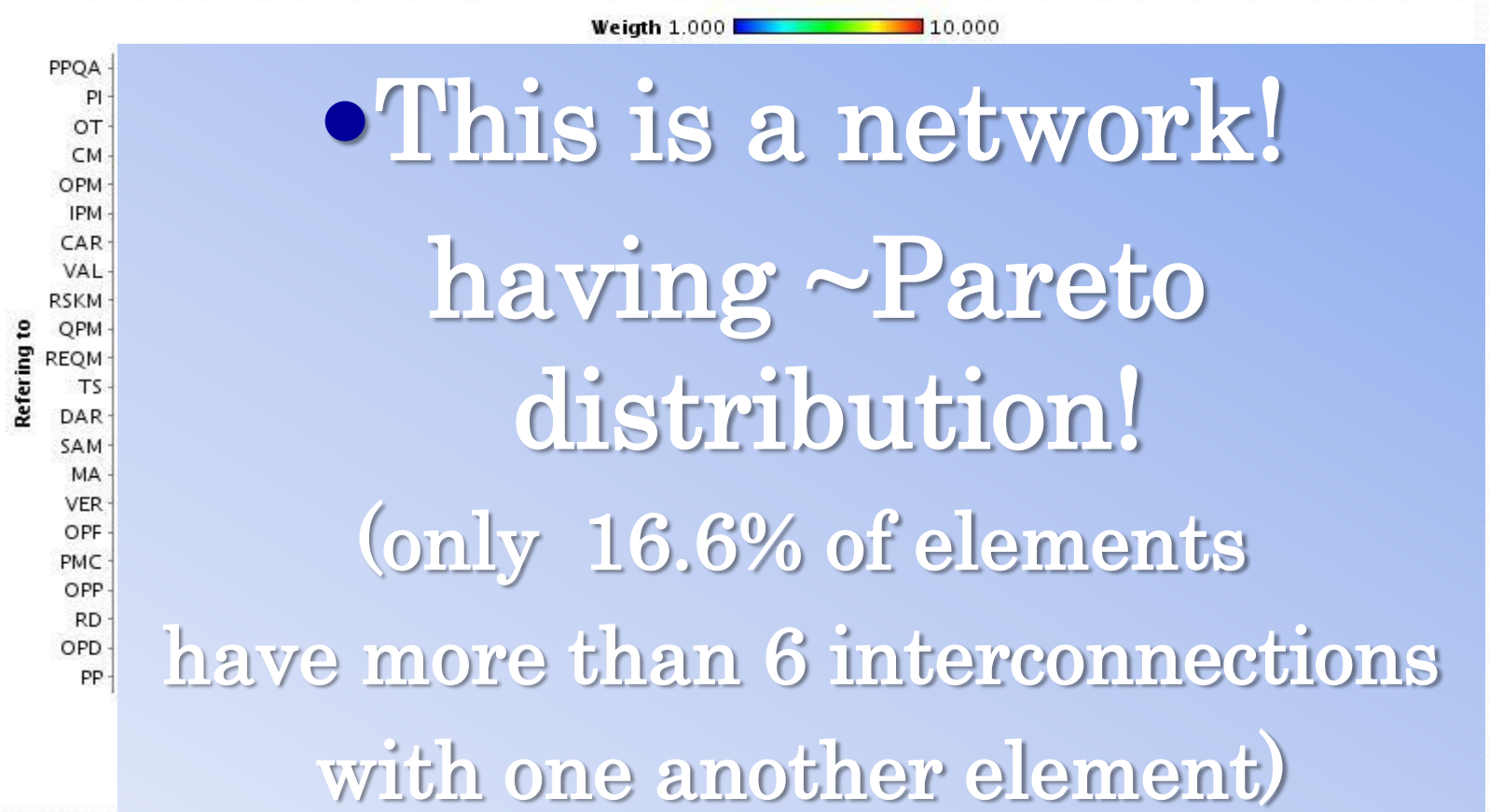


Implementing MA, PP, PMC will have effect on many other elements

PPQA, CAR, OT can probably be implemented more independently.

What about PI?

Coupled elements (139, weighted)



- This is a network!

having ~Pareto distribution!

