Measuring the Quality of Standard Software: the SAP Quality Index

Günther Limböck
Global Quality Governance
SAP AG, Walldorf

Software Quality Days 2009
Wien, 22. Januar 2009

Agenda

Meine Geschichte bei SAP

Messen von Software Qualität: Herausforderungen

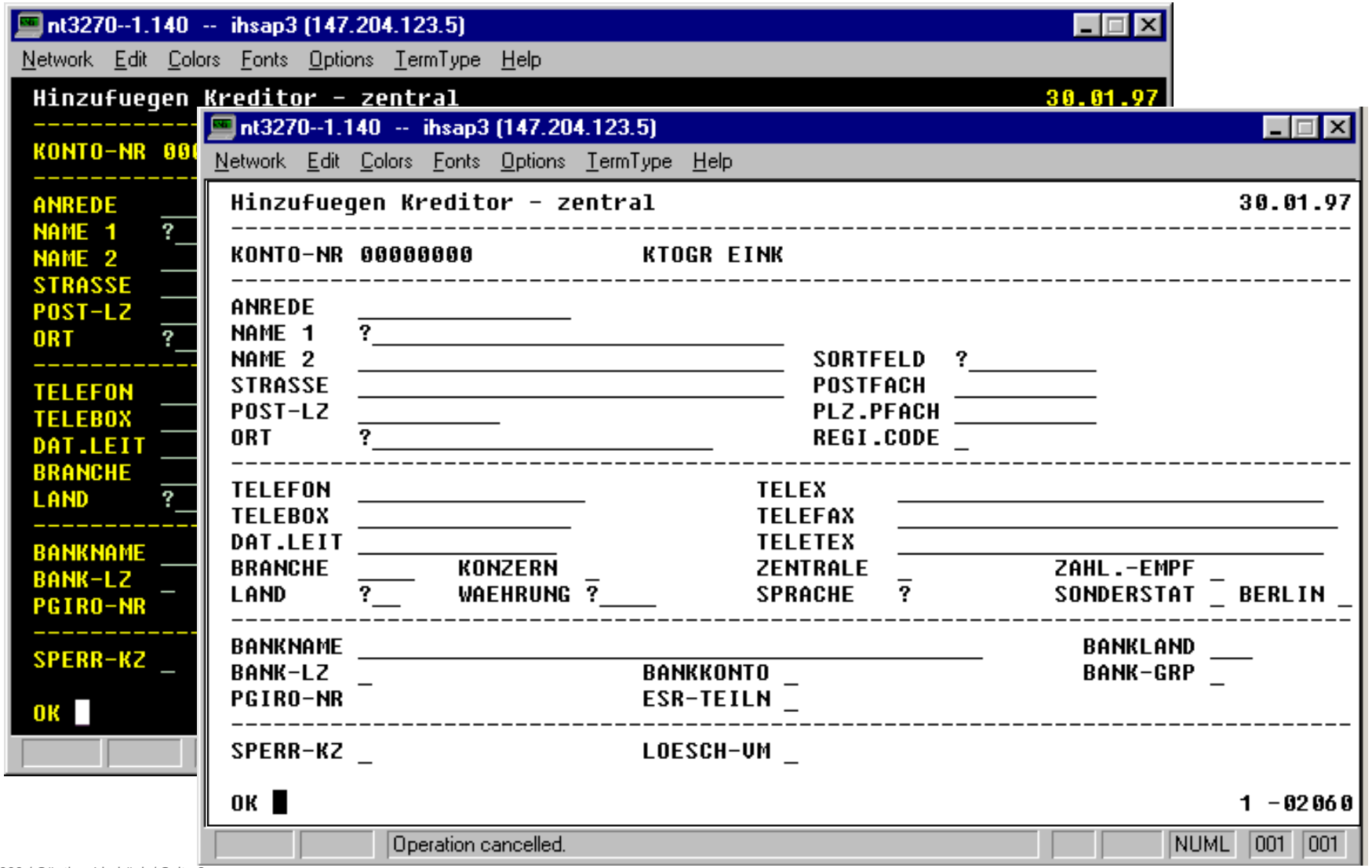
Das Projekt *SAP Quality Index*

Governance und Betrieb des *SAP Quality Index*

Erfahrungen und Ausblick

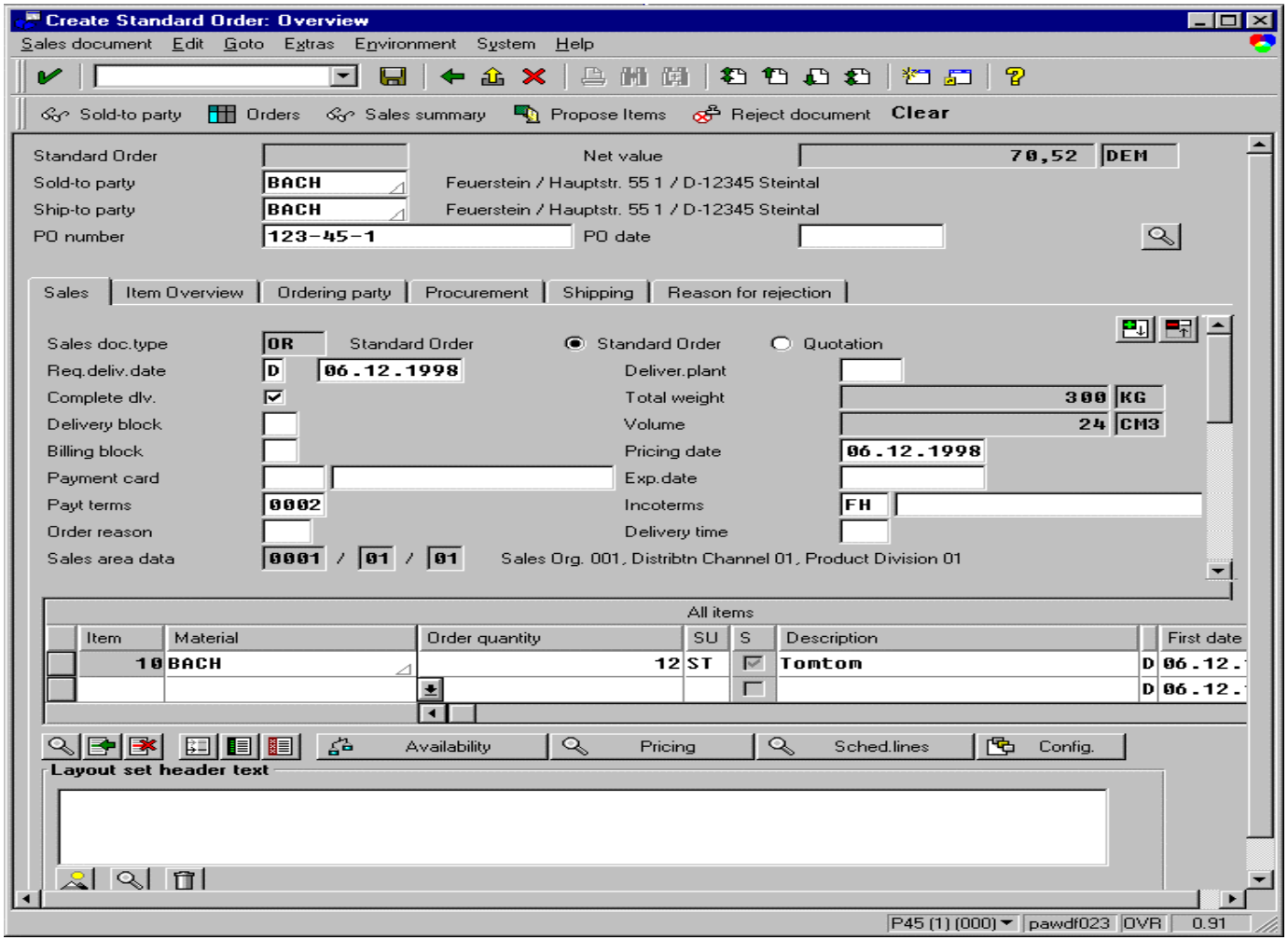
SAP 1989: 1.400 Mitarbeiter
0,19 Mrd. EUR Umsatz

SAP R/2 4.3



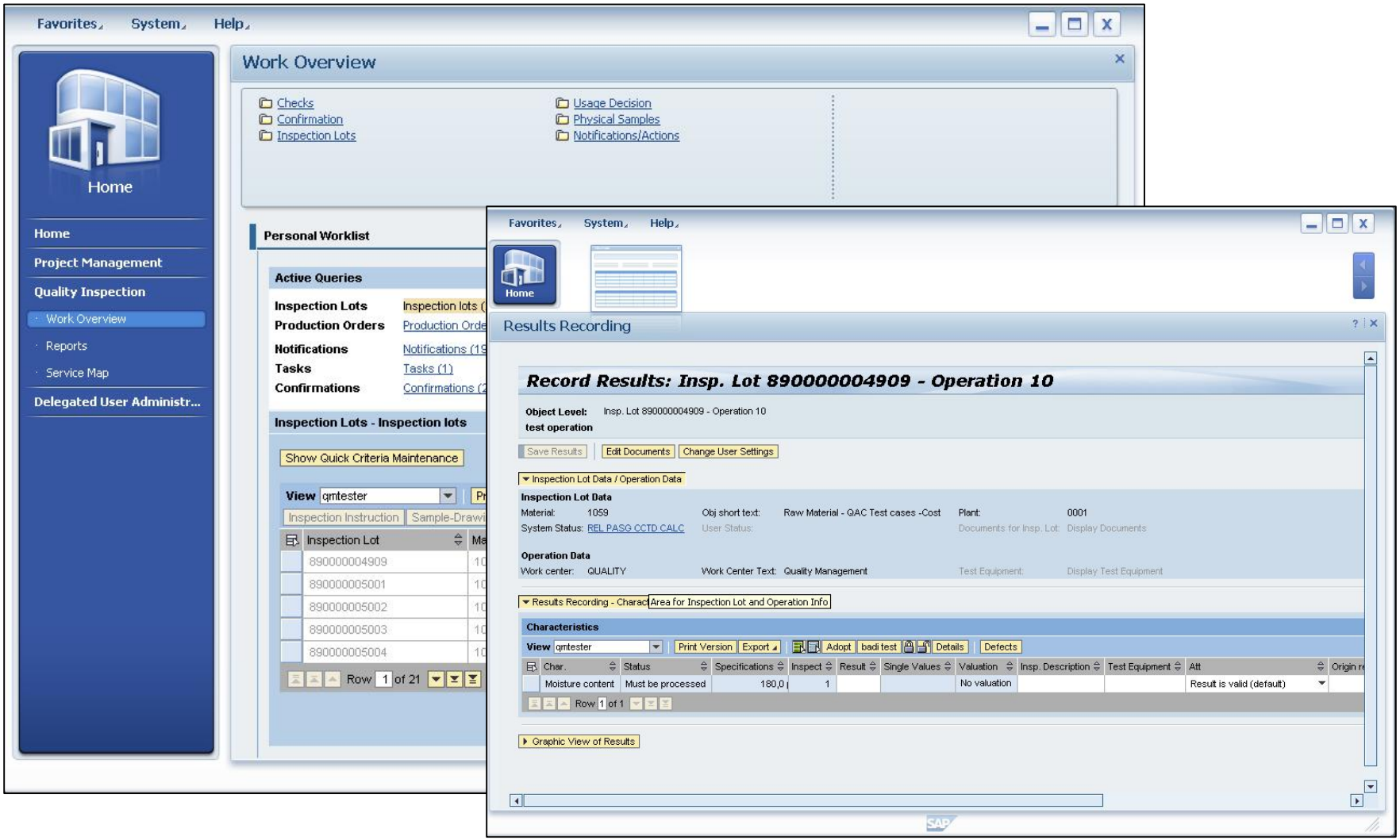
SAP 1998: **19.308** Mitarbeiter
4,33 Mrd. EUR Umsatz

SAP R/3 4.5



SAP 2008: 52.000 Mitarbeiter
10,24 Mrd. EUR Umsatz in 2007

SAP Business Suite 7.0



Agenda

Meine Geschichte bei SAP

Messen von Software Qualität: Herausforderungen

Das Projekt *SAP Quality Index*

Governance und Betrieb des *SAP Quality Index*

Erfahrungen und Ausblick

Soll man messen? Und was? Und wann?

