

# On the Value of CrowdRE in Research and Practice



**Fabiano Dalpiaz**

Requirements Engineering Lab  
Utrecht University, Netherlands

September 21, 2021

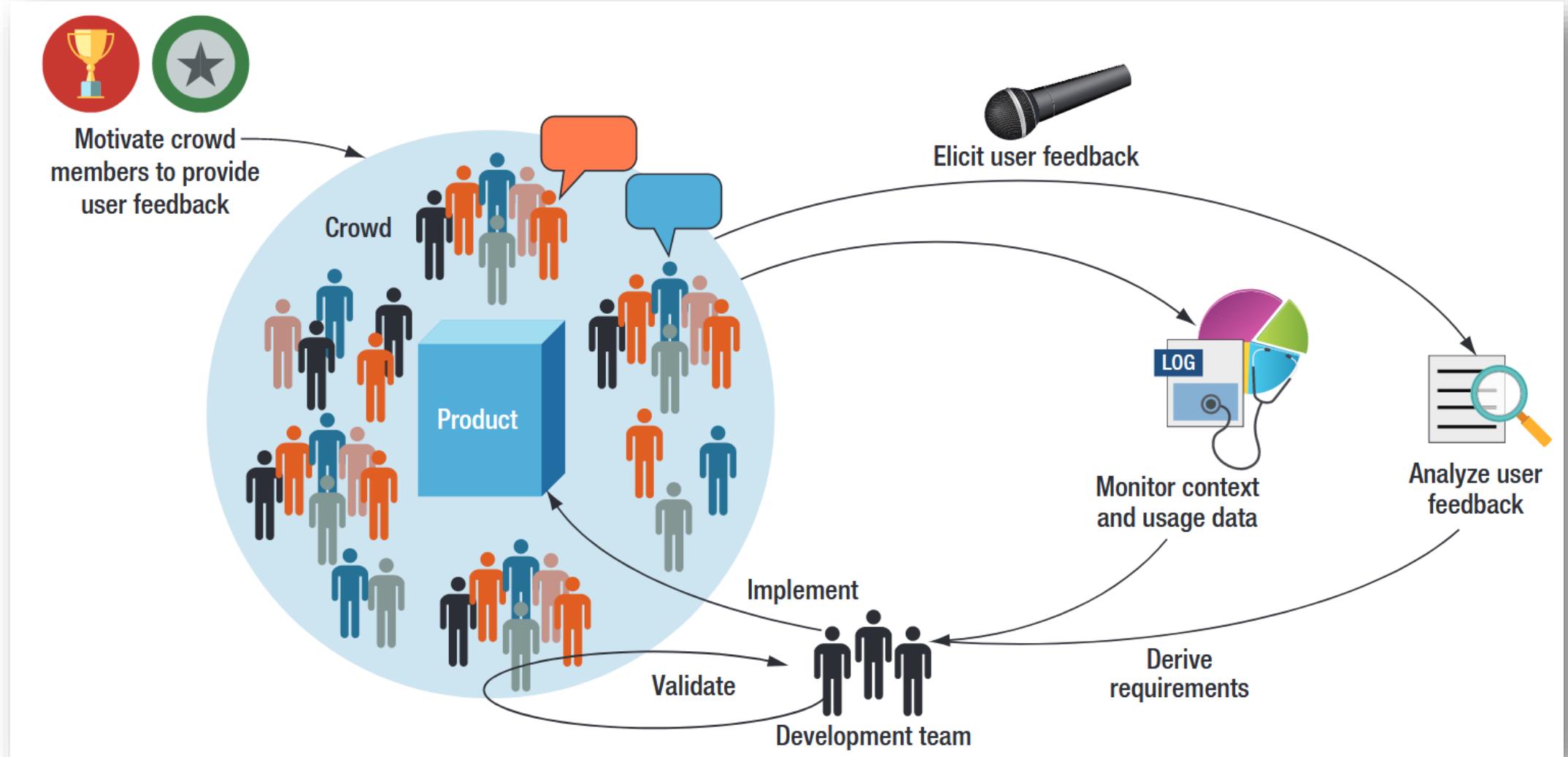


**Utrecht  
University**