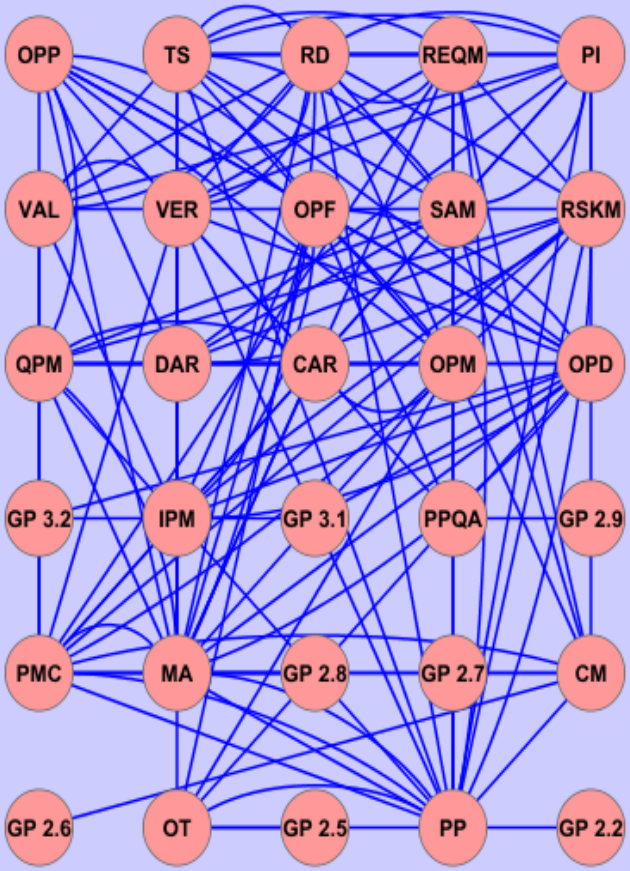
(only 16.6% of elements have more than 6 interconnections with one another element)

10 from PMC to PP, 1 from PP to PMC -> wise to implement them together

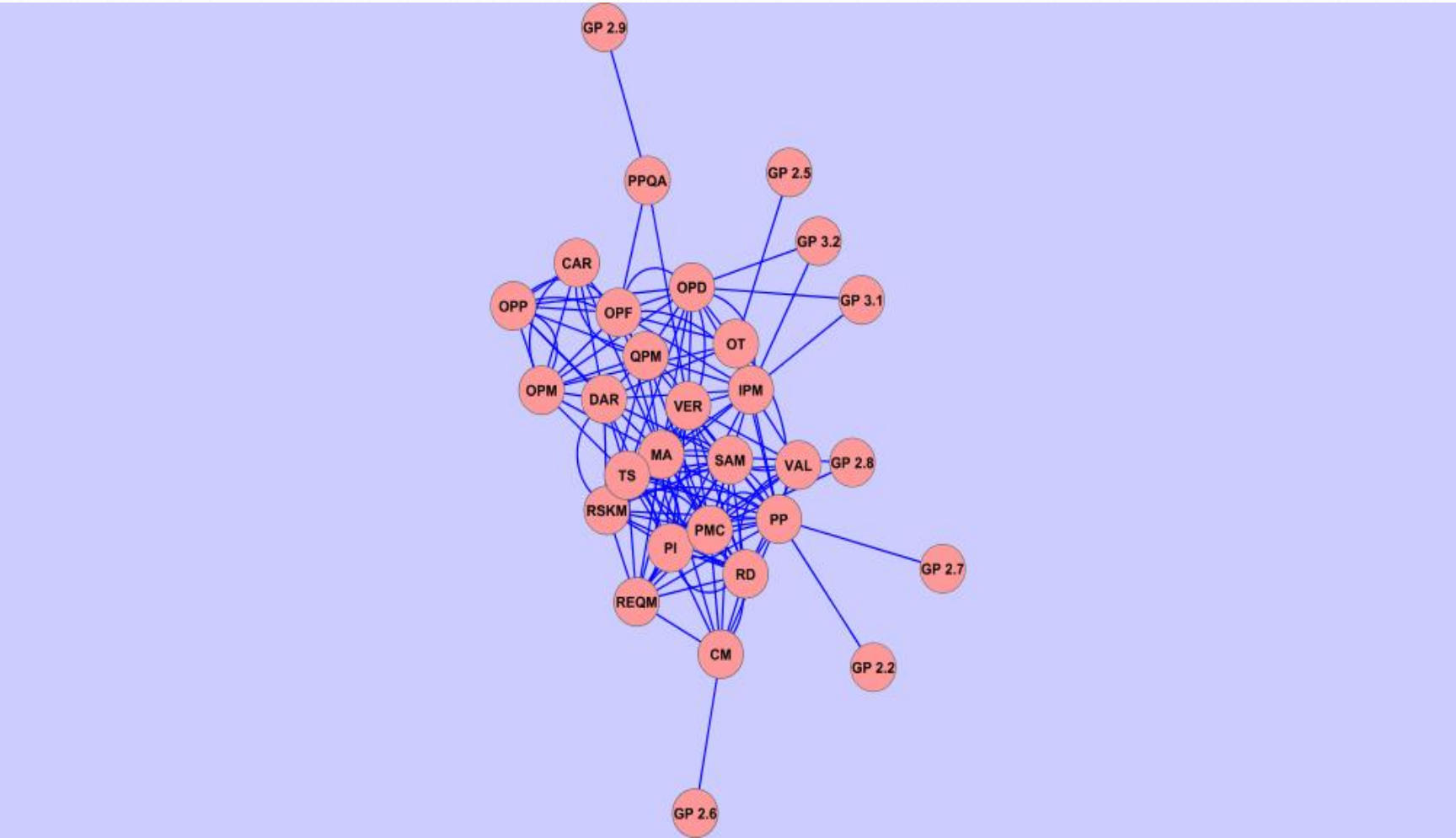
PPQA is the most isolated

TS-RD (8,4) -> 12; IPM-OPD (8, 1); OPM-OPD (6,2); QPM-OPD (4,3);