Zitate zu „Messen“

Messen? Nein!

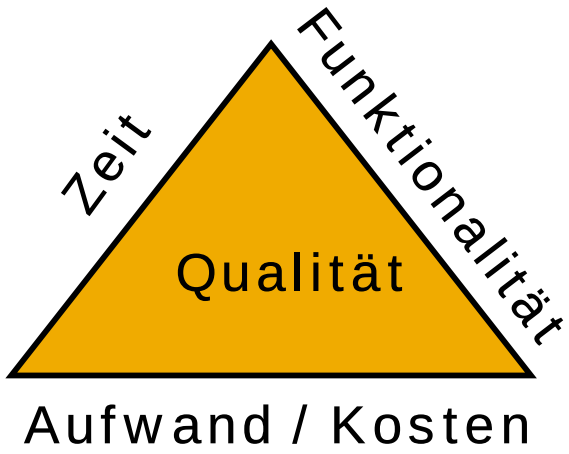
- „Wer viel misst, misst Mist.“
  - *Unbekannt*
- „All the statistics in the world can‘t measure the warmth of a smile.“
  - *Chris Hart*
- „Intense love does not measure, it just gives.“
  - *Muttter Teresa*

Messen: Ja!

- „In fact, the reduction – not necessarily the elimination – of uncertainty is central to the concept of measurement.”
  - *Wikipedia*
- If you don’t measure you’re left with only one reason to believe you’re in control: hysterical optimism.
  - *Tom DeMarco*
- „What gets measured gets managed.“
  - *Peter Drucker*

- Messen soll verwendet werden, um Unsicherheit zu reduzieren und bei Entscheidungen die Fakten zu kennen.
- Messen soll auf Produkte/Prozesse angewendet werden, nicht auf Personen.

Welche Aspekte von Software können wie gut gemessen werden?



Welche Aussagen können üblicherweise über Softwareprojekte gemacht werden?

Unser letztes großes Software Entwicklungsprojekt

- **hatte eine Dauer von 18 Monaten**
- **verursachte einen Aufwand von 80.000 Personentagen (80 Mio. Euro)**
- **hatte eine Funktionalität von 16.000 Functions Points**
- **hatte bei Auslieferung eine Qualität von 8,3 (auf einer Scala von 1 bis 10)**
- **hatte eine verbleibende Defect Density von 0,75 Defects pro Function Point**

Agenda

Meine Geschichte bei SAP

Messen von Software Qualität: Herausforderungen

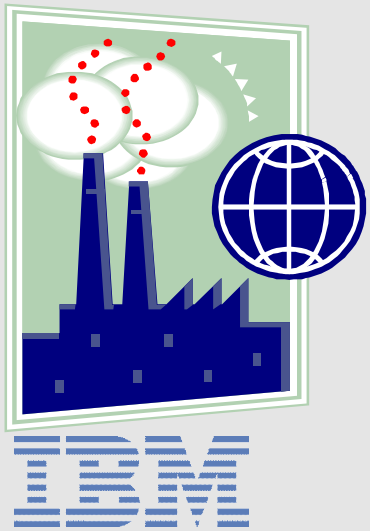
Das Projekt *SAP Quality Index*

Governance und Betrieb des *SAP Quality Index*

Erfahrungen und Ausblick

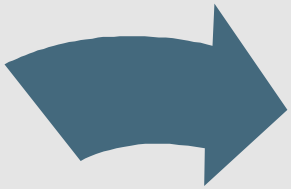
Eine Geschichte zum Messen von Qualität

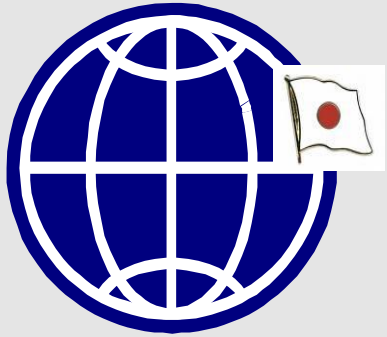
In den 1980er-Jahren bestellte IBM Canada Ltd. in Markham, Ont., Bauteile von einem neuen japanischen Zulieferer

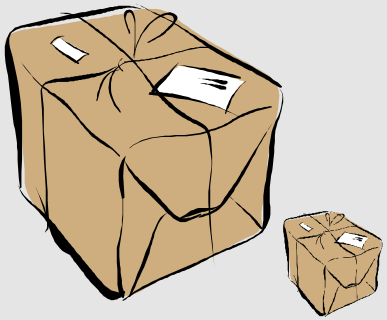


Auftrag

… toleriert werden
maximal 1,5%
fehlerhafte Teile
…



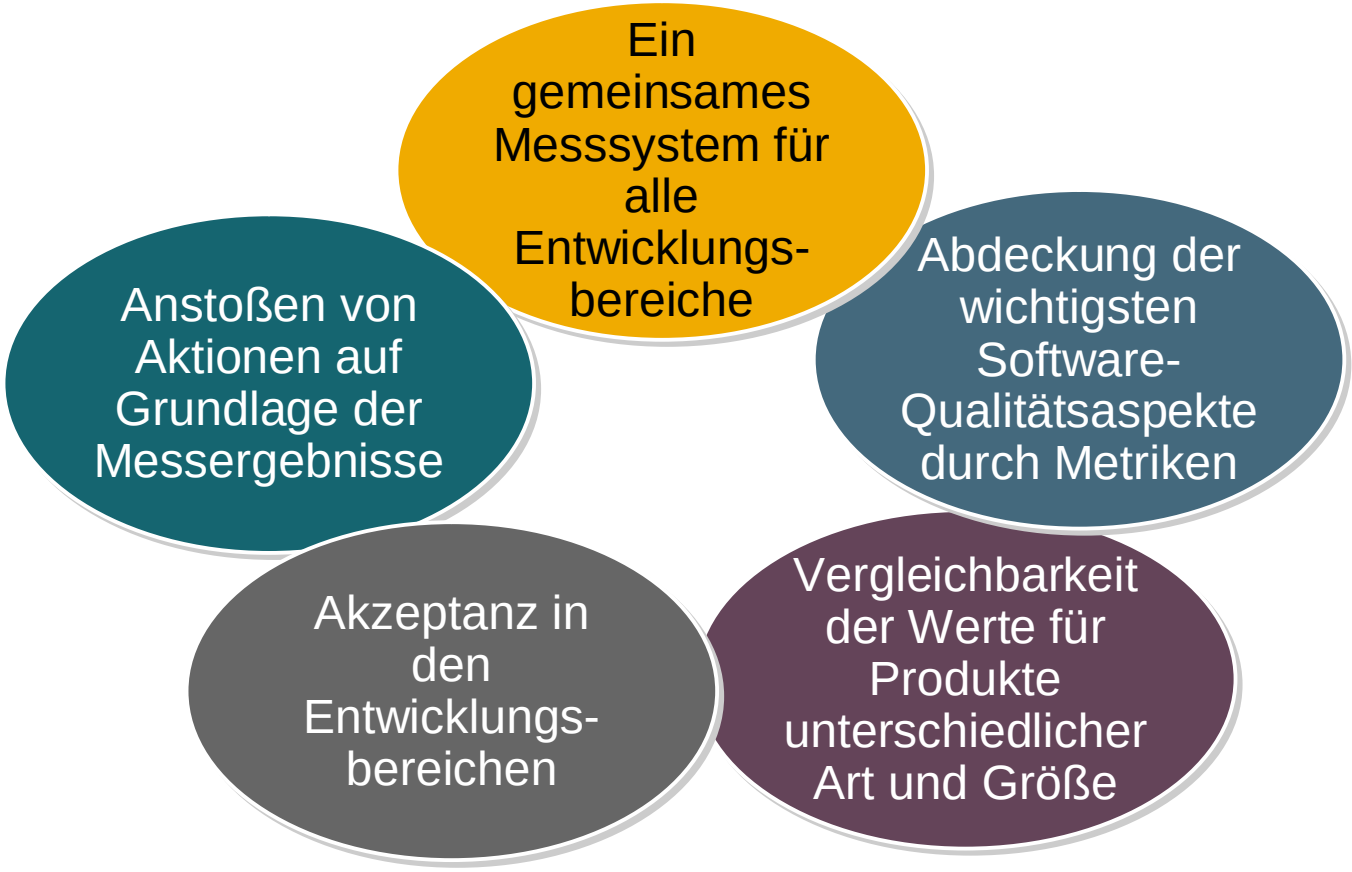




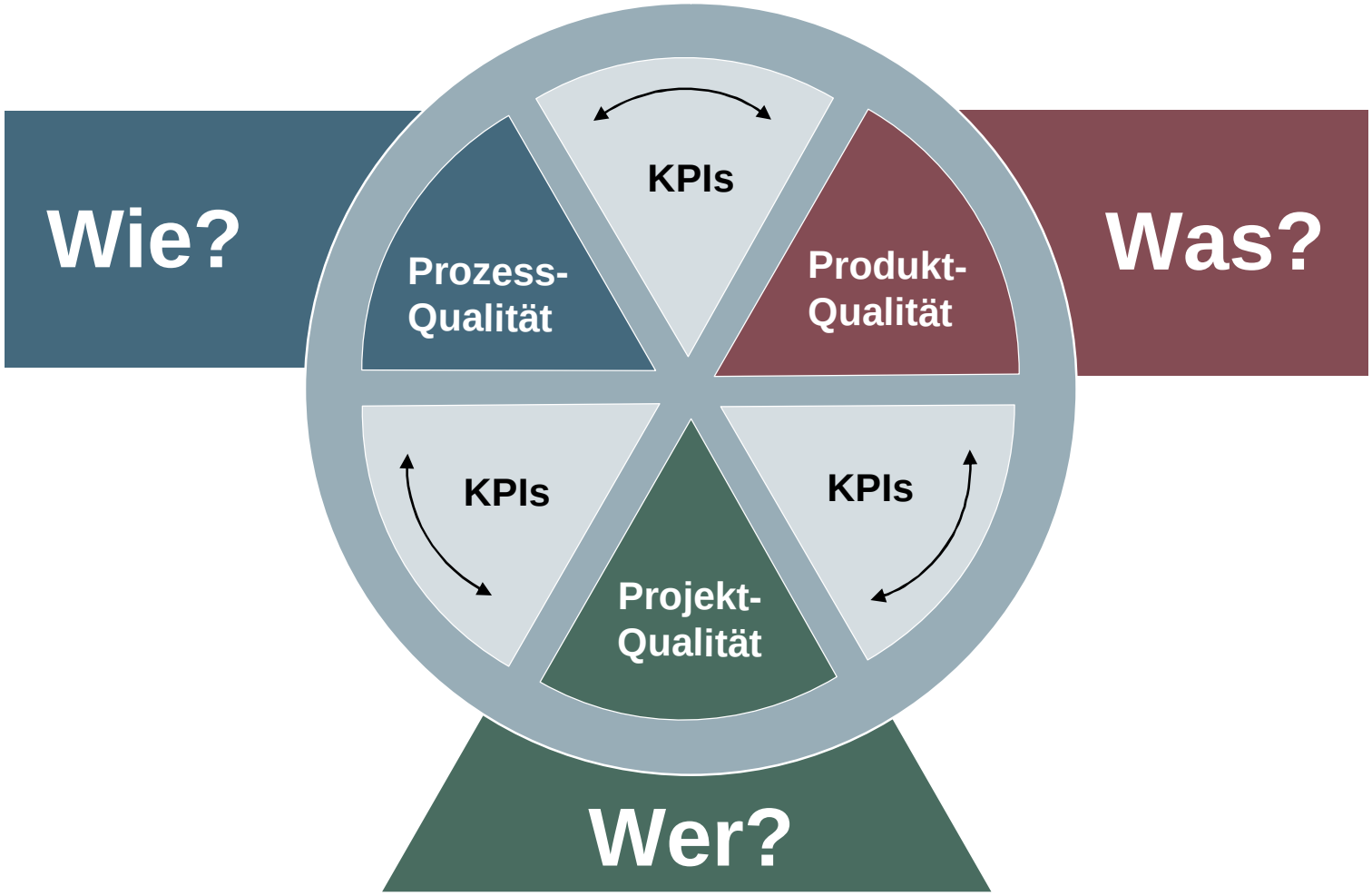
Begleitschreiben

"Wir wissen zwar nicht, warum Sie 1,5 Prozent fehlerhafte Teile haben möchten, aber um Ihnen die Arbeit zu erleichtern, haben wir sie separat verpackt."

