Meten is weten, of toch niet?

@EricBouwers

Waarom zouden we software doormeten?

‘You can’t control what you can't measure.’

DeMarco, Tom. *Controlling Software Projects: Management, Measurement and Estimation*. ISBN 0-13-171711-1.

Waarom zouden we software doormeten?

'You can't improve what you can't measure.'

DeMarco, Tom. *Controlling Software Projects: Management, Measurement and Estimation*. ISBN 0-13-171711-1.

Software metingen worden gebruikt voor:

- Het schatten van kosten en inspanning
- Productiviteit metrieken en modellen
- Data collectie
- Betrouwbaarheid modellen
- Performance evaluaties en modellen
- Structuur en complexiteit metrieken
- Capability-maturity assessments
- Management door metrieken
- Evaluaties van methodieken en tooling
- Kwaliteitsmodellen en -metingen

Fenton et. al., *Software Metrics: a rigorous and practical approach*. ISBN 0534954251

Welke software metrieken gebruiken jullie?

(Software) Metingen

Wat is meten?

*'Formally, we define **measurement** as a mapping from the empirical world to the formal, relational world. '*

*'A **measure** is the number or symbol assigned to an entity by this mapping in order to characterize an attribute'*

Fenton et. al., *Software Metrics: a rigorous and practical approach*. ISBN 0534954251

Entiteit

Attribuut

Mapping

Meting

Entiteit

Product:

- **Specificaties, Architecture diagrammen, Designs, Code, Test Data, ...**

Proces:

- **Maken van specificaties, Gedetailleerd ontwerp, Testen, ...**

Resources:

- **Personeel, Teams, Software, Hardware, Kantoren, ...**

Fenton et. al., *Software Metrics: a rigorous and practical approach*. ISBN 0534954251

Attributen

Extern

Bruikbaarheid,
Betrouwbaarheid

Intern

Volume,
Gestructureerdheid,
Functionaliteit

Mapping

Definition Checklist for Source Statement Counts

Definition name: *Physical Source Lines of Code (basic definition)*

Date: *8/7/92*

Originator: *SEI*

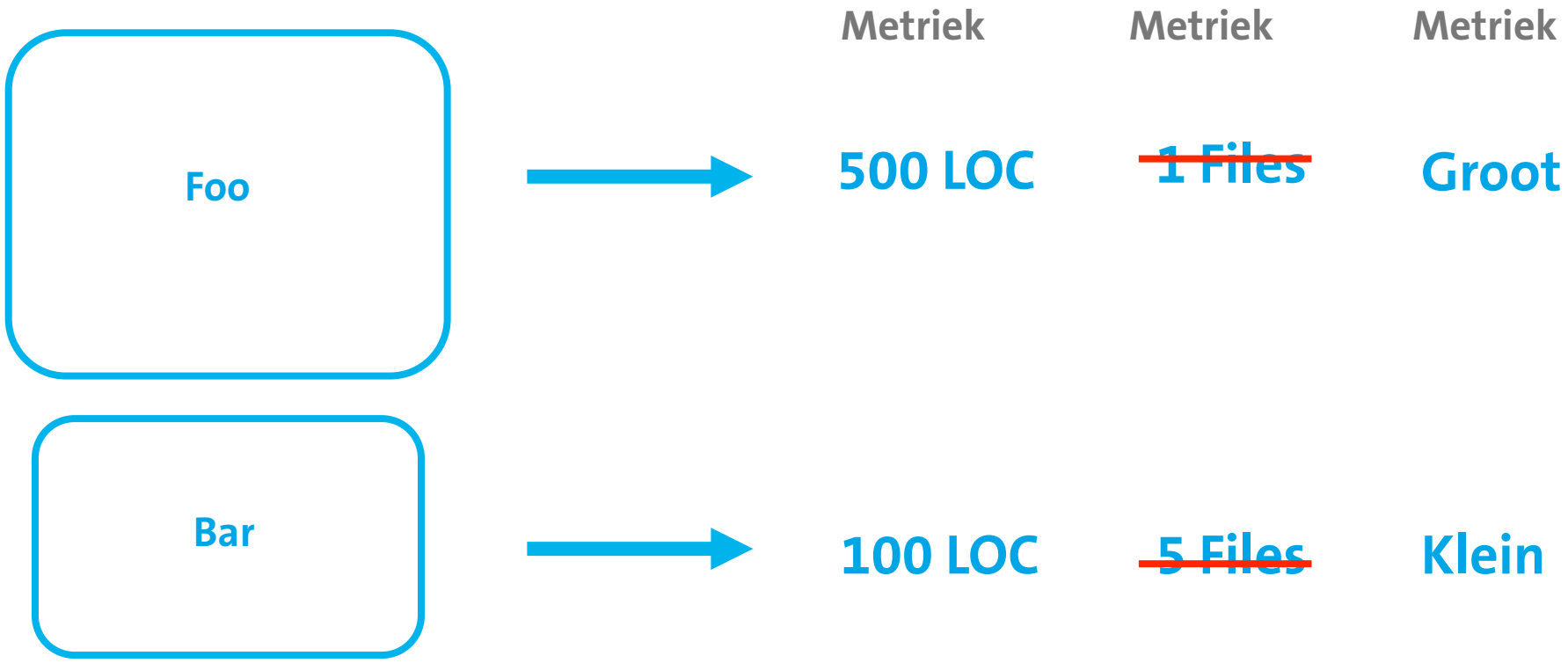
Measurement unit: **Physical source lines** [4] **Logical source statements** [ ]