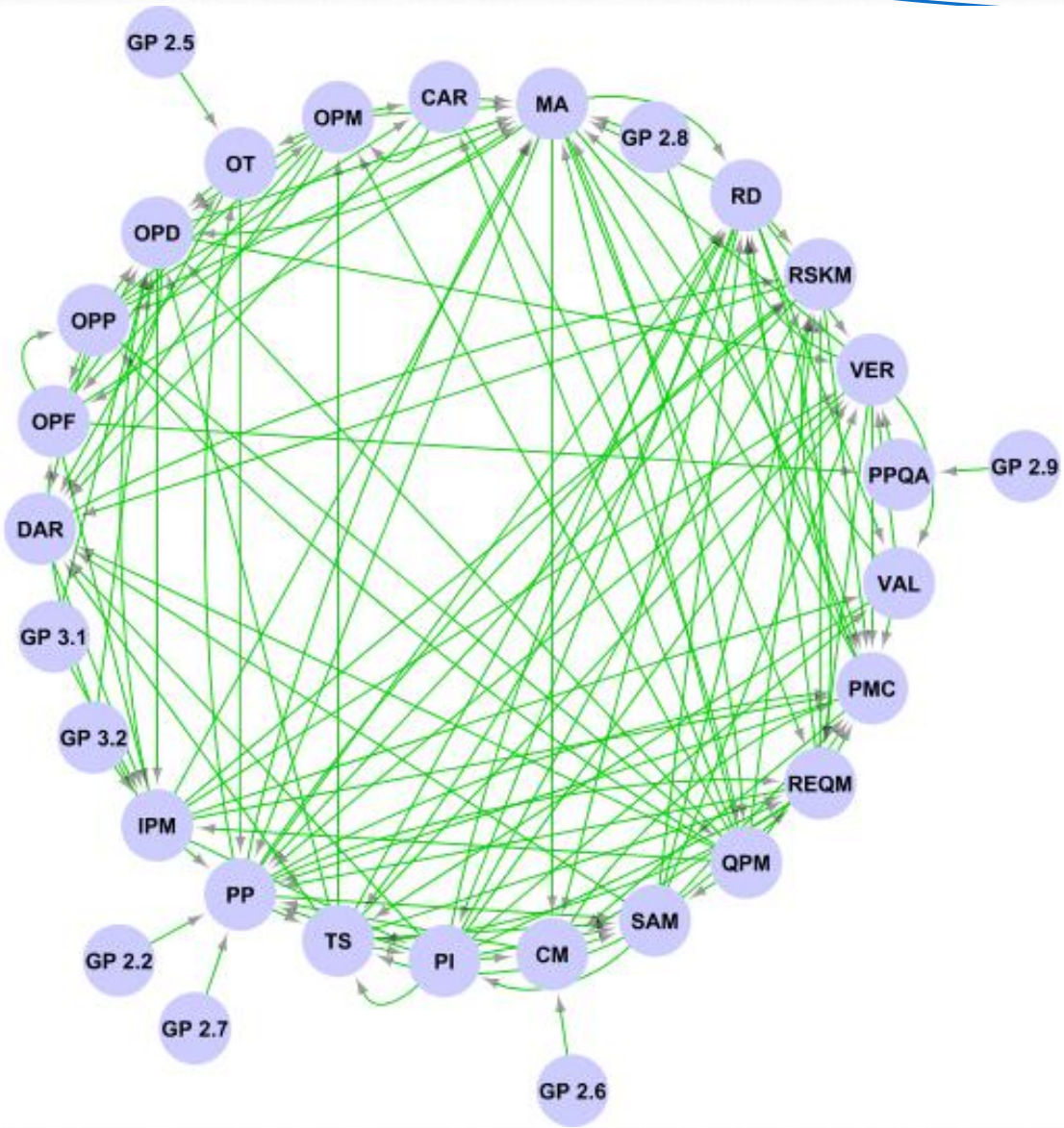
Views of the network



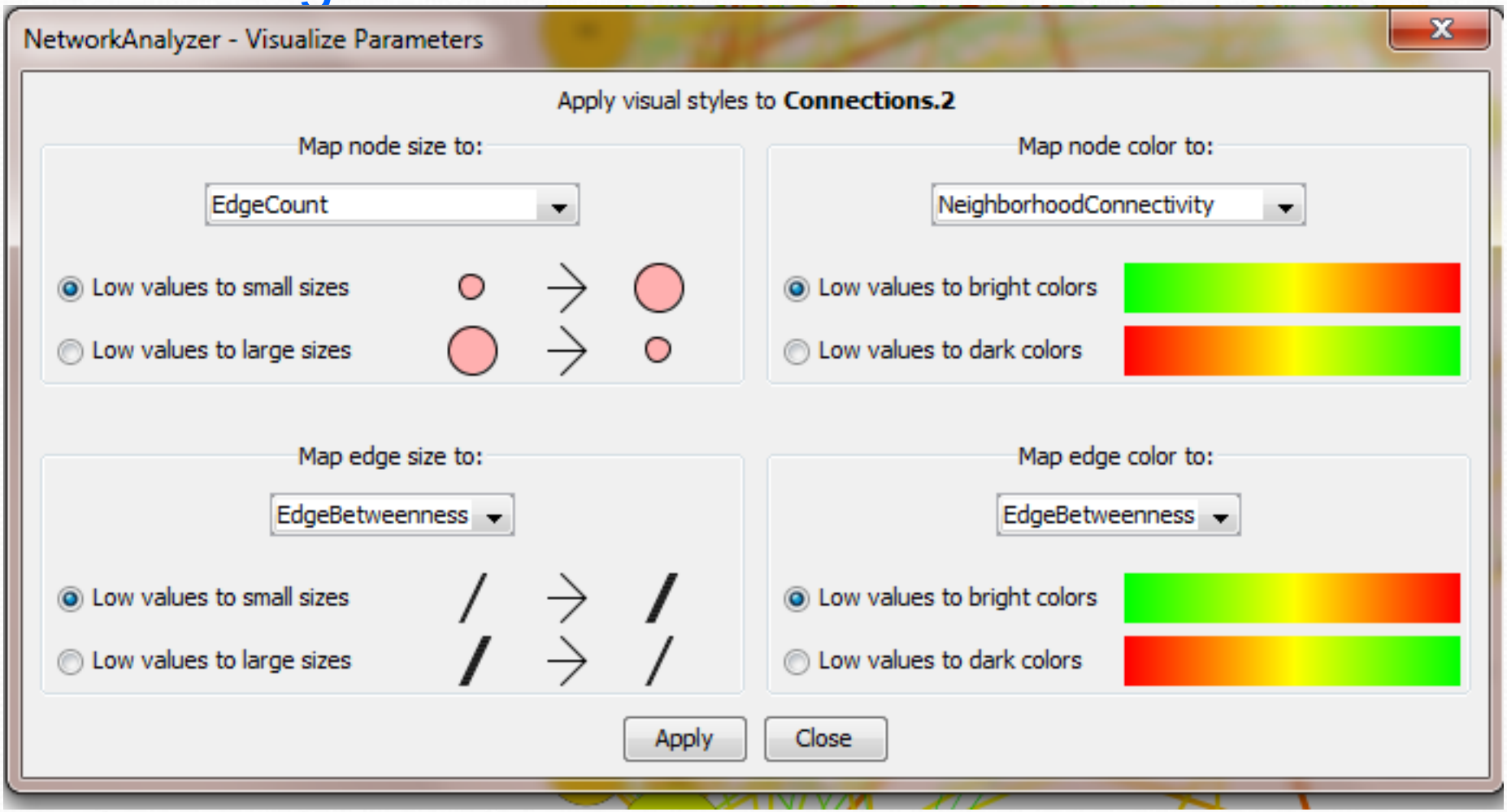
Views of the network



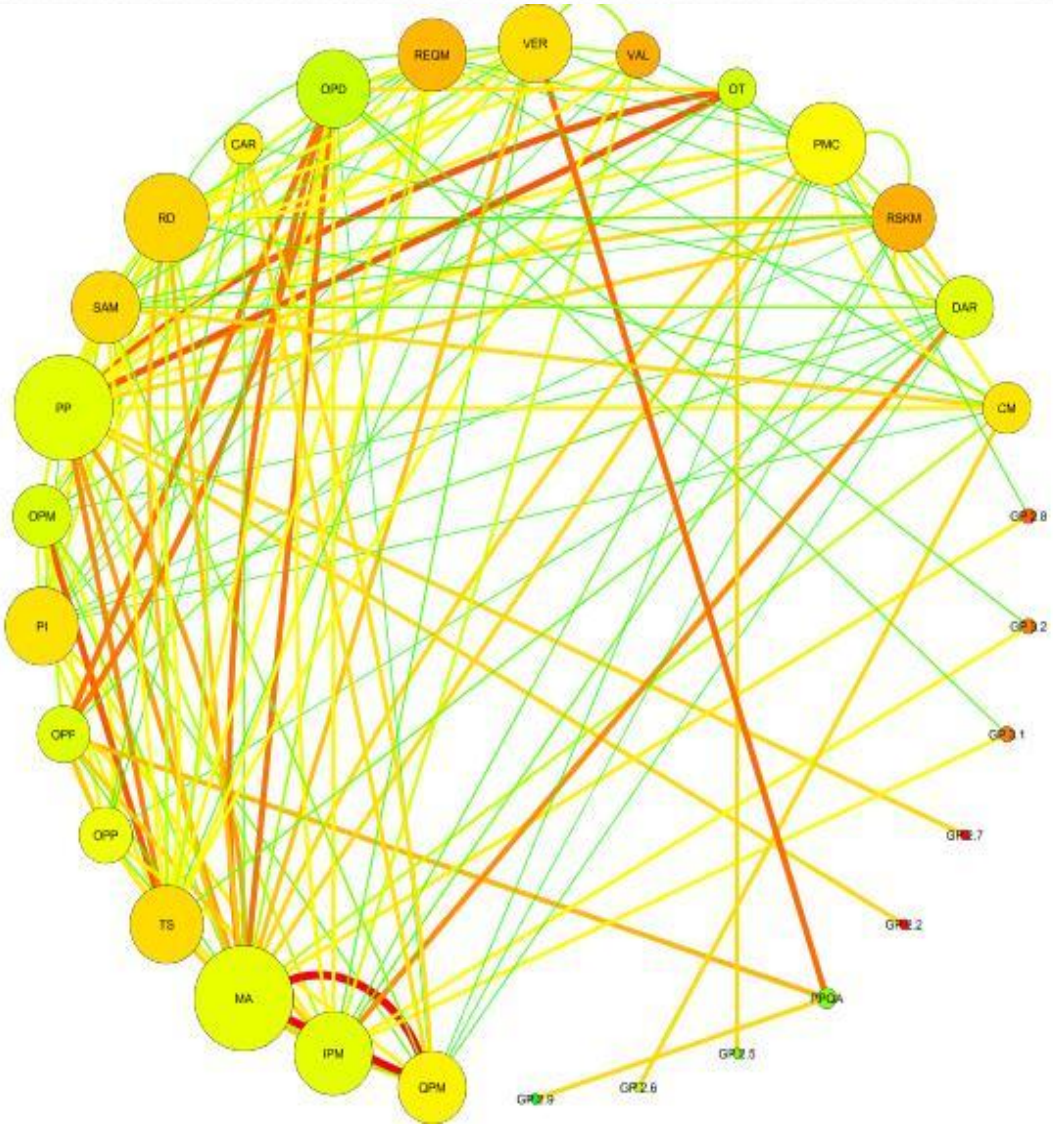
Views of the network



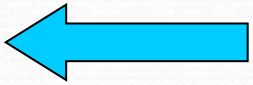
Visualizing characteristics



Visualizing characteristics



Contents

- Introduction
  - Ways of understanding CMMI
- Text mining in CMMI
  - What we did
  - Results of text mining
- Difficulties in applying CMMI concepts 
  - What people frequently misunderstand
    - Based on work with companies implementing CMMI