SAP Quality Index: Der Projektauftrag im August 2005



Welche Aspekte sollen in einen Quality Index?



Qualität aus Kundensicht → Fokus auf Produkt-Qualität

Welche Daten werden für die Definition einer KPI gebraucht?

Vorlage zur Definition der KPIs

<table>
<tr><td>Name:</td><td rowspan="2">Type:
■ Product / Process / Project
■ Objective / Subjective
■ Direct / Indirect

Breakdown granularity
■ Project
■ Application hierarchy
■ Others:</td><td rowspan="2">Relevance to:
■ PIL phases
Invent / Define / Develop / Deploy / Optimize
■ SAP product standard
■ Significance
(10 = high, medium, 0 = low)</td></tr>
<tr><td>Goal:

Question:

Metric description:</td></tr>
<tr><td>Measurement / Calculation procedure

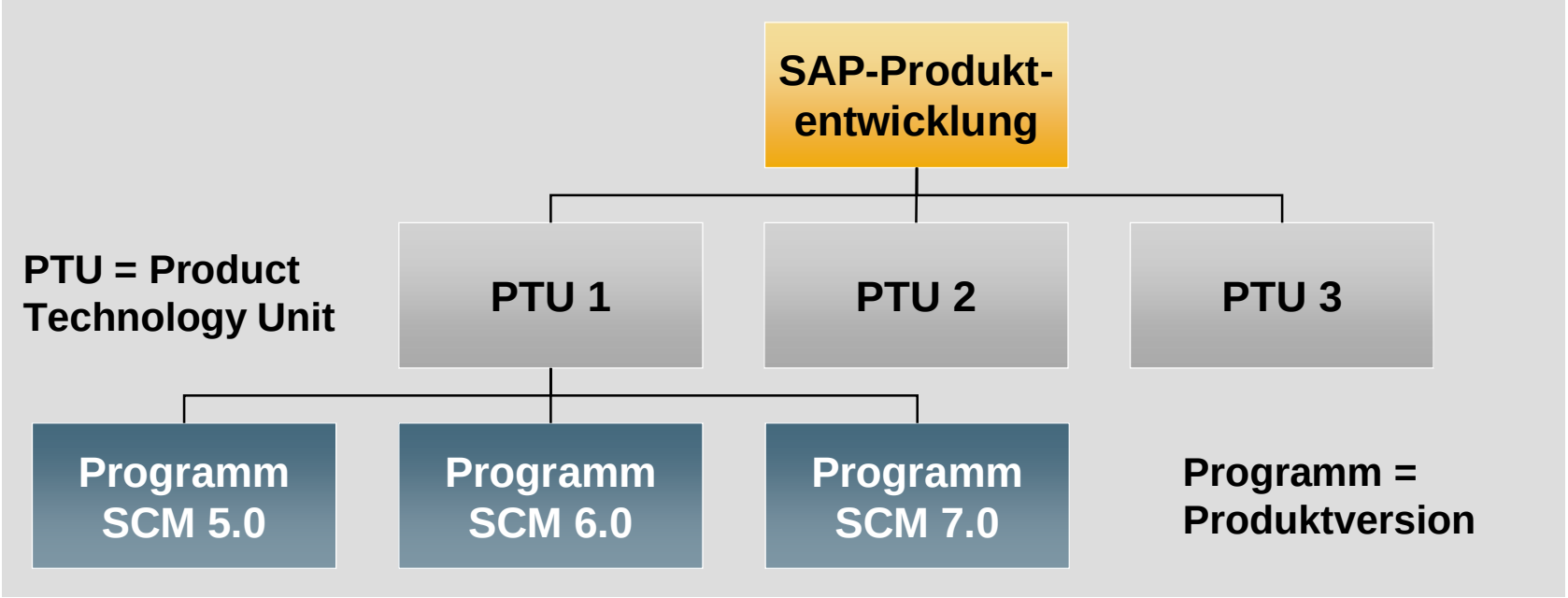
Aggregation rules

Measurement:
■ Start of measurement
■ Start of judgment, end of judgment
■ Maximum frequency</td><td colspan="2">Implementation details

■ Source (tool)
■ Tool changes
■ Process changes
■ Business Information Development</td></tr>
<tr><td colspan="3">Comments
■</td></tr>
</table>

Für welche Entität soll Qualität gemessen werden?

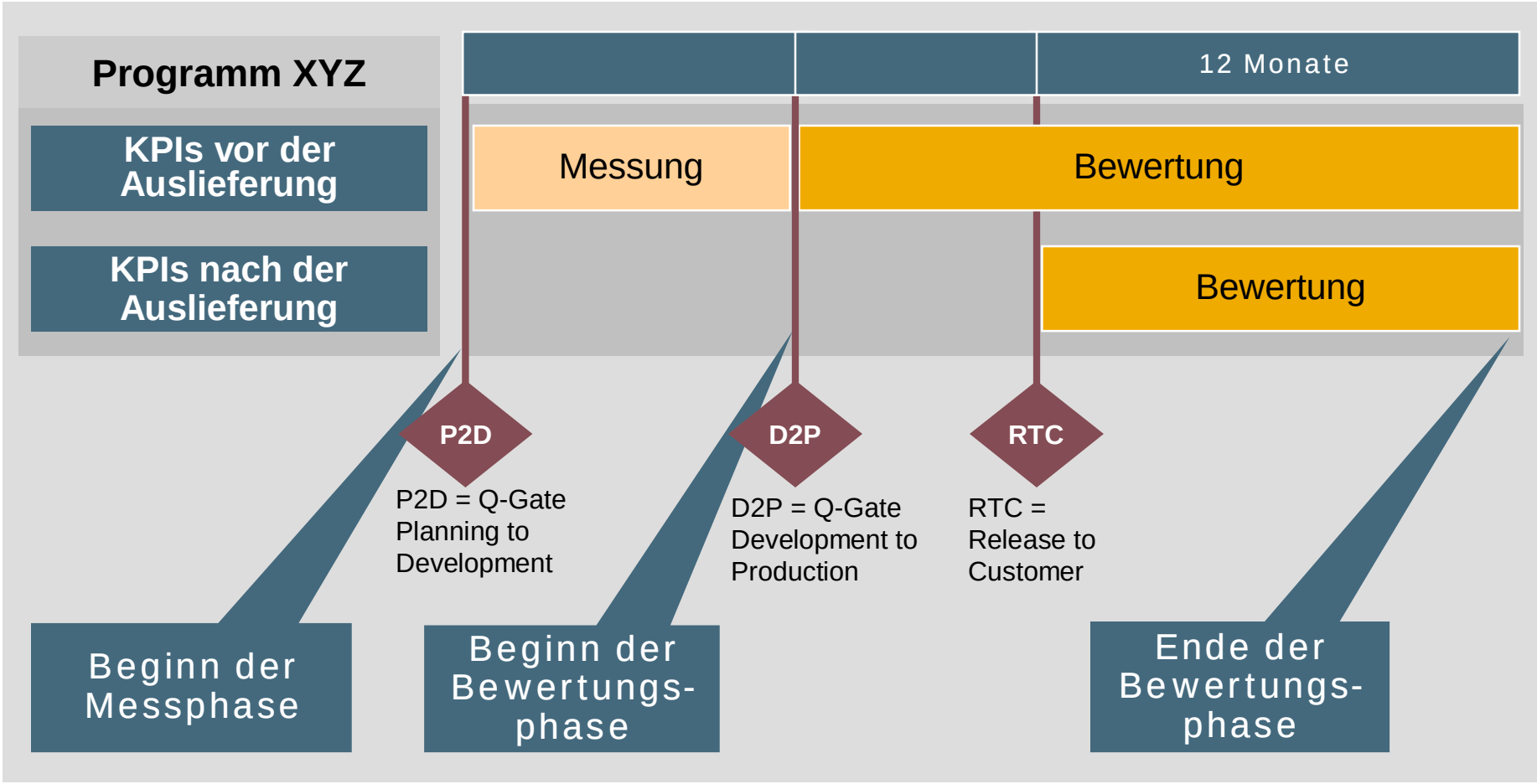
Organisationsstruktur der SAP-Entwicklung



- Programme repräsentieren das, was der Kunde auf der DVD bekommt
- Ein Programm ist daher die zentrale Entität, auf der Qualität von Software bei SAP gemessen wird

Wann soll gemessen werden, wann soll bewertet?

Zeitachse pro Programm:



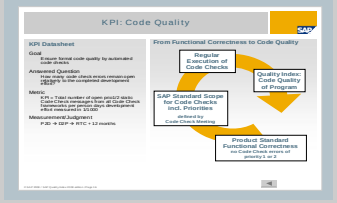
SAP Quality Index

Überblick über die Key Performance Indikatoren–

KPIs während der Development-Phase: Forecast der Produkt Qualität

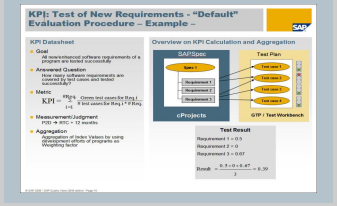
Code Quality

Ensure formal code quality by automated code checks



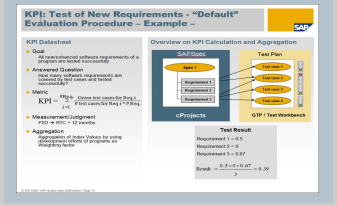
Test of New Requirements

All new/enhanced software requirements of a program are tested successfully



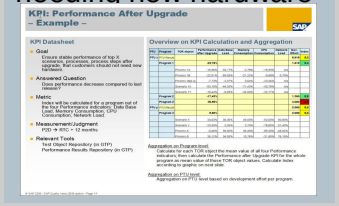
Regression Test Coverage

Ensure stability of previously delivered product (code) with regression tests



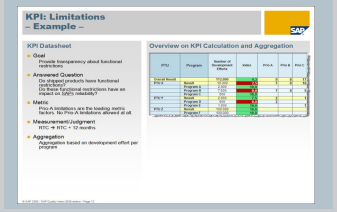
Performance after Upgrade

Stable performance of top X processes after upgrade precludes customers from needing new hardware



Limitations

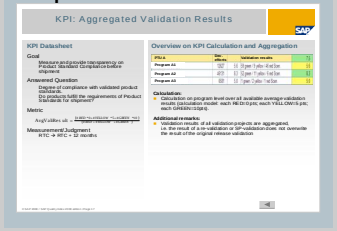
Provide transparency about functional restrictions