	Statement type	**Definition** [4] **Data array** [ ]		**Includes**	**Excludes**
	When a line or statement contains more than one type, classify it as the type with the highest precedence.				
1	Executable	**Order of precedence ->**	1	4	
2	Nonexecutable				
3	Declarations		2	4	
4	Compiler directives		3	4	
5	Comments				
6	On their own lines		4		4
7	On lines with source code		5		4
8	Banners and nonblank spacers		6		4
9	Blank (empty) comments		7		4
10	Blank lines		8		4
11					

Park, Rober E., *Software Size Measurement: A Framework for Counting Source Statements,* Technical Report CMU/SEI-92-TR-020

Representation Condition

Attribuut: Volume



Meetschalen

Type	Toegestane operaties	Voorbeelden
Nominaal	= , ≠	A, B, C, D, E
Ordinaal	= , ≠, < , >	Small, large
Interval	= , ≠, < , > , + , -	Start date
Ratio	All	LOC
Absoluut	All	-

Samenvatting concepten

Attribuut
(Lengte)

Entiteit
(Kind)

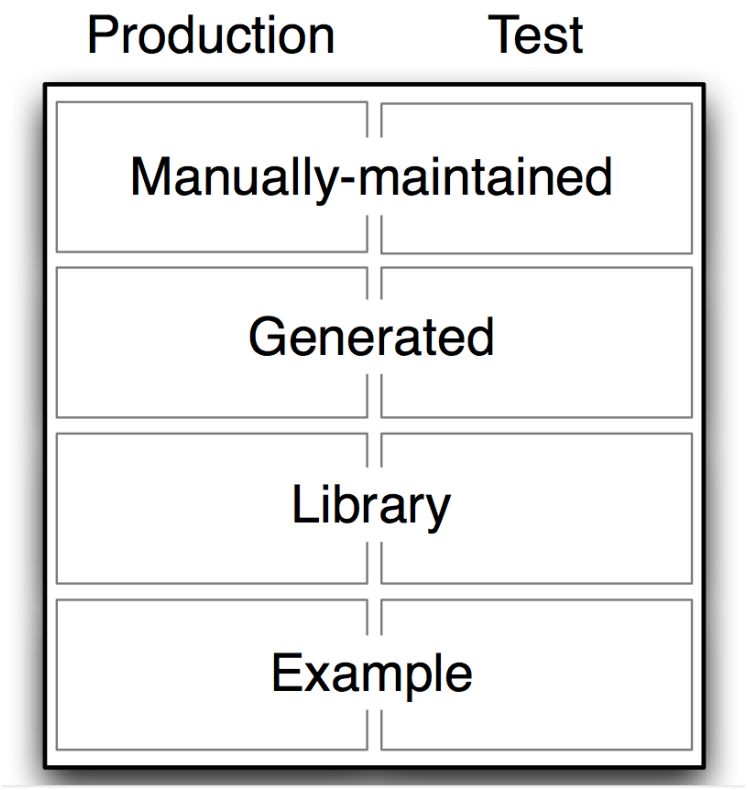
Metriek
(cm)

Mapping
(Voeten op de grond)

Waarom maakt dit uit?

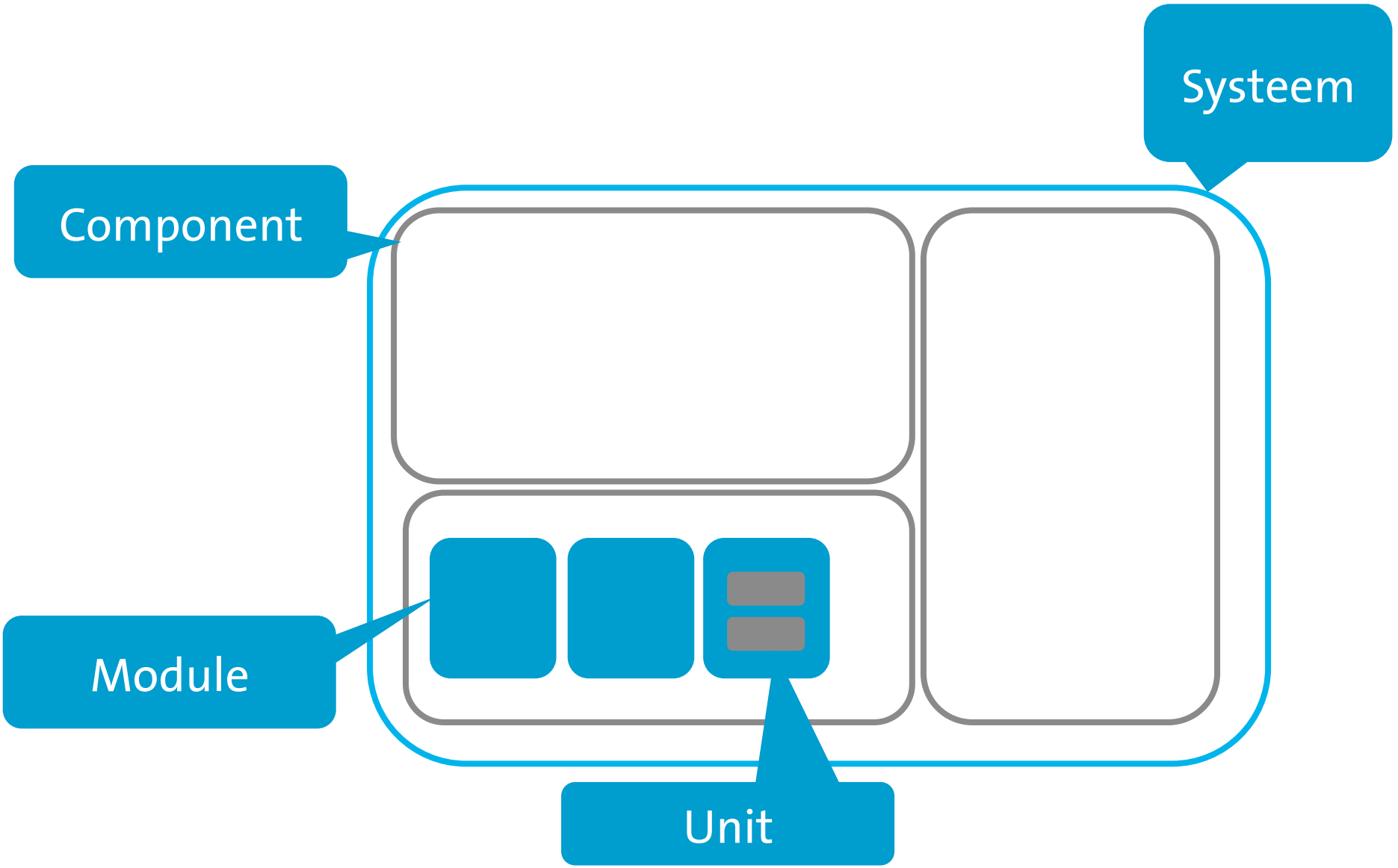
Het beïnvloedt wat je zou willen ...