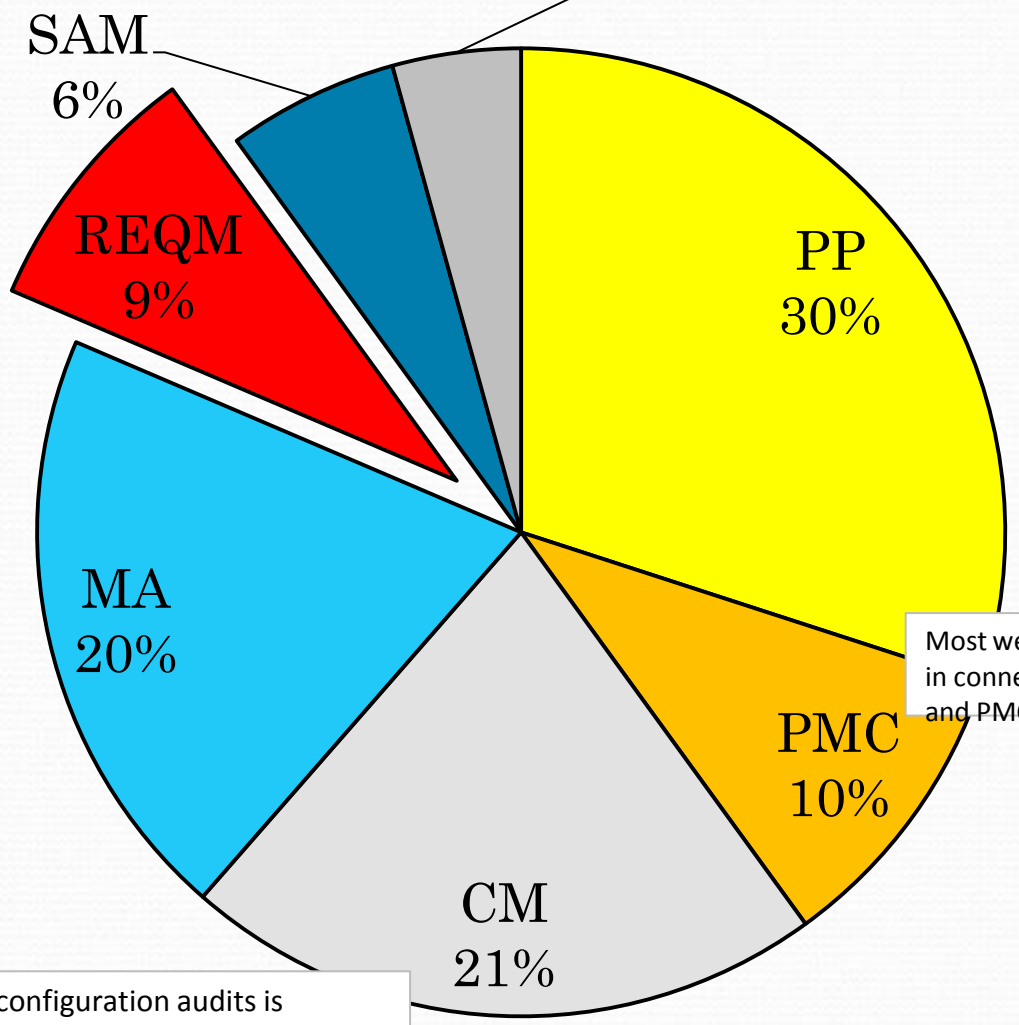
Difficult to understand for CMMI users…

- (Top 5 - based on own experience)…
  - Planning and monitoring activities different from engineering ones
  - CM for things different from code
  - Data management – often confused with CM
  - IPM as a whole (often argued that if PP, PMC is done for all projects, IPM results – so, e.g. no process description is needed)