KPIs nach Auslieferung: KPIs: *Produkt Qualität aus Kundensicht*

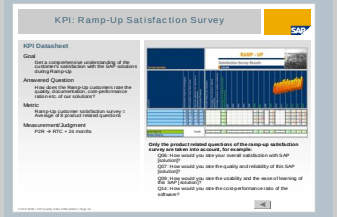
Aggregated Validation Results

Measure and provide transparency on Product Standard Compliance before shipment



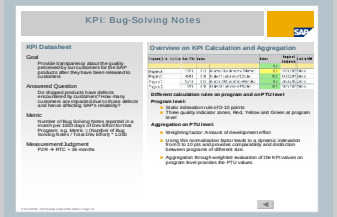
Ramp-Up Survey

Get comprehensive understanding of customer's satisfaction with SAP solutions during Ramp-Up



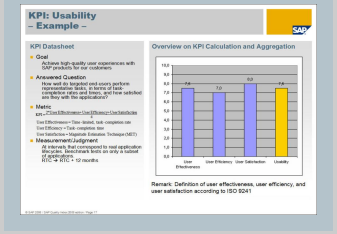
Bug Solving Notes

Provide transparency about customer quality perception for products after release to customers



Usability

Achieve high-quality user experiences with SAP products for our customers



Beschreibung der KPI: Test of New Requirements

Beispiel

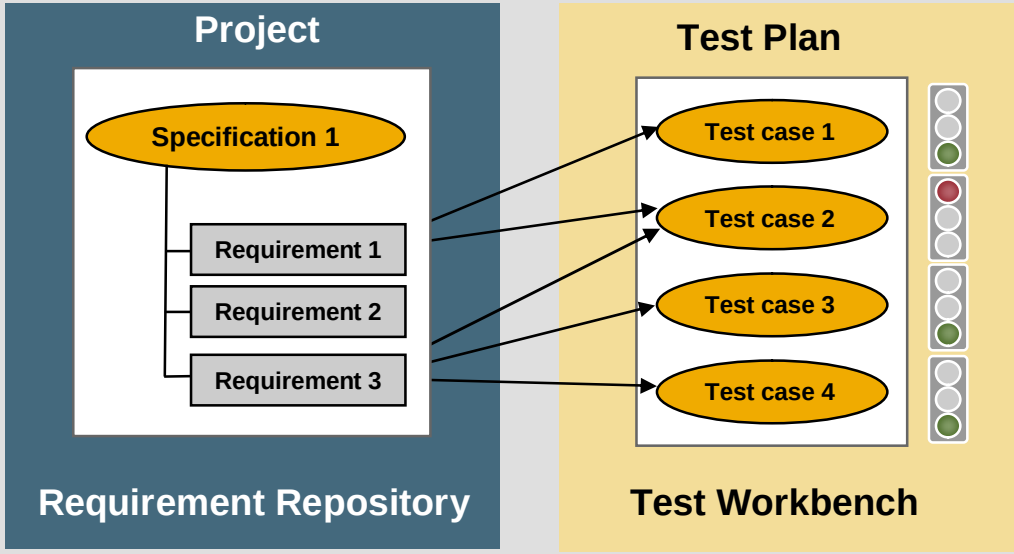
KPI Datasheet:

- **Goal**
  All new software requirements of a program are tested successfully
- **Answered Question**
  How many software requirements are covered by test cases and tested successfully?
- **Metric**

$$KPI = \sum_{i=1}^{\#Req.} \frac{\text{Green test cases for Req. i}}{\text{\# test cases for Req. i} * \text{\# Req.}}$$

- **Measurement/judgment**
  P2D → RTC + 12 months

Overview on KPI Calculation and Aggregation



Test Result

Requirement 1 = 0.5
Requirement 2 = 0
Requirement 3 = 0.67

$$\text{Result} = \frac{0.5+0+0.67}{3} = 0.39$$

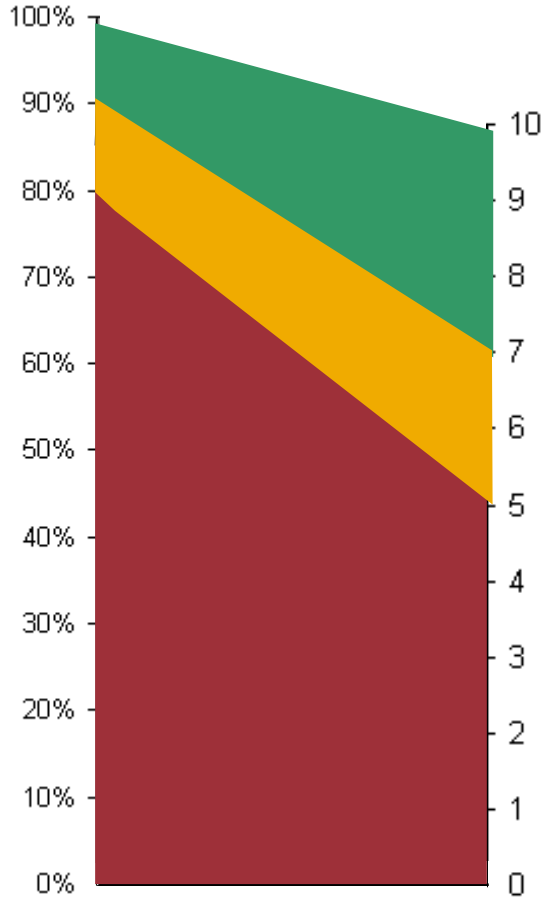
Bewertung jeder KPI und Zuordnung von KPI-Werten zur Index-Skala

Für alle KPIs wurden Regeln definiert, die die KPI–spezifischen Werte einer gemeinsamen Q–Index-Skala von 0 bis 10 zuordnet

KPI-Wertebereiche

KPI–spezifische Bewertung, z.B.:

0 % < x ≤ 80% → rot
80 % < x ≤ 90% → gelb
90 % < x ≤ 100% → grün



Q-Index-Bereiche

Für alle KPIs identisch

0 ≤ x < 5 → rot
5 ≤ x < 7 → gelb
7 ≤ x ≤ 10 → grün

Aggregationsregeln zur Berechnung der KPI-Werte auf verschiedenen Ebenen

Für alle KPIs gilt der **Entwicklungsaufwand** als Gewichtungsfaktor. Alle KPIs werden im Index gleich gewichtet.

Beispiel einer Aggregation für eine KPI:

PTU	Program	Entwicklungs-aufwand	Gewicht-ung	Index-Wert	Gewichte-ter Index
PTU 1	Program 1	4.000	80,00%	6,0	4,8
PTU 1	Program 2	1.000	20,00%	8,0	1,6
Summe PTU 1		**5.000**	**100,00%**		**6,40**
PTU 2	Program 3	20.000	16,67%	6.0	1,0
PTU 2	Program 4	100.000	83,33%	8.0	6,67
Summe PTU 2		**120.000**	**100,00%**		**7,67**
SAP Total		**125.000**			**7,62**