Entiteit: systeem



Alves. *Categories of Source Code in Industrial Systems,* ESEM 2011: 335-338

Het maakt uit wie het wil weten ...

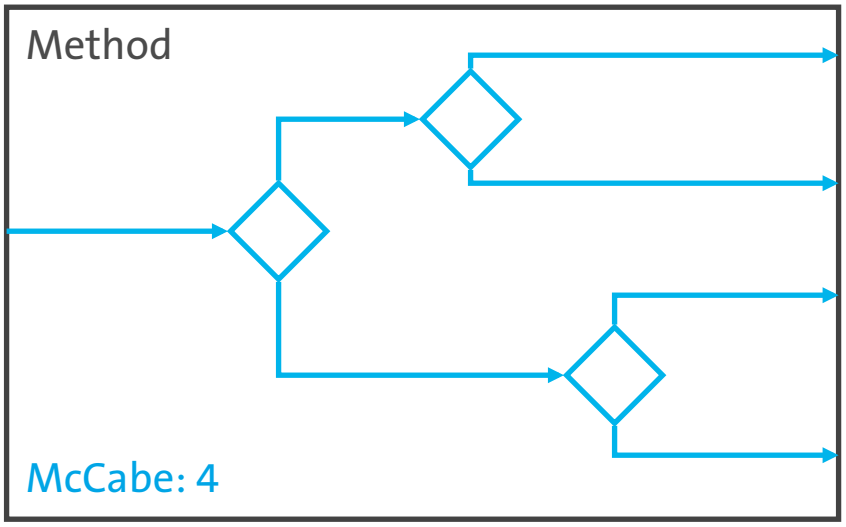


Aggregatie oefening

Van Unit naar Systeem

Unit meting:
T. McCabe, *IEEE Transactions on Software Engineering*, 1976

- Academisch: aantal onafhankelijke paden in een methode
- Intuïtief: aantal beslispunten in een methode
- Realiteit: het aantal if (for, while,...) statements



Beschikbare data

Voor 4 projecten, per unit:

- **Lines of Code**
- **McCabe complexity**

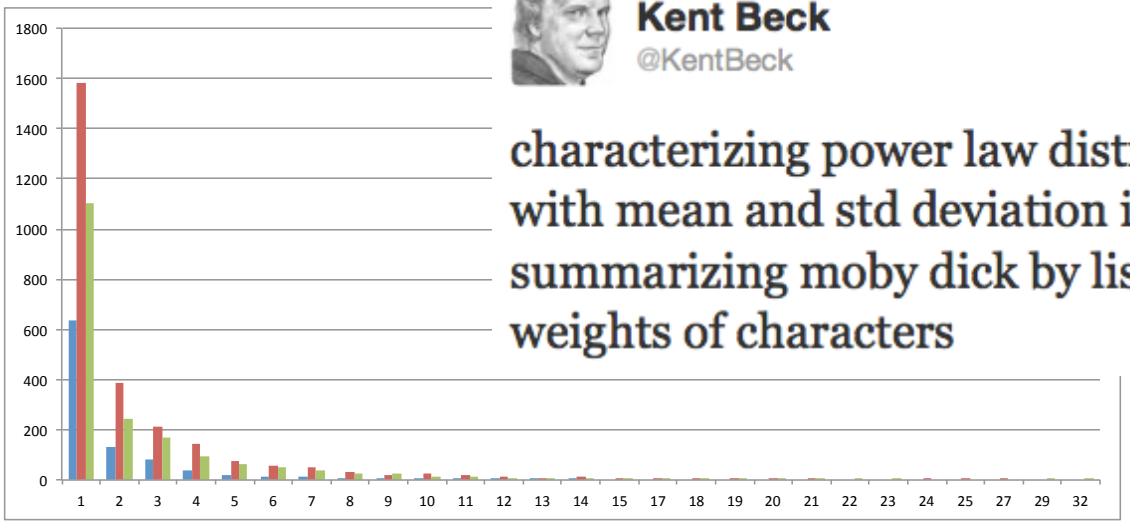
In welk systeem is het testen van de units over het algemeen uitdagender?