  - GP-s, mainly planning of PP and PMC, measurement on measurement, CM on CM, relevant stakeholders for PA-s …

Some statistics

Weaknesses and notes referring to CMMI ML2 PA-s

(based on 9 SCAMPI A-s, all weaknesses+notes = 100%, 2007-2010)



Most of weaknesses refer to "maintain bidirectional traceability".

Measurement has little feedback. More PA-s measured "together".

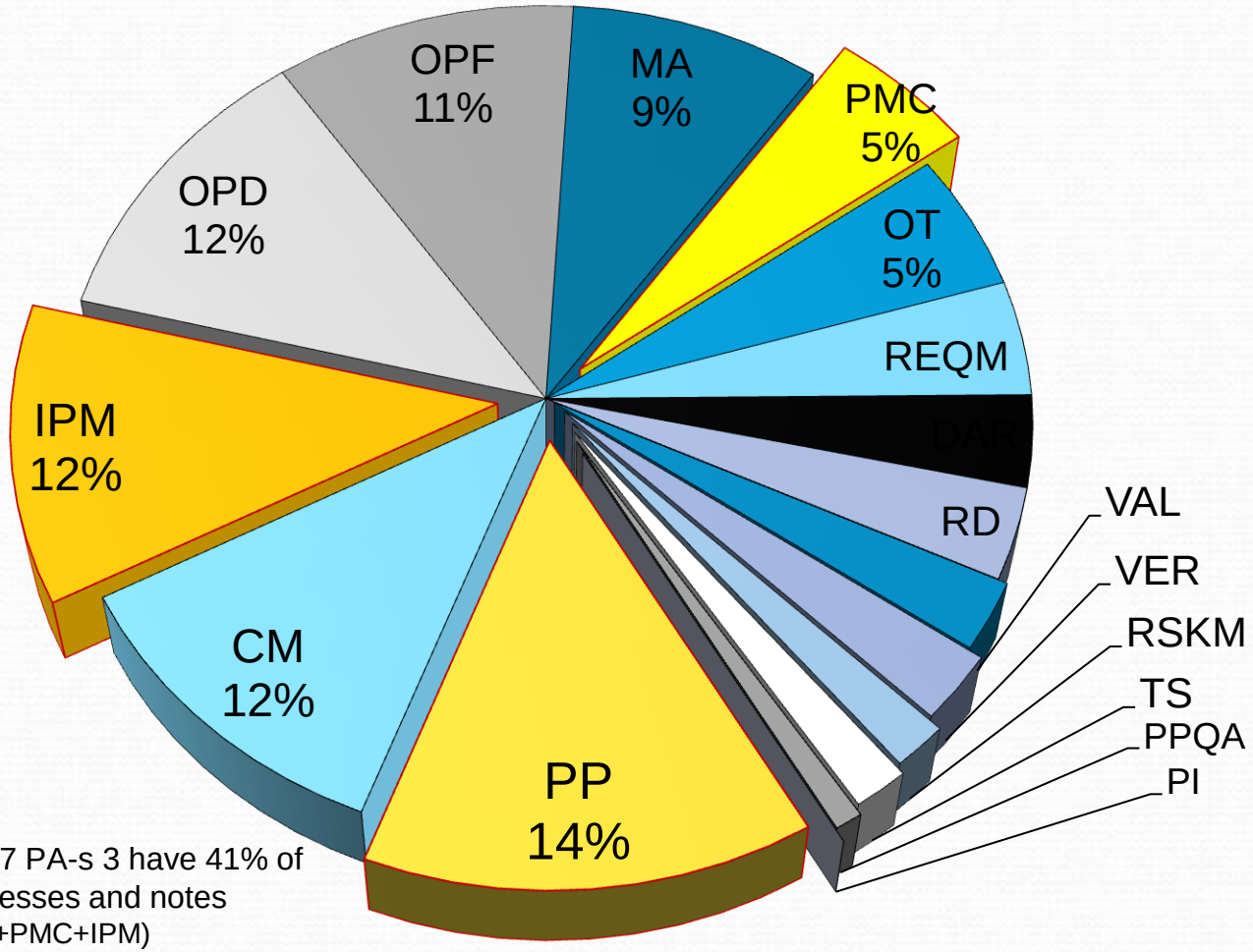
Most weaknesses are in connection with PP and PMC.