# I.The status-quo in CrowdRE research

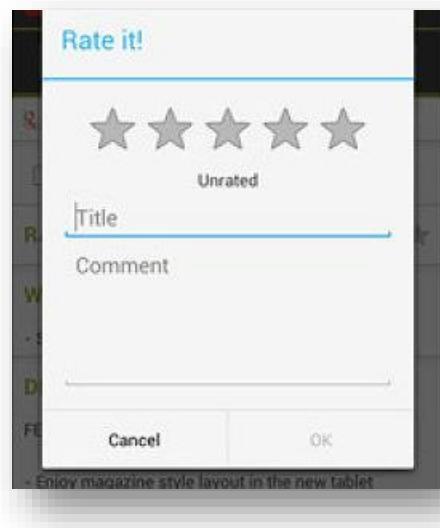


# CrowdRE, illustrated

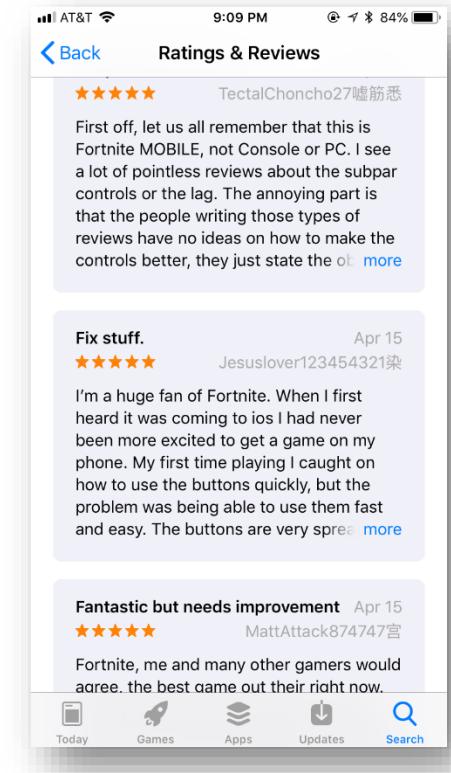


# Pull and Push feedback

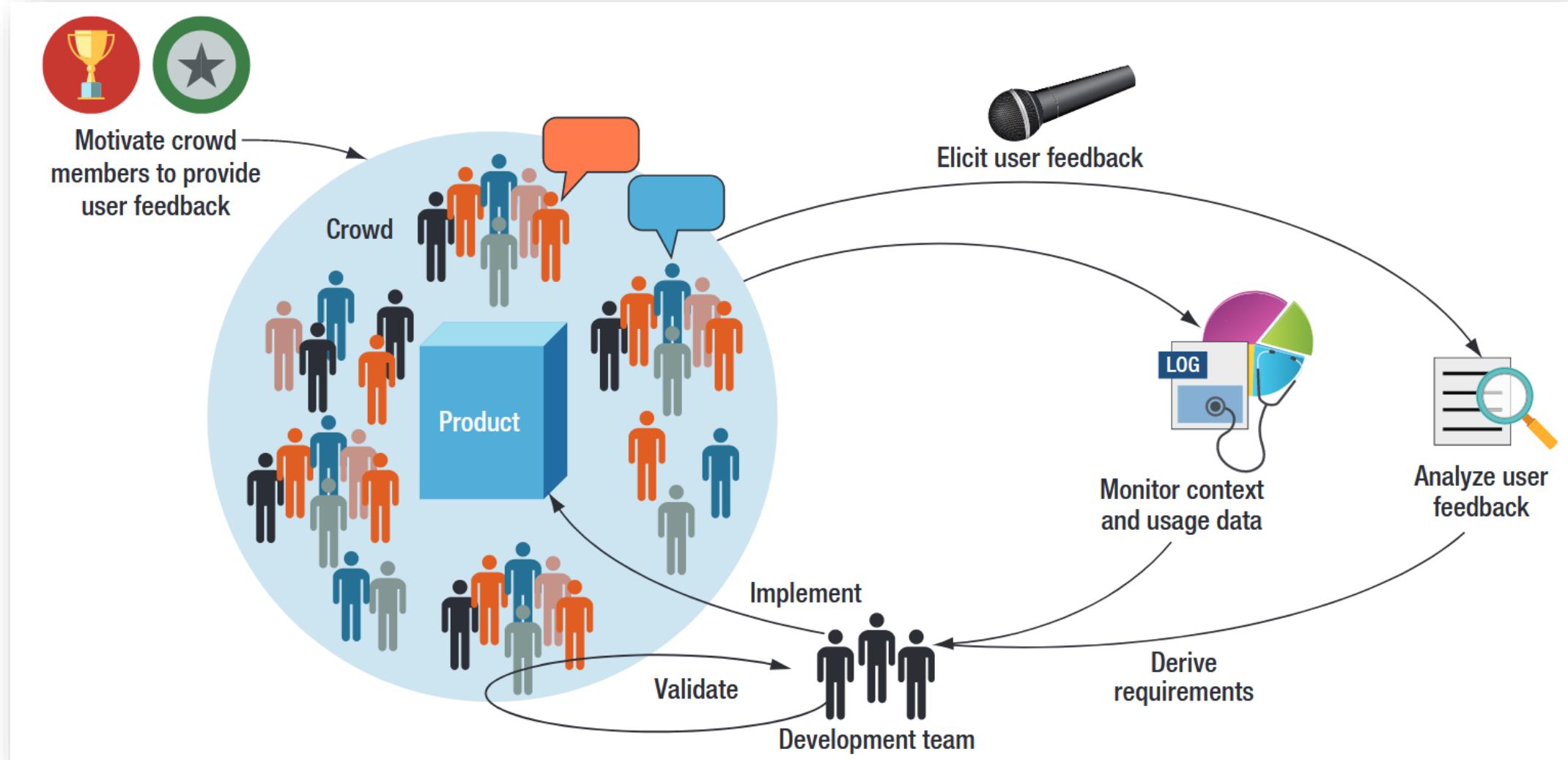
- ▶ **Pull feedback:** the company requests feedback to the users