Optie 1: Optellen

	Crawljax	GOAL	Checkstyle	Springframework
Totaal McCabe	1814	6560	4611	22937
Totaal LOC	6972	25312	15994	79474
Ratio	0,260	0,259	0,288	0,288

Optie 2: Gemiddelde

	Crawljax	GOAL	Checkstyle	Springframework
Gemiddelde McCabe	1,87	2,45	2,46	1,99



Kent Beck
@KentBeck

Following

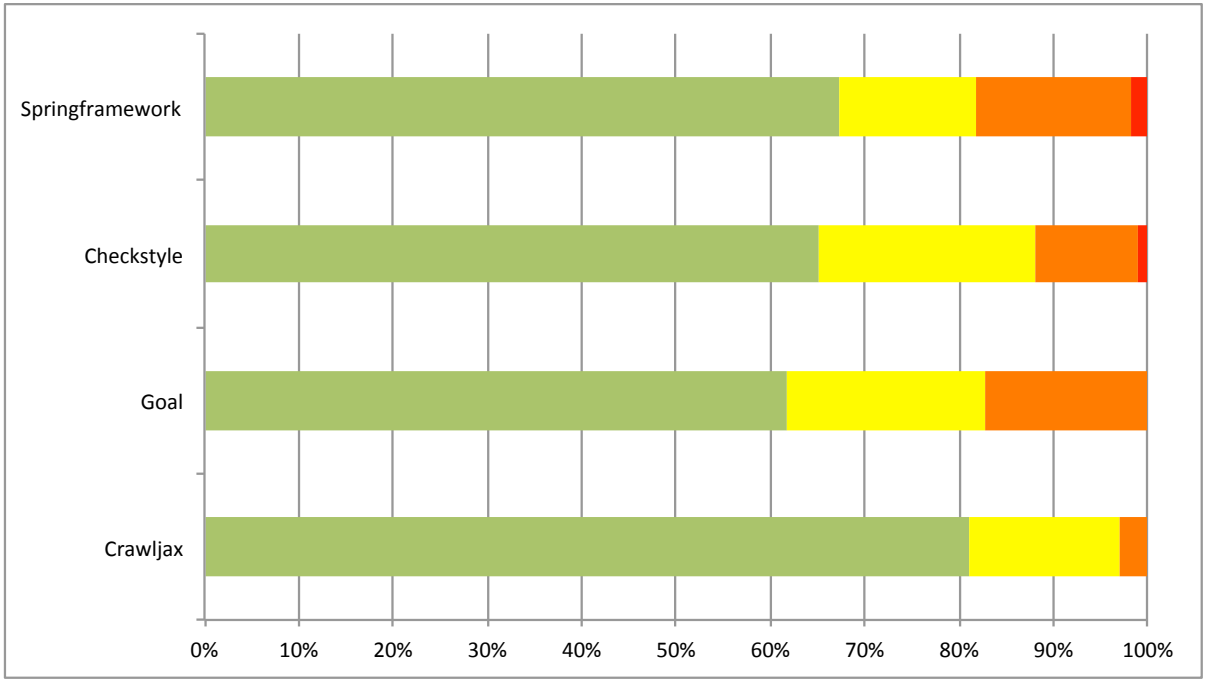
characterizing power law distributed data with mean and std deviation is like summarizing moby dick by listing the weights of characters

Optie 3: Kwaliteitsprofiel

Cyclomatic complexity	Risk category
1 - 5	Low
6 - 10	Moderate
11 - 25	High
> 25	Very high

Tel Lines of Code per categorie

Lines of code per risk category			
Low	Moderate	High	Very high
70 %	12 %	13 %	5 %