Configuration management, and mainly configuration audits is sometimes formal for work products different from code.

Some statistics

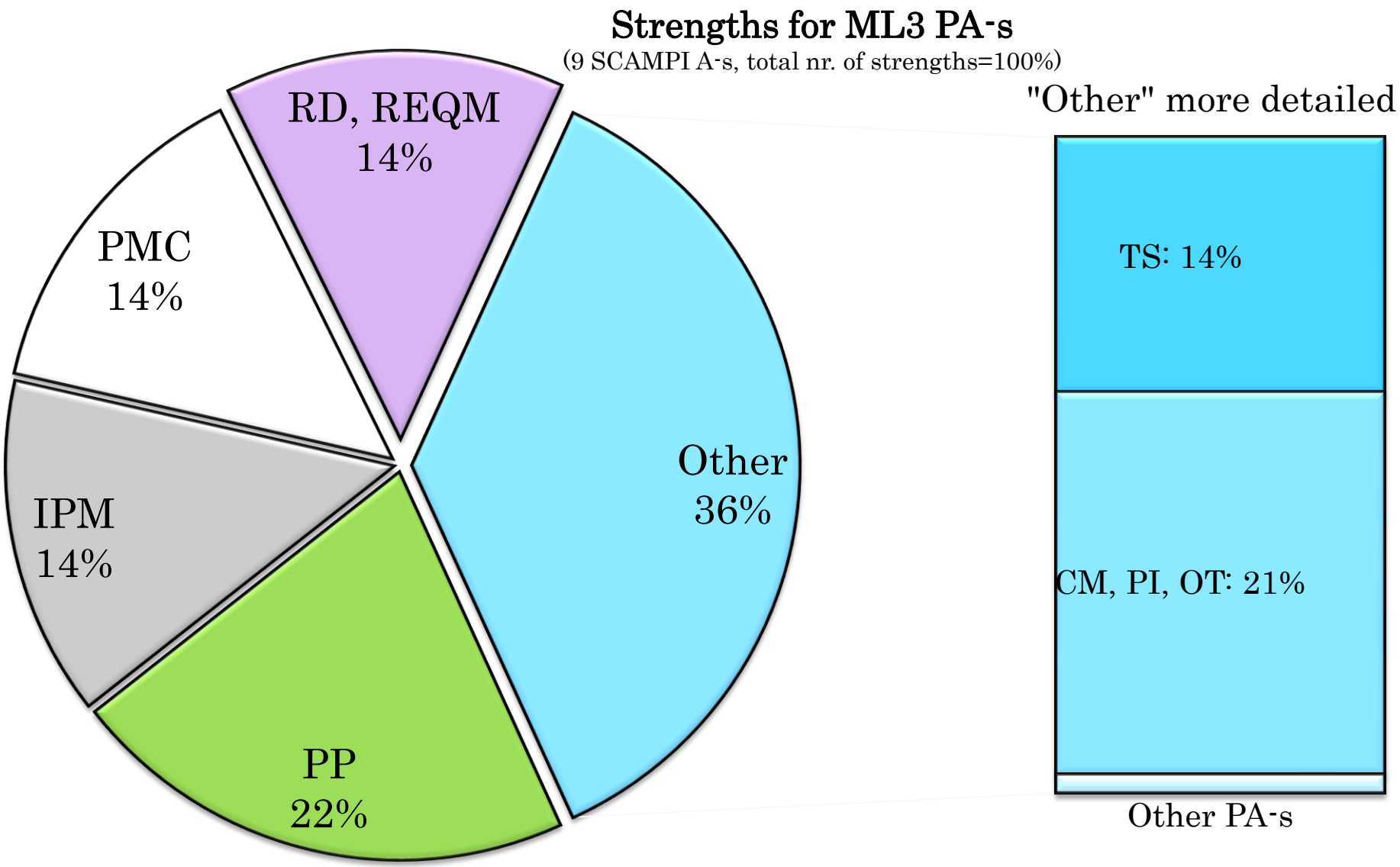
Weaknesses and notes referring to CMMI ML3 PA-s