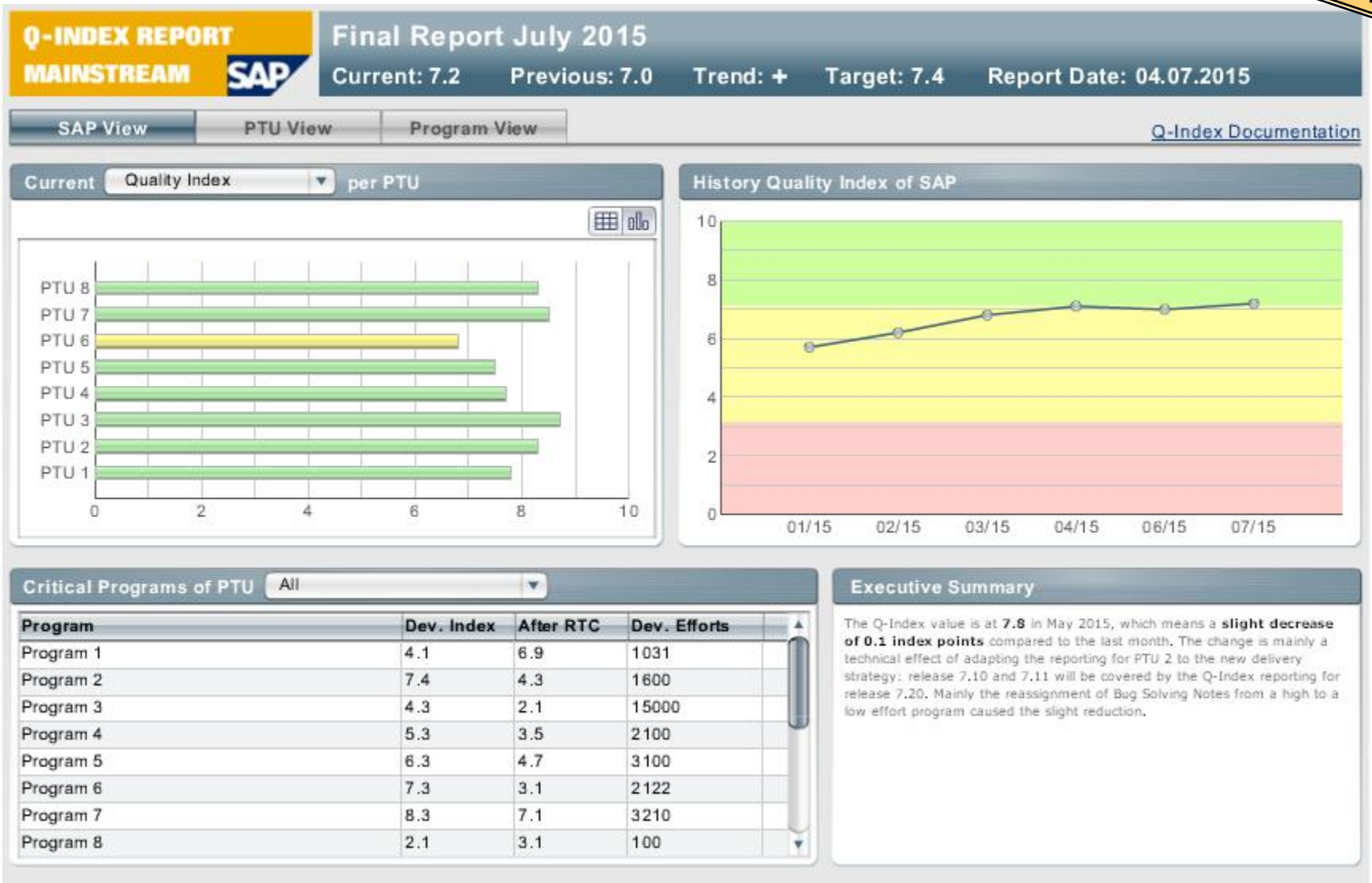
Q-Index-Wert PTU 1

Q-Index-Wert PTU 2

SAP-Q-Index-Wert

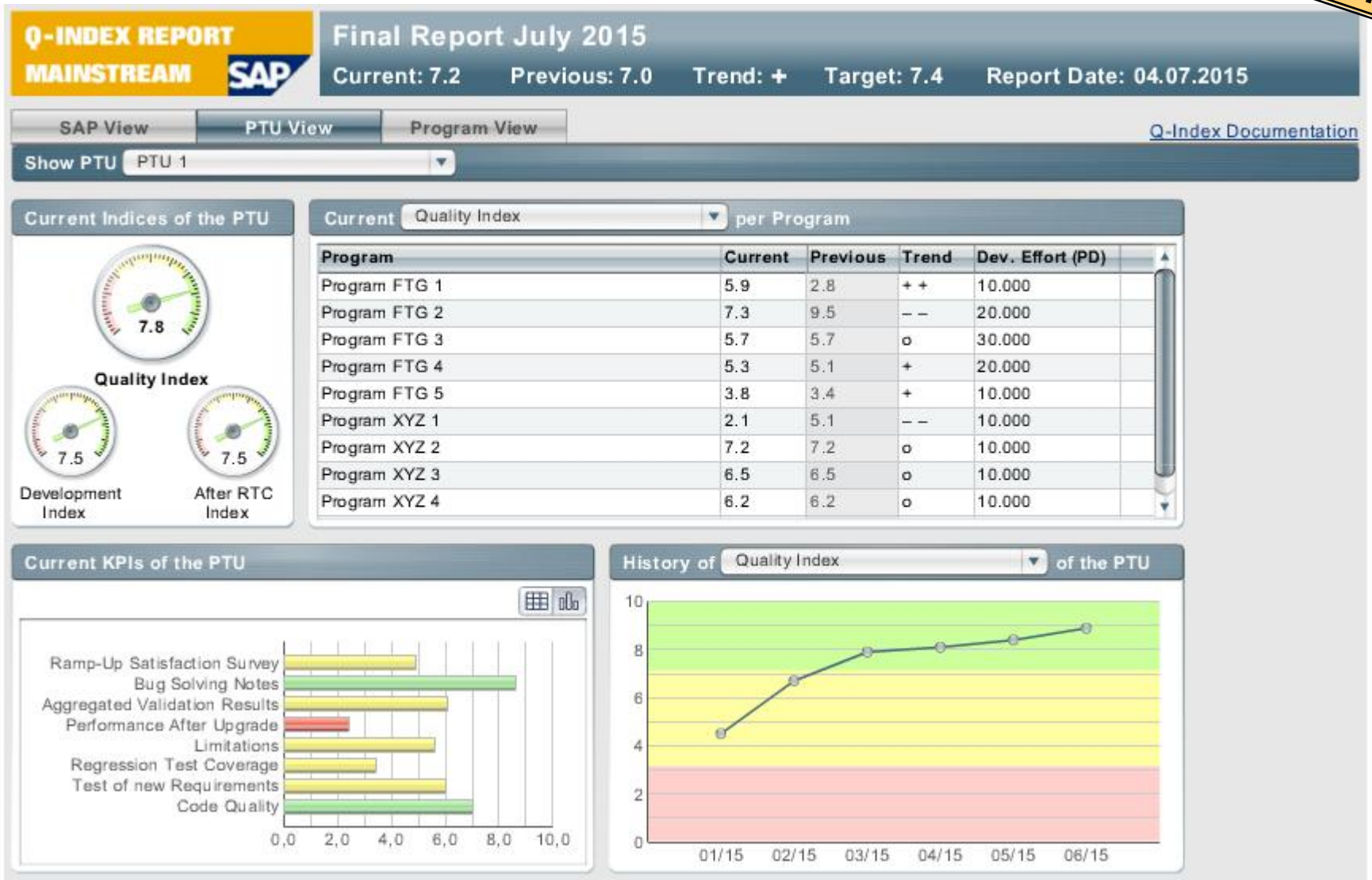
SAP Quality Index User Interface – Sicht SAP gesamt

Beispiel



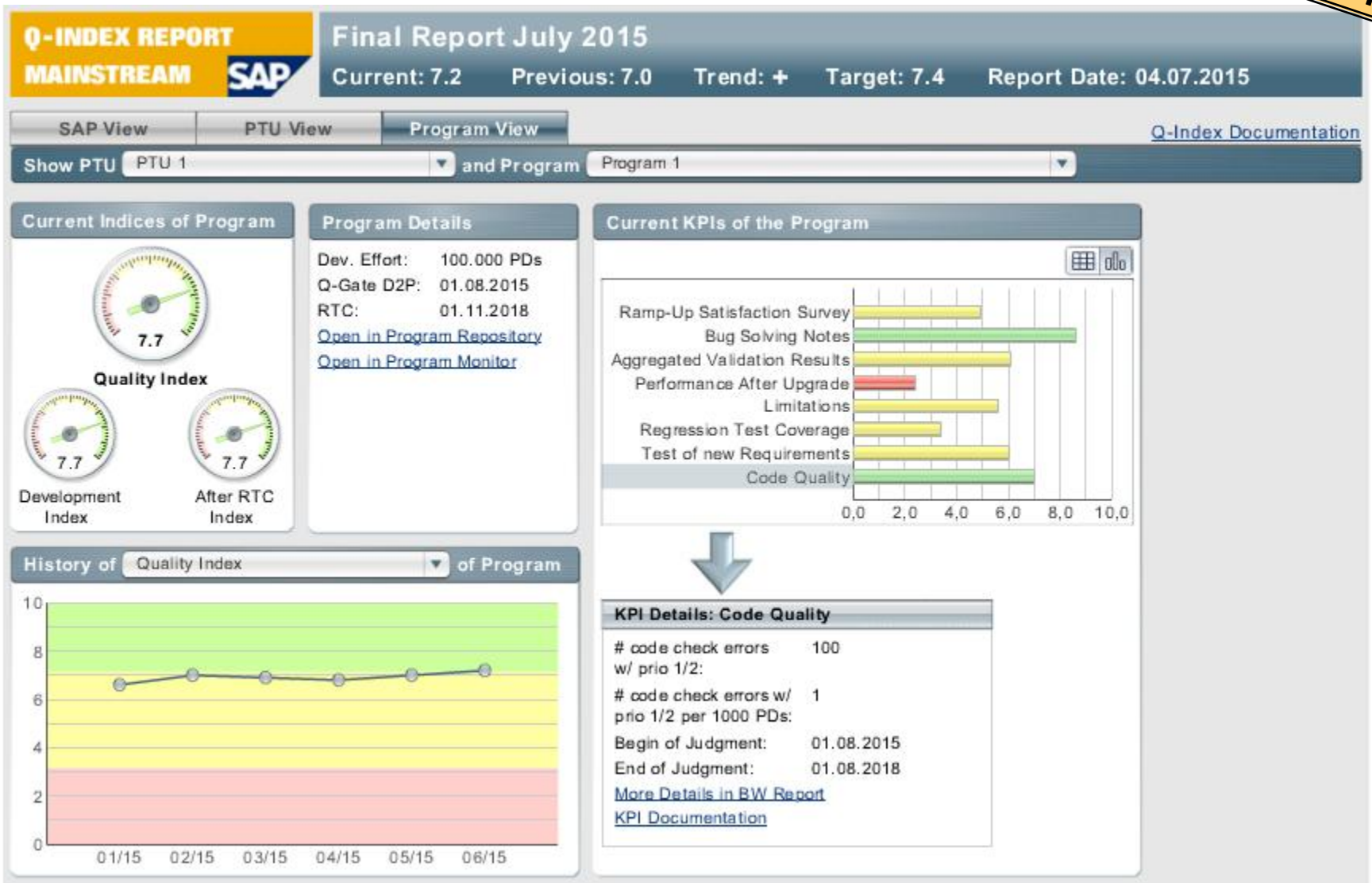
SAP Quality Index User Interface – Sicht je PTU

Beispiel

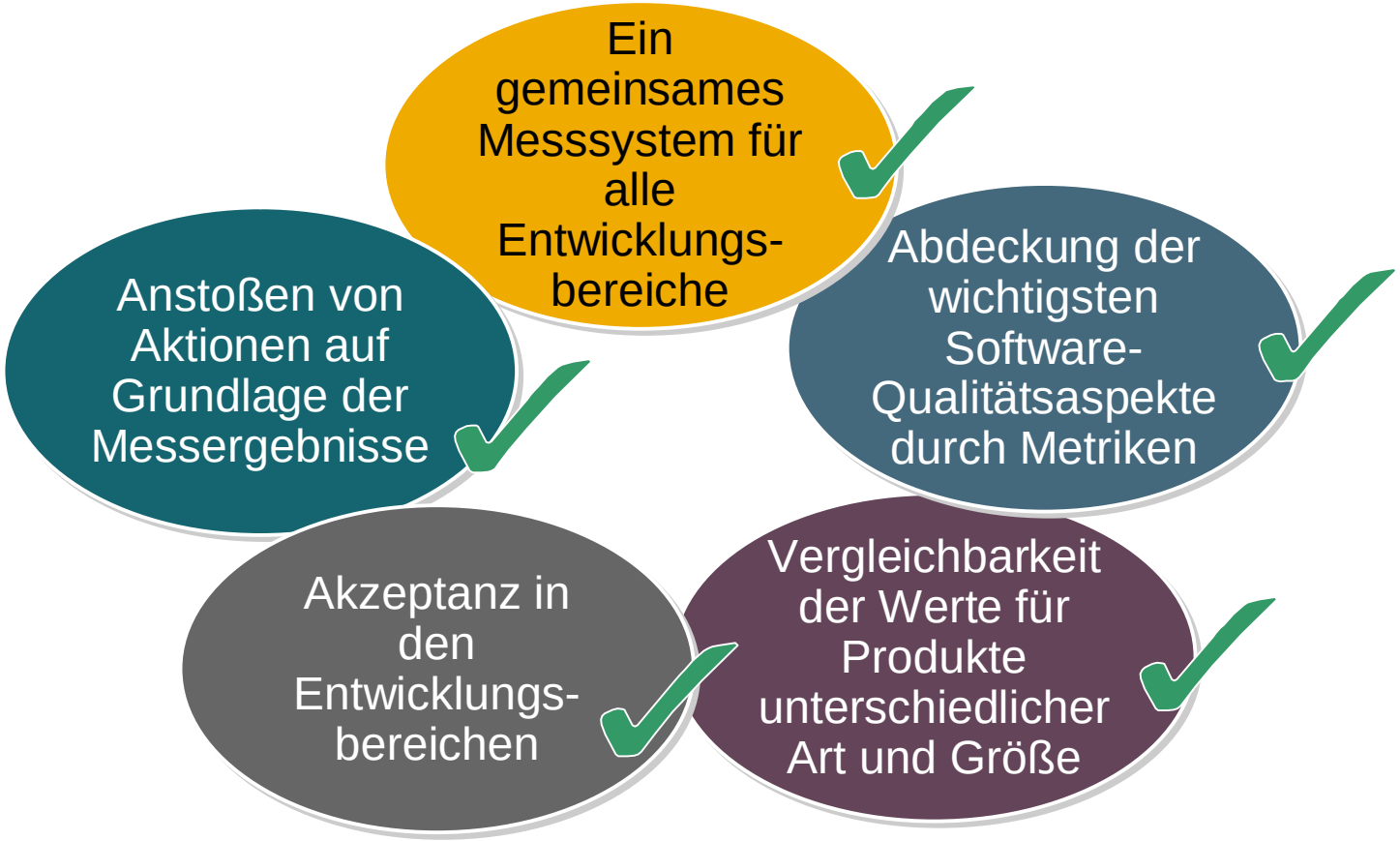


SAP Quality Index User Interface – Sicht je Programm

Beispiel



SAP Quality Index: Zusammenfassung



Agenda

Meine Geschichte bei SAP

Messen von Software Qualität: Herausforderungen

Das Projekt *SAP Quality Index*

Governance und Betrieb des *SAP Quality Index*

Erfahrungen und Ausblick

SAP „Quality Goal Compliance Declaration“ (inkl. toolunterstütztem Ausnahmeprozess)