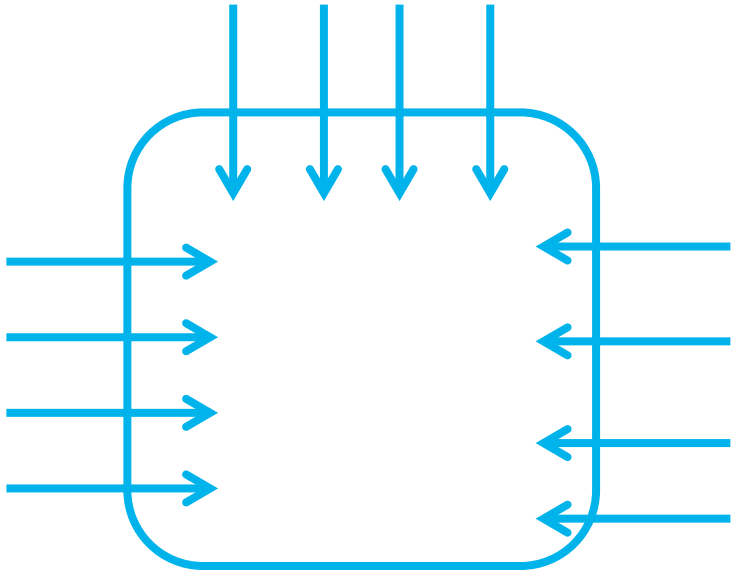
Belangrijke zaken voor metingen

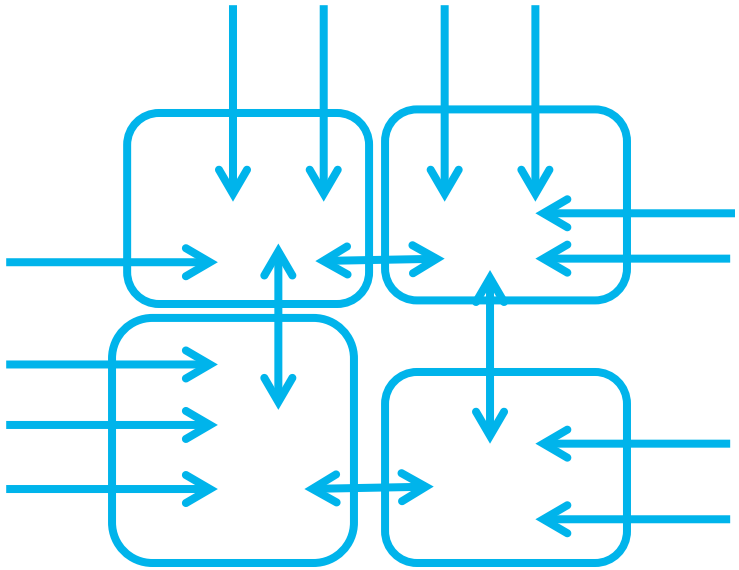
- Volume
- Uitlegbaarheid
- Distributie

Valkuilen van het meten

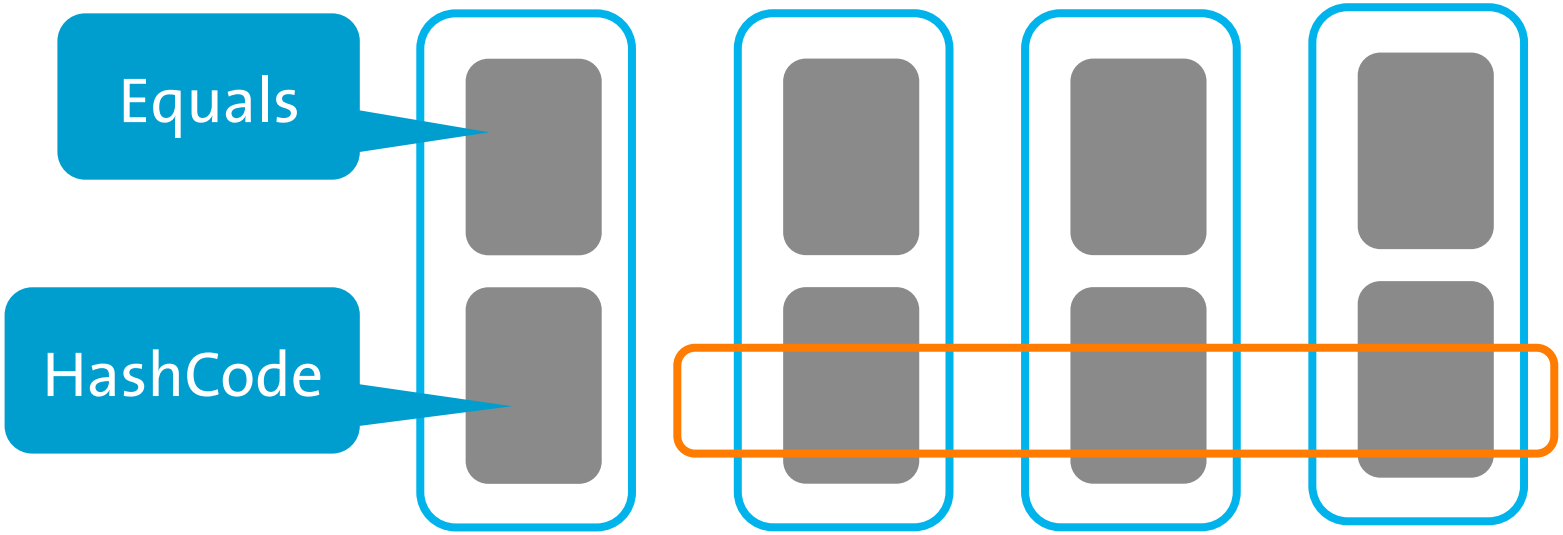
One-track metric

Alleen kijken naar het volume



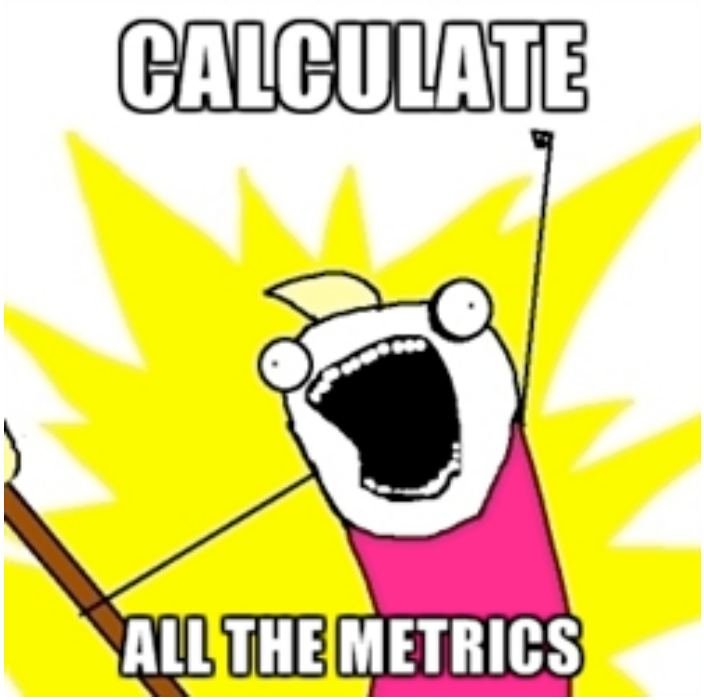


Combinatie van metingen levert meer inzichten op



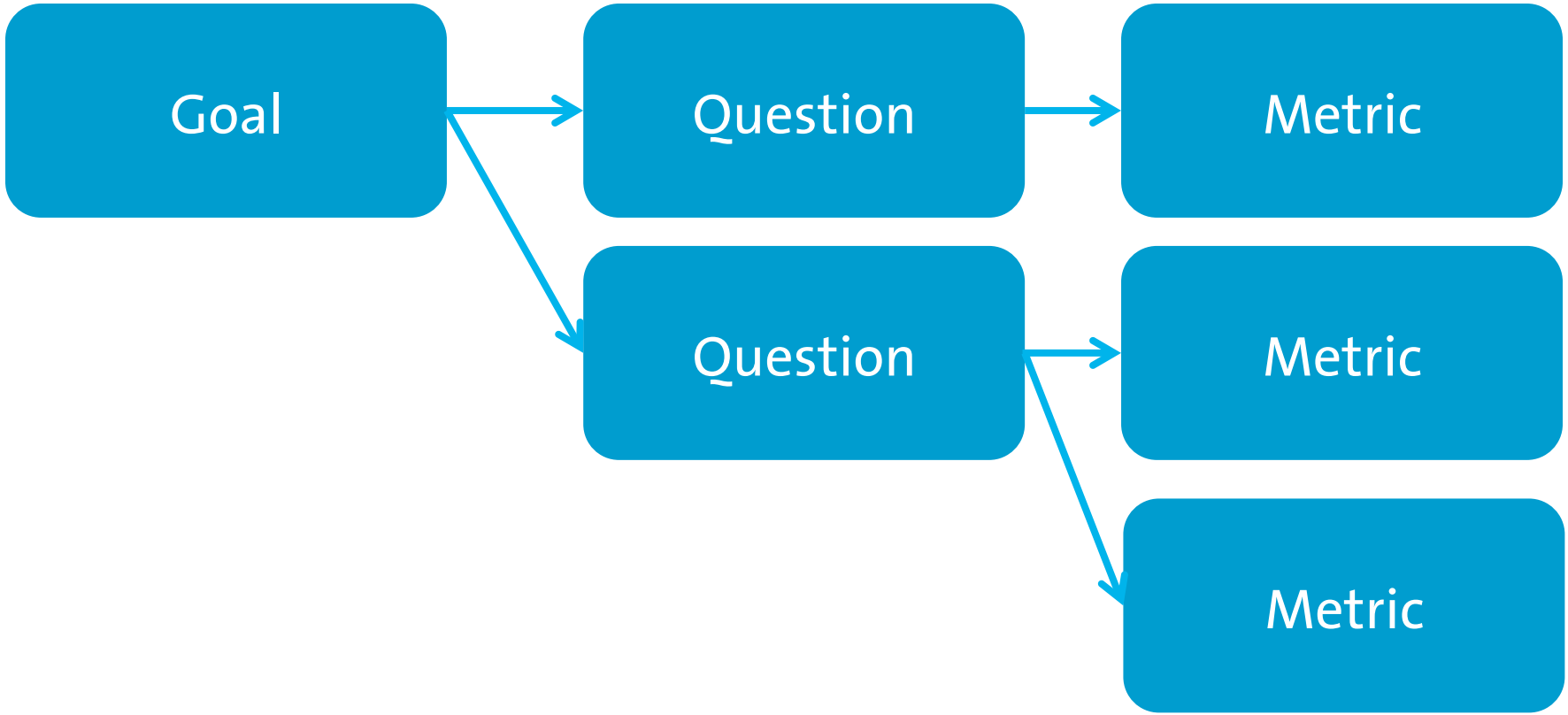

```
private String aap;
private Number noot;
private Date mies;

@Override
public int hashCode() {
    final int prime = 31;
    int result = 1;
    result = prime * result + ((aap == null) ? 0 : aap.hashCode());
    result = prime * result + ((mies == null) ? 0 : mies.hashCode());
    result = prime * result + ((noot == null) ? 0 : noot.hashCode());
    return result;
}
```

Metrics Galore



Wat moeten we dan meten?

GQM



Basili, Et. Al. , *The Goal Question Metric Approach*, Chapter in Encyclopedia of Software Engineering, Wiley, 1994.

GQM - Voorbeeld

Goal	Purpose Issue Object (process) Viewpoint	Improve the timeliness of change request processing from the project manager's viewpoint
Question	Q1	What is the current change request processing speed?
Metrics	M1 M2 M3	Average cycle time Standard deviation % cases outside of the upper limit
Question	Q2	Is the (documented) change request process actually performed?
Metrics	M4 M5	Subjective rating by the project manager % of exceptions identified during reviews

http://www.cs.umd.edu/users/mvz/handouts/gqm.pdf

Treating the metric

Metric in a bubble

Scope and Scale of NSA Collection

According to figures published by a major tech provider, the Internet carries 1,826 Petabytes of information per day. In its foreign intelligence mission, NSA touches about 1.6% of that. However, of the 1.6% of the data, only 0.025% is actually selected for review. The net effect is that NSA analysts look at 0.00004% of the world's traffic in conducting their mission – that's less than one part in a million. Put another way, if a standard basketball court represented the global communications environment, NSA's total collection would be represented by an area smaller than a dime on that basketball court.