(Excl. SAM , 8 SCAMPI A-s , 2008-2010. All weaknesses + notes = 100%)



3 PA-s out of 17 PA-s 3 have 41% of all weaknesses and notes (PP+PMC+IPM)

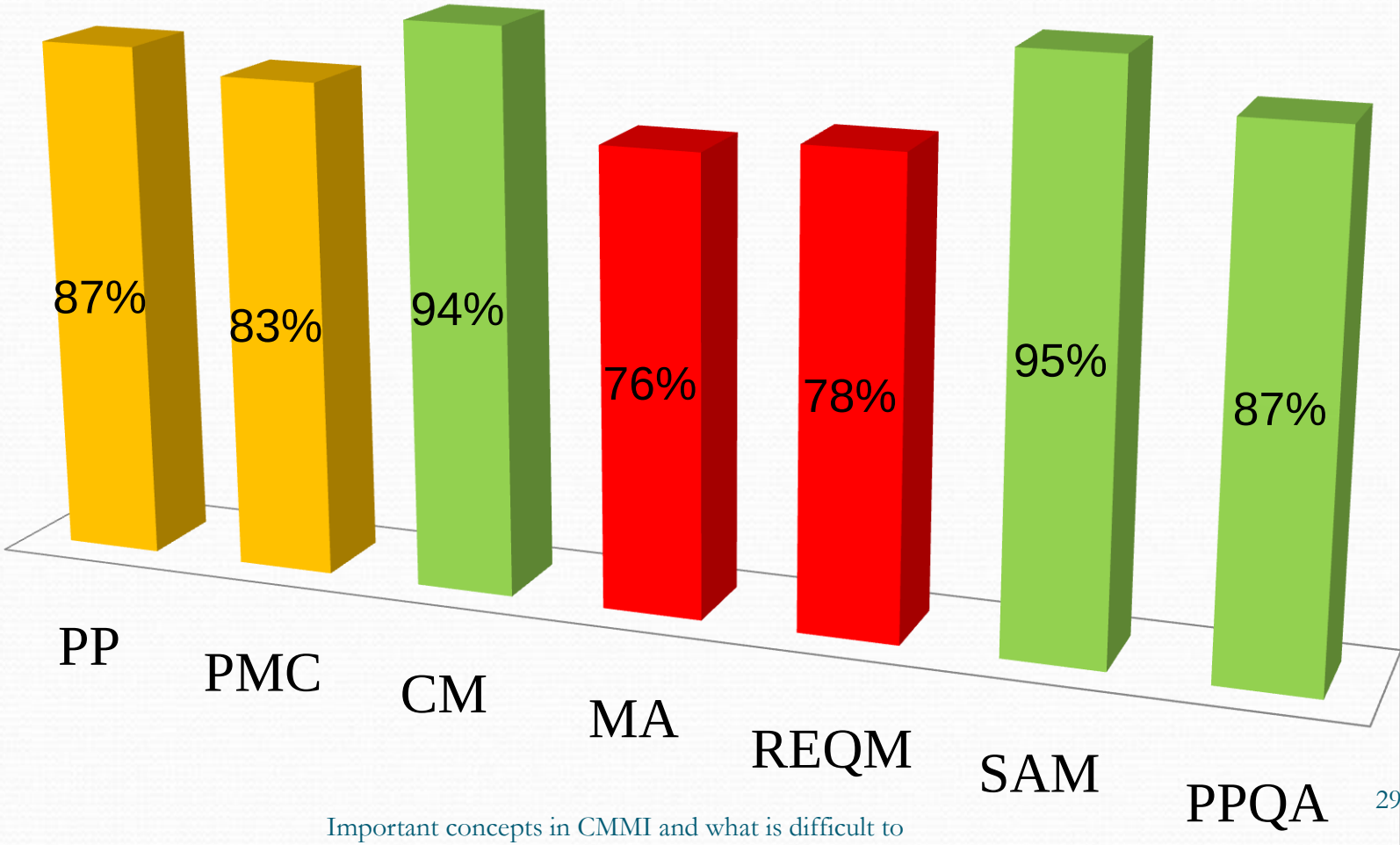
Some statistics



Some statistics

Average of ML2 PA SP-conformity for pre-appraisals

(15 pre-appraisals, PA SP conformity = 100%)



What we talked about…

- What is important in CMMI – based on text analysis
  - Confirms our previous research
- How process areas are interconnected – based on text analysis
  - Results could be useful to help companies use continuous representation
- What are the most difficult and misunderstood concepts when implementing CMMI – based on historical data

Thank you for your attention!

balla.katalin@sqi.hu
kelemen.daniel@sqi.hu
www.sqi.hu
www.iit.bme.hu