Beispiel

SAP Quality Goal Compliance Declaration and Exception Requests

Program: Supply Chain Management 2010

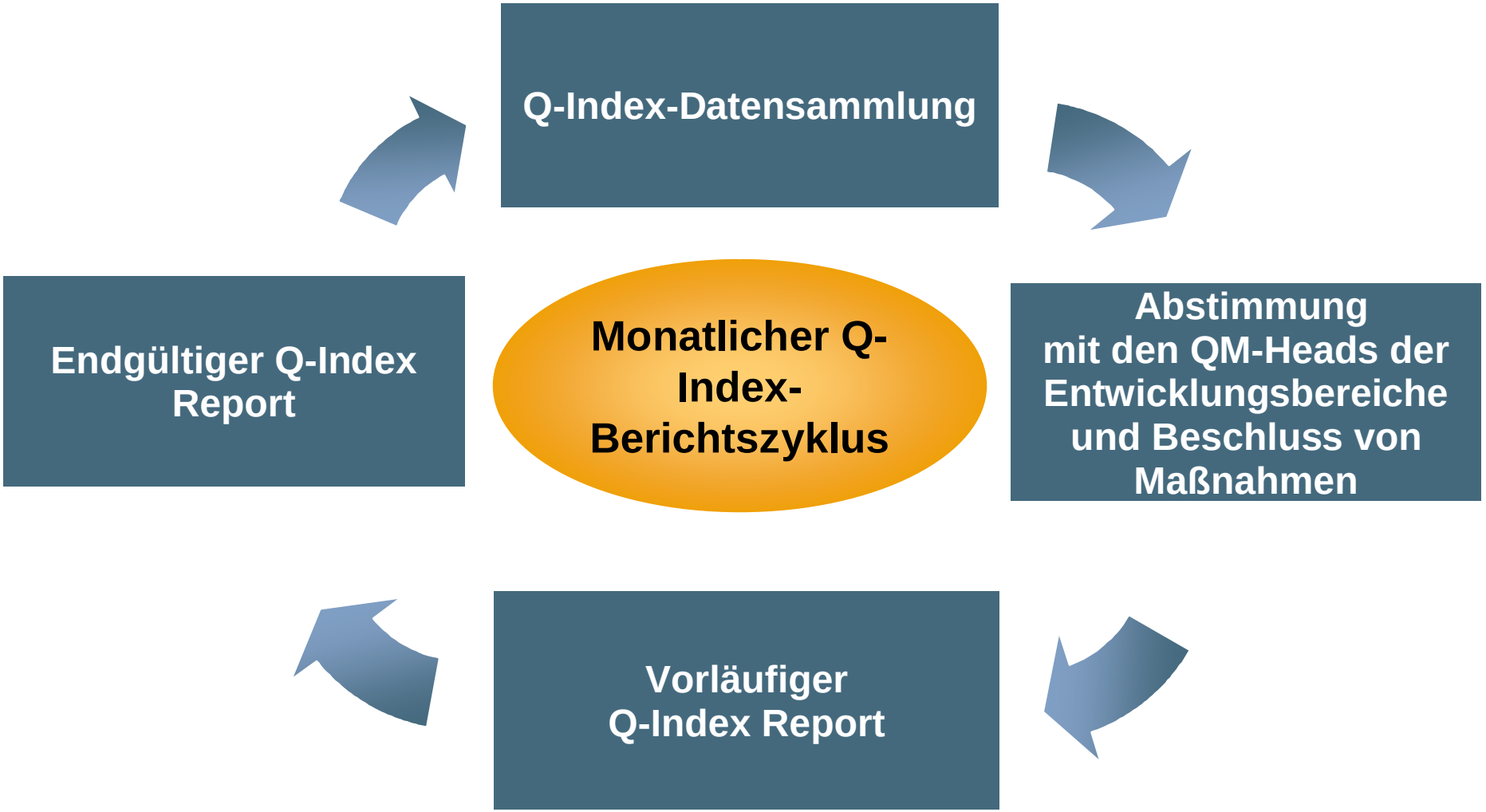
PTU: PTU1

Planned compliance per KPI

Quality Index KPI Name	Do you plan to comply?	Relevant business reason for non-compliance	Additional Information for business reason	Responsible (User ID)	Approver	Approval status	Approved on
Code Quality	yes			Hoffmann	Limboeck	n/a	
Test of New Requirements	yes			Hoffmann	Limboeck	n/a	
Regression Test Coverage	no	Initial shipment of a new product	Top x processes for the next release will be defined.	Hoffmann	Limboeck	Approved	06.08.2008
Limitations	yes			Hoffmann	Limboeck	n/a	
Aggreg. Validation Results	yes			Hoffmann	Limboeck	n/a	
Performance after Upgrade	yes			Hoffmann	Limboeck	n/a	
Bug Solving Notes	yes			Meyer	Limboeck	n/a	
Ramp-Up Survey	yes			Meyer	Limboeck	n/a	
Usability	yes			Meyer	Limboeck	n/a	

Save & Send

Q-Index-Betrieb: Vorläufiger und endgültiger Bericht



Agenda

Meine Geschichte bei SAP

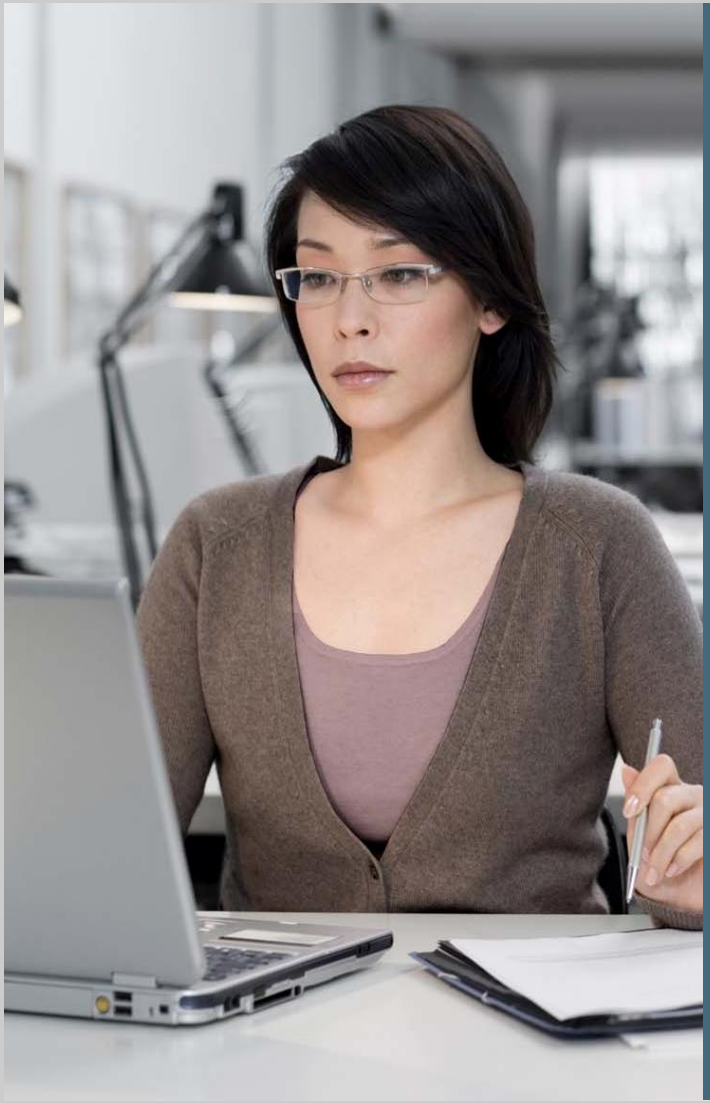
Messen von Software Qualität: Herausforderungen

Das Projekt *SAP Quality Index*

Governance und Betrieb des *SAP Quality Index*

Erfahrungen und Ausblick

Erfahrungen – 1 –



KPI-Definition und Indexberechnung

- *Goal-Question-Metric*-Methode ist sehr nützlich bei der KPI-Definition
- In dynamischem Umfeld sind Produkt-KPIs weitaus beständiger als Prozess-KPIs
- Aggregationsregeln und Zeitachsen: *einfach* ist besser!

Regelmäßiges Reporting

- Erwartungen an Steigerung der Transparenz wurden übertroffen
- Wichtigster Erfolgsfaktor: Sichtbarkeit auf Vorstandsebene
- Hohe Datenqualität ist erforderlich
- Benutzerfreundliche Darstellung erhöht Akzeptanz (vor allem im Senior Management)

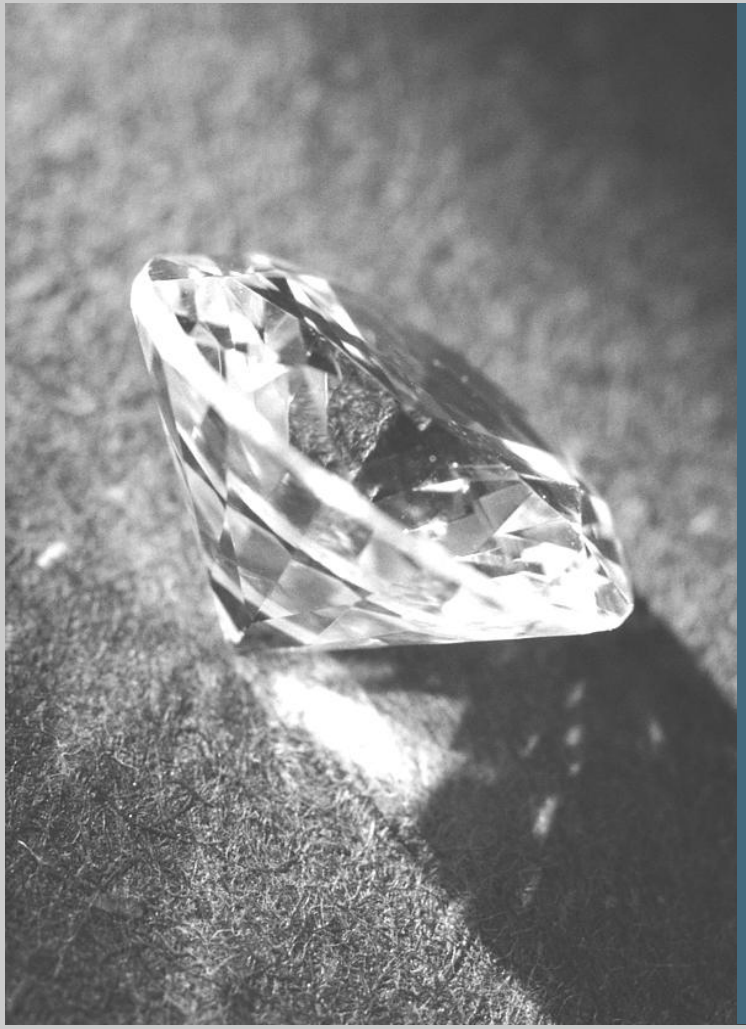
Erfahrungen – 2 –

Governance und Betrieb

- Reports ohne Maßnahmen sind Datenfriedhöfe
- Regelmäßige Überprüfung der KPI-Definitionen ist notwendig, um sicher zu stellen, dass die Qualität optimiert wird und nicht der Index
- Betrieb des Index erfordert hohe Qualifikation
- Betrieb des Index erfordert hohe Kommunikationsfähigkeit

→ **Betrieb des Q-Index benötigt ein starkes Governance Team**

Ausblick



- **Systematische Datenanalyse basierend auf den Q-Index Daten zu einem frühen Zeitpunkt**
- **Erhöhung der Vorhersagbarkeit von Qualität**
- **Integration von Akquisitionen (z.B. SAP Business Objects)**
- **Unterstützung von Technologien neben ABAP und JAVA**
- **Separate KPIs für On-Demand Produkte**

Danke für Ihre Aufmerksamkeit!