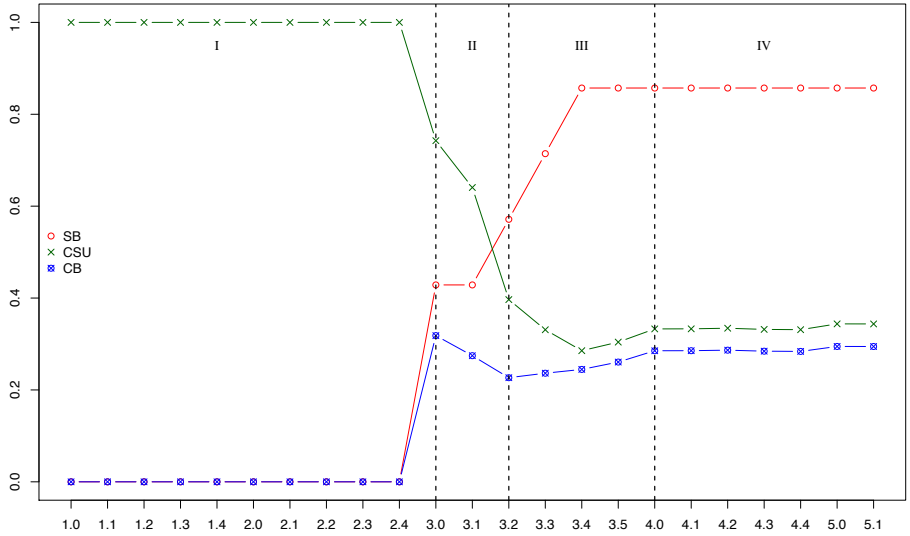
0.23 petabytes aan Facebook foto's

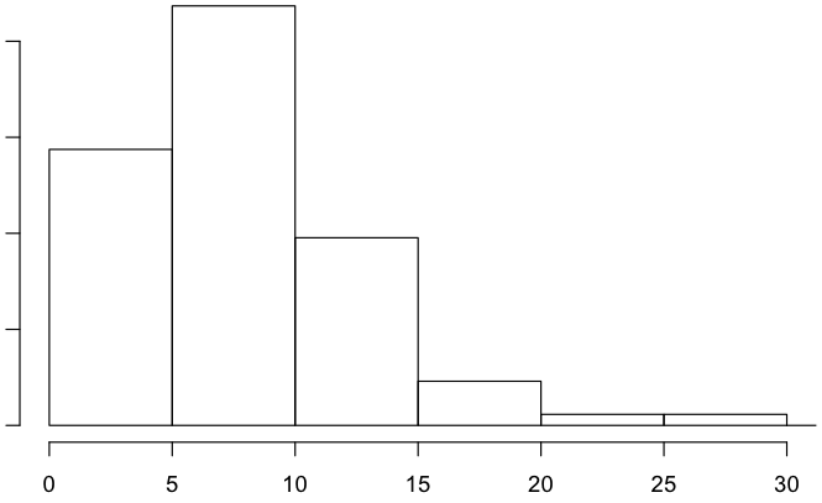
Realtime entertainment is 62% van internetverkeer

HTTP verkeer is 11.8% van het internetverkeer

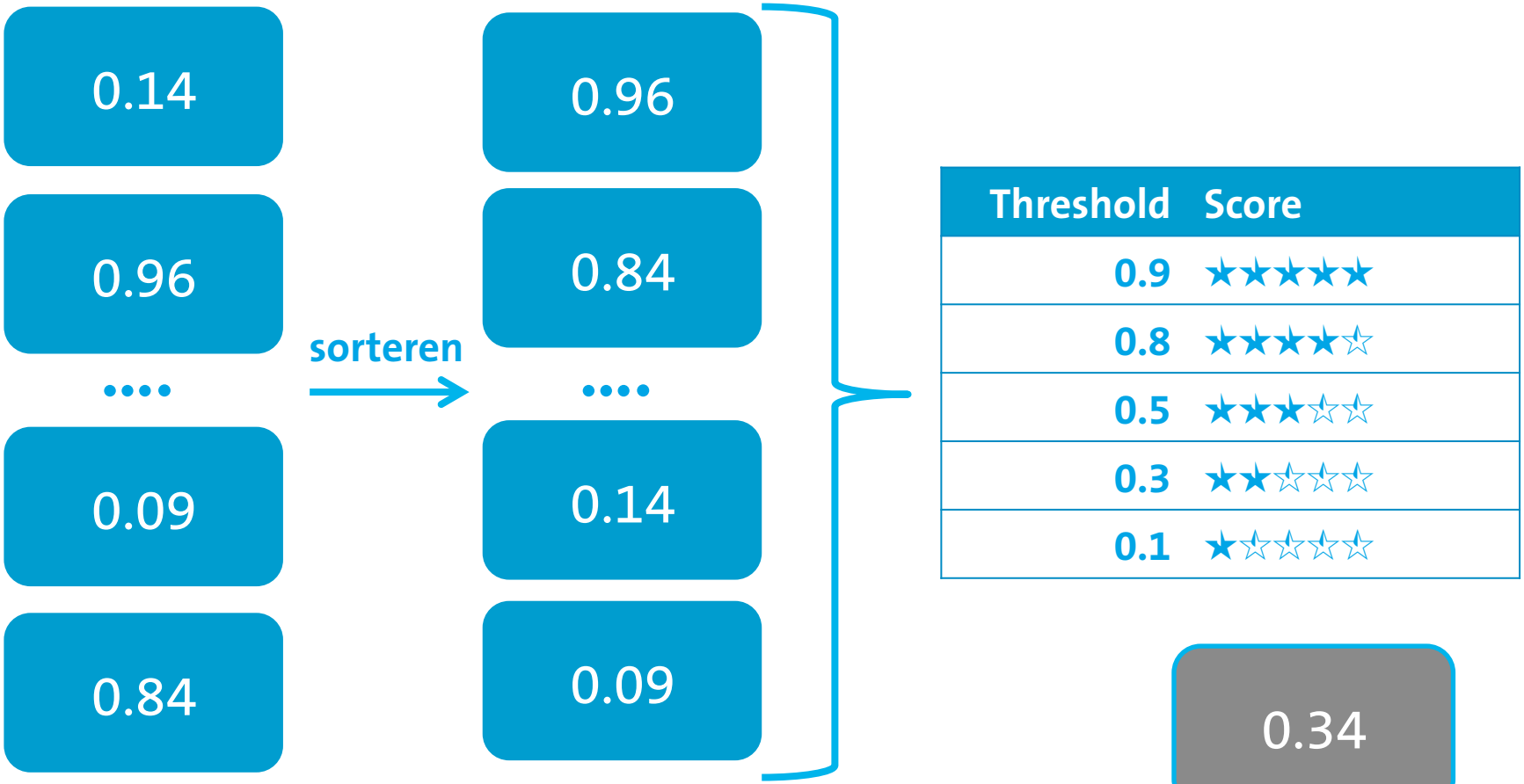
http://www.nsa.gov/public_info/_files/speeches_testimonies/2013_08_09_the_nsa_story.pdf
http://buzzmachine.com/2013/08/10/nsa-by-the-numbers/