Anhang

SAP aktuell

Jahresumsatz der SAP AG 2007: 10,25 Mrd. Euro

- Mehr als 76.000 Unternehmen setzen SAP ein
- Mehr als 25 Branchenlösungen
- 51.863 SAP-Mitarbeiter (Stand: September 2008)

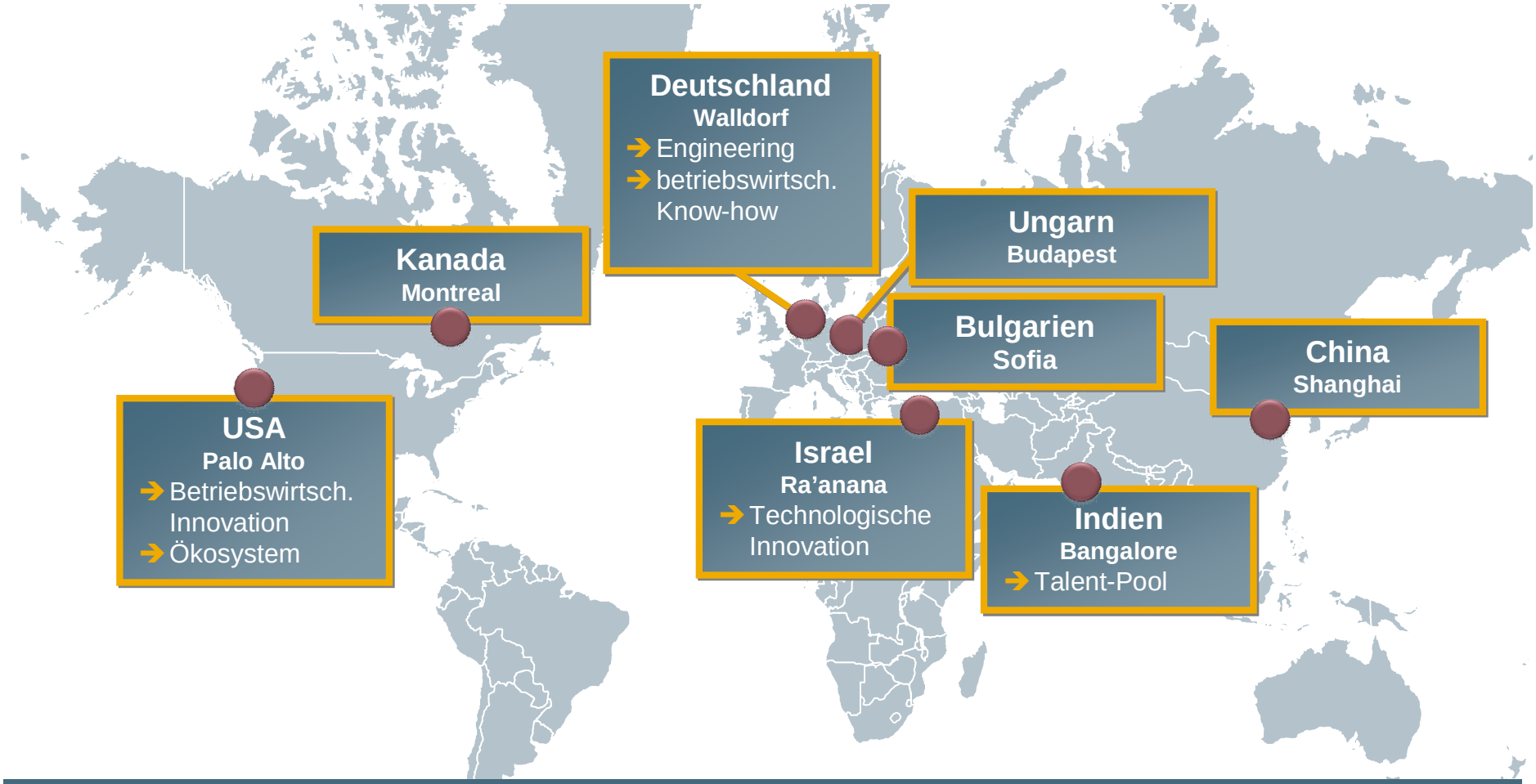
12 Millionen Anwender in über 120 Ländern nutzen SAP-Lösungen zur

- Integration von Geschäftsprozessen
- Stärkung der Wettbewerbsfähigkeit
- schnelleren Rendite bei niedrigeren Systemgesamtkosten

Einzigartiges Partnernetzwerk

- Mehr als 1,3 Millionen Community Mitglieder (SDN und BPX)
- Mehr als 1.800 zertifizierte Partnerlösungen für SAP NetWeaver
- 15 Industry Value Networks

SAPs Forschungs- und Entwicklungsorganisation:
ca. 15.000 Entwickler an 8 Entwicklungsstandorten



Alle Standorte arbeiten gemäß einem Qualitätsmanagement-System (ISO 9001-zertifiziert)

KPI: Code Quality

Beispiel

KPI Datasheet

- **Goal**

  Ensure formal code quality by automated code checks

- **Answered Question**

  How many code check errors remain open relatively to the completed development effort?

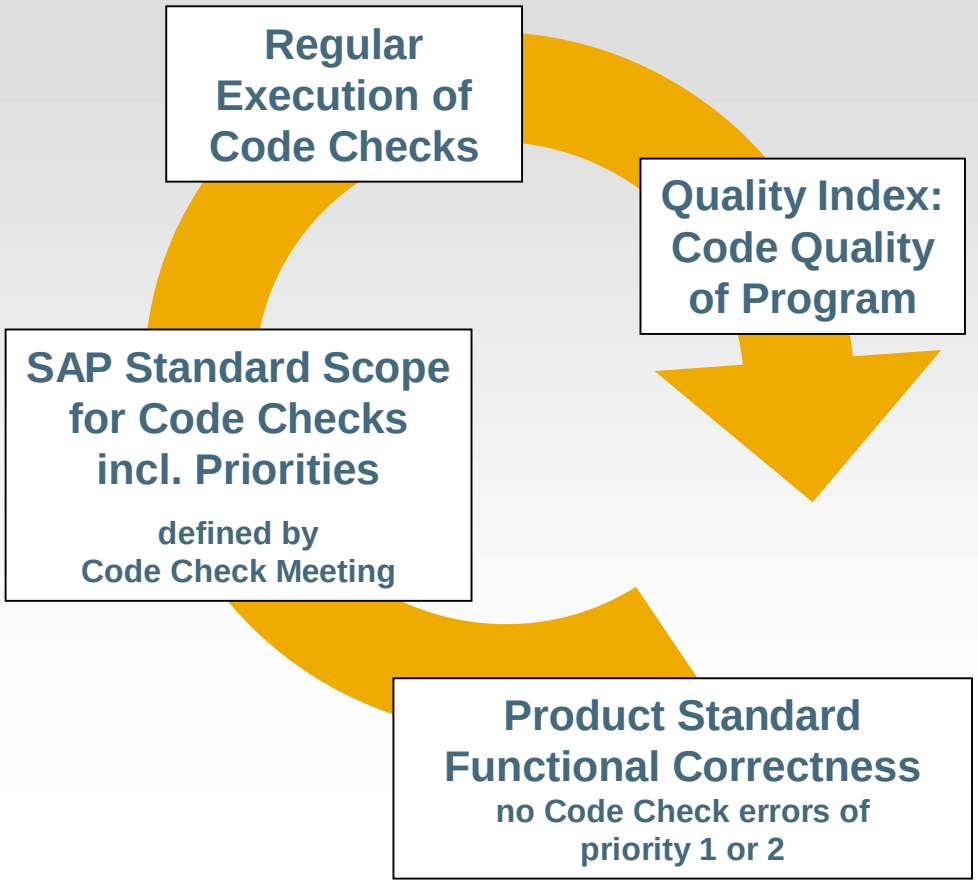
- **Metric**

  KPI = Total number of open prio1/2 static Code Check messages from all Code Check frameworks per person days development effort measured in 1/1000

- **Measurement/Judgment**

  P2D → D2P → RTC + 12 months

From Functional Correctness to Code Quality



KPI: Test of New Requirements

Beispiel

KPI Datasheet

- **Goal**

  All new software requirements of a program are tested successfully

- **Answered Question**

  How many software requirements are covered by test cases and tested successfully?

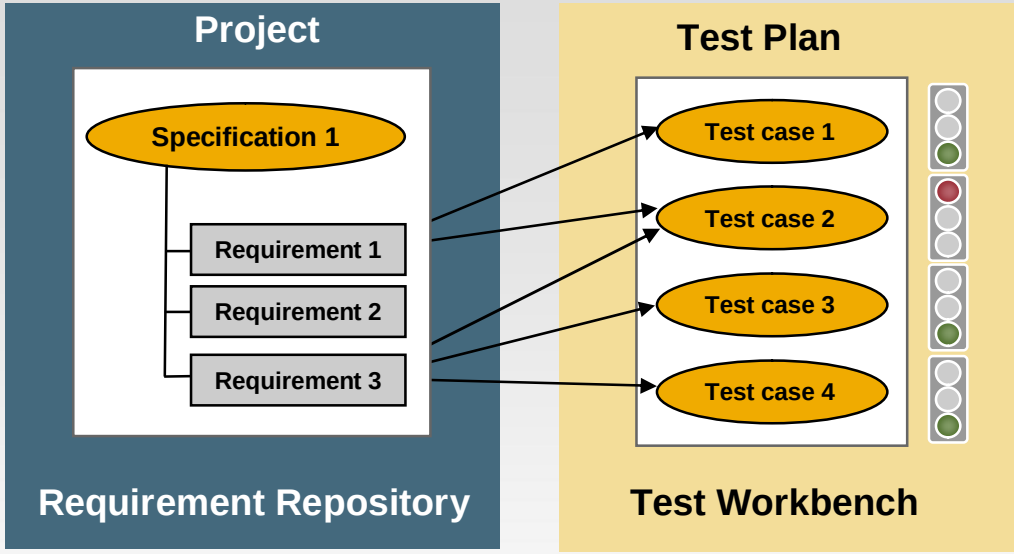
- **Metric**

$$KPI = \sum_{i=1}^{\#Req.} \frac{\text{Green test cases for Req. i}}{\text{\# test cases for Req. i} * \text{\# Req.}}$$

- **Measurement/judgment**

  P2D → RTC + 12 months

Overview on KPI Calculation and Aggregation



Test Result

Requirement 1 = 0.5

Requirement 2 = 0

Requirement 3 = 0.67

$$\text{Result} = \frac{0.5+0+0.67}{3} = 0.39$$

KPI: Regression Test Coverage

Beispiel

KPI Datasheet

- **Goal**

  Ensure stability of previously delivered product (code) with regression tests

- **Answered Question**

  Are regression tests successfully running for the top X processes, process steps?

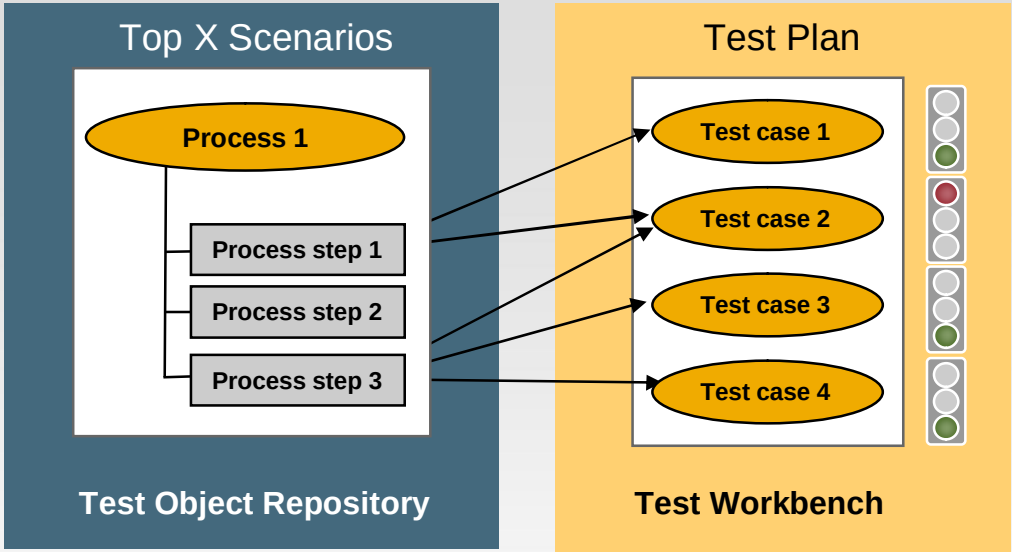
- **Metric**

$$KPI = \sum_{i=1}^{\#Pr.step} \frac{\text{Green test cases Proc.step i}}{\text{\# Test cases Proc.step i} * \text{\# Proc.steps}}$$

- **Measurement/Judgment**

  P2D → RTC + 12 months

Overview on KPI Calculation and Aggregation



Test Result

Requirement 1 = 0.5

Requirement 2 = 0

Requirement 3 = 0.67

$$\text{Result} = \frac{0.5+0+0.67}{3} = 0.39$$

KPI: Limitations

Beispiel

KPI Datasheet

- **Goal**

  Provide transparency about functional restrictions

- **Answered Question**

  Do shipped products have functional restrictions?
  Do these functional restrictions have an impact on SAPs reliability?