Metric in a bubble





Een voorbeeld van context
Benchmarking



Threshold	Score
0.9	★★★★★
0.8	★★★★☆
0.5	★★★☆☆
0.3	★★☆☆☆
0.1	★☆☆☆☆

0.34

★★☆☆☆

Let op: voorbeeld thresholds

Valkuilen van het meten

	Te weinig	Te veel
Betekenis	Metric in a bubble	Treating the metric
# metrics	One-track metric	Metrics galore

E. Bouwers, A. van Deursen, and J. Visser. Getting What You Measure. Communications of the ACM, Vol. 55 No. 7, Pages 54-59, July 2012

Samenvatting

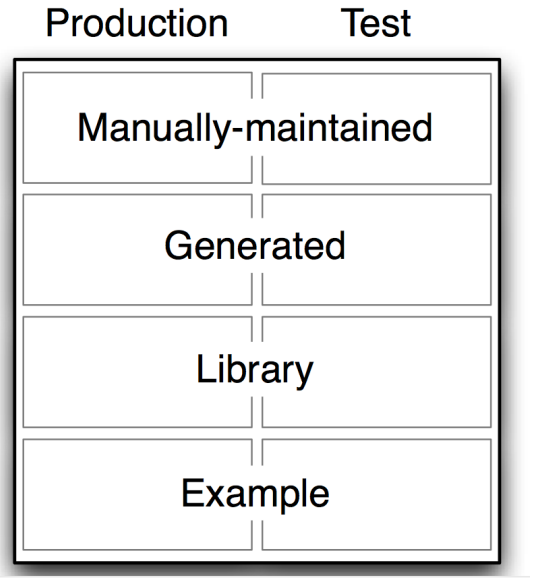
Uitdagingen in meten

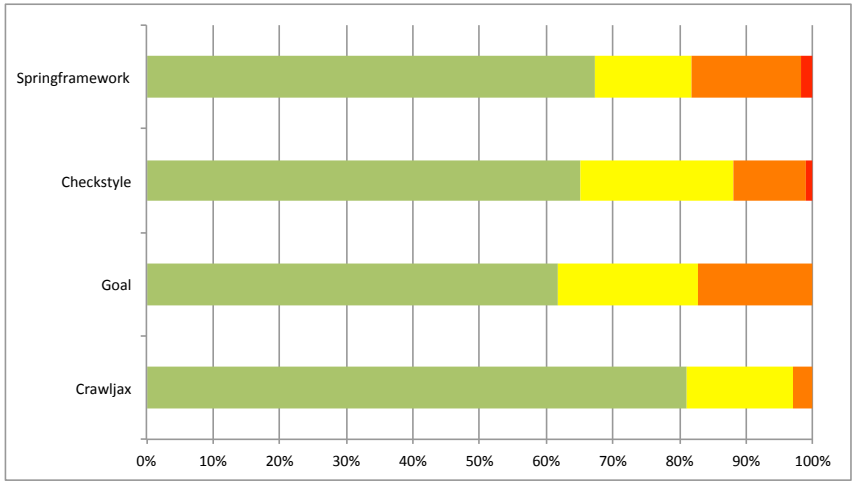
Entiteit

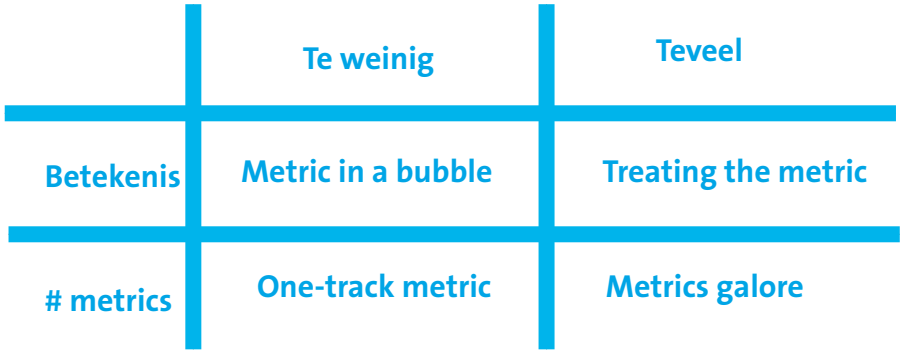
Attribuut

Mapping

Measure







Meten om te weten

Goal

Entiteit – Attribuut – Mapping

Context