- **Metric**

  Prio-A limitations are the leading metric factors. No Prio-A limitations allowed at all.

- **Measurement/Judgment**

  RTC → RTC + 12 months

Overview on KPI Calculation and Aggregation

PTU	Program	Number of Development Efforts	Index	Prio A	Prio B	Prio C
Overall Result		112.000	9,3	9	8	17
PTU X	Result	10.000	2,5	7	8	16
	Program A	2.000	10,0			
	Program B	7.500	0,0	7	8	5
	Program C	500	10,0			
PTU Y	Result	2.000	7,5	2		1
	Program D	500	0,0	2		
	Program E	1.500	10,0			1
PTU Z	Result	100.000	10,0			
	Program F	100.000	10,0			

KPI: Aggregated Validation Results

Beispiel

KPI Datasheet

- **Goal**

  Measure and provide transparency on Product Standard Compliance before shipment

- **Answered Question**

  Degree of compliance with validated product standards.
  Do products fulfill the requirements of Product Standards for shipment?

- **Metric**

$$\text{AvgValiRes ult} = \frac{(\#RED*0+\#YELLOW*5+\#GREEN*10)}{(\#RED+\#YELLOW+\#GREEN)}$$

- **Measurement/Judgment**

  RTC → RTC + 12 months

Overview on KPI Calculation and Aggregation

PTU A	Dev. efforts		Validation results	7,5
Program A1	12627	5,6	59 green / 9 yellow / 46 red Scen.	5,6
Program A2	46131	8,3	52 green / 11 yellow / 6 red Scen.	8,3
Program A3	6581	5,0	1 green / 2 yellow / 1 red Scen.	5,0

Calculation:

- Calculation on program level over all available average validation results (calculation model: each RED=0 pts; each YELLOW=5 pts; each GREEN=10pts).

Additional remarks:

- Validation results of all validation projects are aggregated, i.e. the result of a re-validation or SP-validation does not overwrite the result of the original release validation

KPI: Performance After Upgrade

KPI Datasheet

- **Goal**

  Ensure stable performance of top X scenarios, processes, and process steps after upgrade, so customers shouldn't need new hardware.

- **Answered Question**

  How is the performance compared to last release?

- **Metric**

  Index will be calculated for a program out of the four Performance indicators: Data Base Load, Memory Consumption, CPU Consumption, Network Load.

- **Measurement/Judgment**

  P2D → RTC + 12 months

Overview on KPI Calculation and Aggregation

PTU	Program	Top X process	Performance after Upgrade	Data Base Load	Memory Consumption	CPU Comsumption	Network Load	Dev. Effort	Index
PTU x	PTU Result							6.915	5,2
	Program 1		-23,76%					1.615	9,4
		Process 1A	-16,96%	-32,17%	-2,78%	-15,93%	n/a		
		Process 1B	-22,51%	-59,53%	-21,22%	-9,98%	0,70%		
		Process step xy	-7,73%	-3,37%	3,64%	-23,46%	n/a		
		Scenario 1X	-53,18%	-44,33%	-71,43%	-43,78%	n/a		
		Scenario 1Y	-18,44%	-8,05%	-24,50%	-22,77%	n/a		
	Program 2		-17,43%					1.300	8,8
	Program 3		36,56%					4.000	2,3
PTU y	PTU Result							2.000	6,0
	Program X		5,60%					2.000	6,0
		Scenario X	20,03%	36,30%	45,00%	-52,00%	50,80%		
		Scenario Y	-23,30%	3,30%	3,70%	-78,80%	-21,40%		
		Process A	-0,46%	96,00%	36,40%	-89,20%	-45,02%		
		Process B	26,12%	34,50%	15,78%	-21,89%	76,10%		

Aggregation on Program level:

Calculate the mean value of all four performance indicators for each top x process; then calculate the performance after upgrade KPI for the whole program as a mean value of those top x processes.

KPI: Ramp-Up Satisfaction Survey

Beispiel

KPI Datasheet

- **Goal**

  Get a comprehensive understanding of the customer's satisfaction with the SAP solutions during Ramp-Up

- **Answered Question**

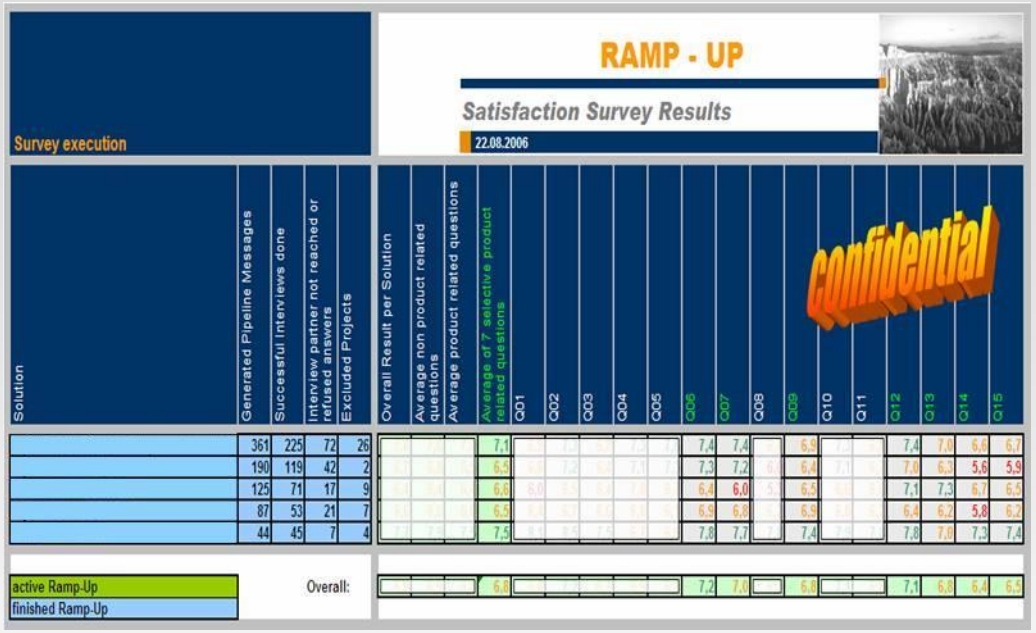
  How does the Ramp-Up customers rate the quality, documentation, cost-performance ration etc. of our solutions?

- **Metric**

  Ramp-Up customer satisfaction survey = Average of 8 product related questions

- **Measurement/Judgment**

  P2R → RTC + 24 months



Only the product related questions of the ramp-up satisfaction survey are taken into account, for example:

Q06: How would you rate your overall satisfaction with SAP [solution]?

Q07: How would you rate the quality and reliability of this SAP [solution]?

Q09: How would you rate the usability and the ease of learning of this SAP [solution]?

Q14: How would you rate the cost-performance ratio of the software?

KPI: Bug-Solving Notes

KPI Datasheet

- **Goal**

  Provide transparency about the quality perceived by our customers for the SAP products after they have been released to customers

- **Answered Question**

  Do shipped products have defects encountered by customers? How many customers are impacted due to those defects and hence affecting SAP's reliability?

- **Metric**

  Number of Bug Solving Notes reported in a month per 1000 days of Dev Effort for that Program; e.g. Metric = (Number of Bug Solving Notes / Total Dev Effort) * 1.000

- **Measurement/Judgment**

  P2R → RTC + 36 months

Overview on KPI Calculation and Aggregation

Program (LinkProgrRep)	Dev. PDs	Value		Index	Begin of Judgment	Link to BW
				8,5		
Program A	12530	5,43	68 notes / 38 customers w/ BSNotes	6,3	02.08.2007	Details
Program B	45468	0,00	0 notes / 7 customers w/ BSNotes	10,0	14.12.2007	Details
Program C	95776	0,42	40 notes / 104 customers w/ BSNotes	8,9	26.06.2006	Details
Program D	3285	0,00	0 notes / 0 customers w/ BSNotes	10,0	30.06.2005	Details

Different calculation rules on program and on PTU level

Program level:

- Static indexation rule of 0-10 points
- Three quality indicator zones, Red, Yellow and Green at program level

Aggregation on PTU level:

- Weighting factor: Amount of development effort
- Using this normalization factor leads to a dynamic indexation from 0 to 10 pts and provides comparability and distinction between programs of different size.
- Aggregation through weighted evaluation of the KPI values on program level provides the PTU values.

KPI: Usability

Beispiel

KPI Datasheet

- **Goal**

  Achieve high-quality user experiences with SAP products for our customers

- **Answered Question**

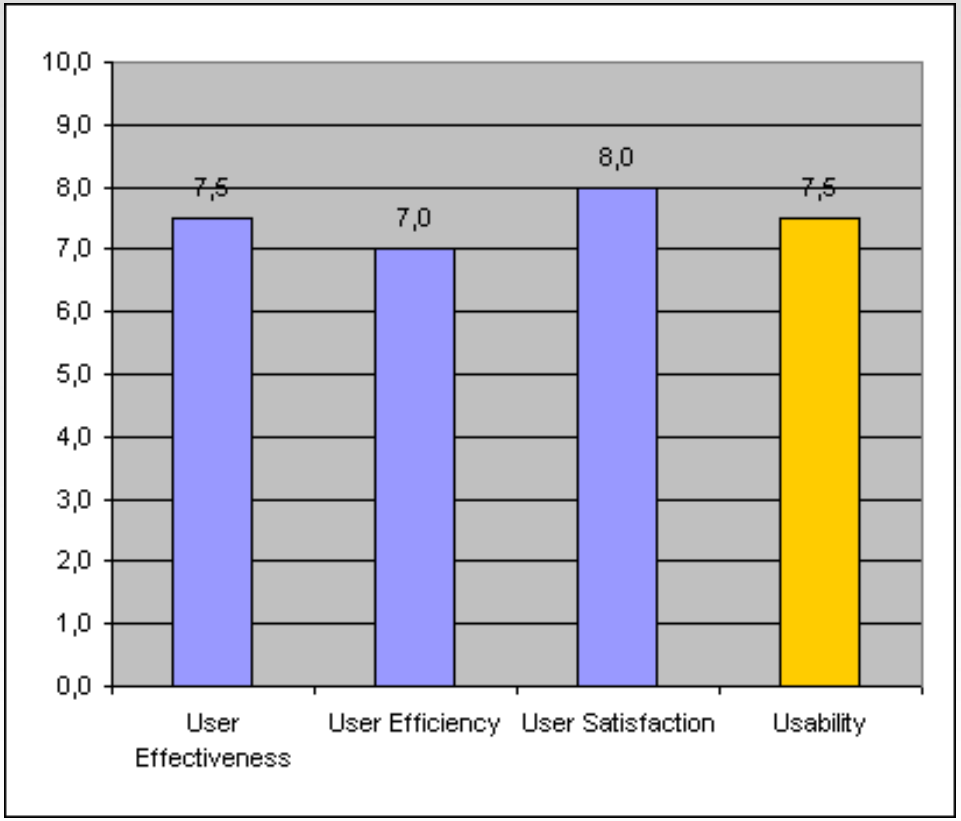
  How well do targeted end-users perform representative tasks, in terms of task-completion rates and times, and how satisfied are they with the applications?

- **Metric**

- **Measurement/Judgment**

  At intervals that correspond to real application lifecycles. Benchmark tests on only a subset of applications.
  RTC → RTC + 12 months

Overview on KPI Calculation and Aggregation



Remark: Definition of user effectiveness, user efficiency, and user satisfaction according to ISO 9241