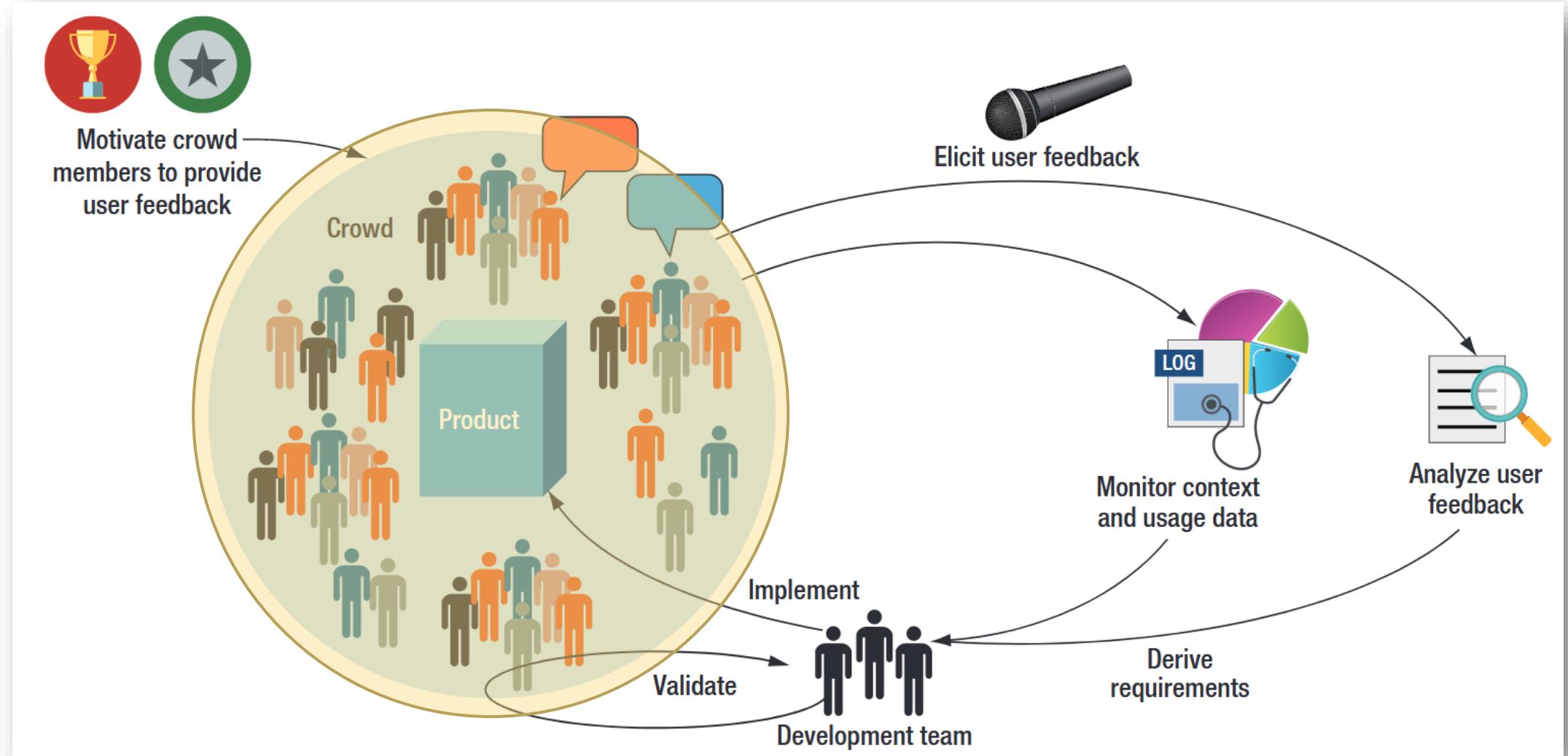
- ▶ **Push feedback:** the users spontaneously provide feedback



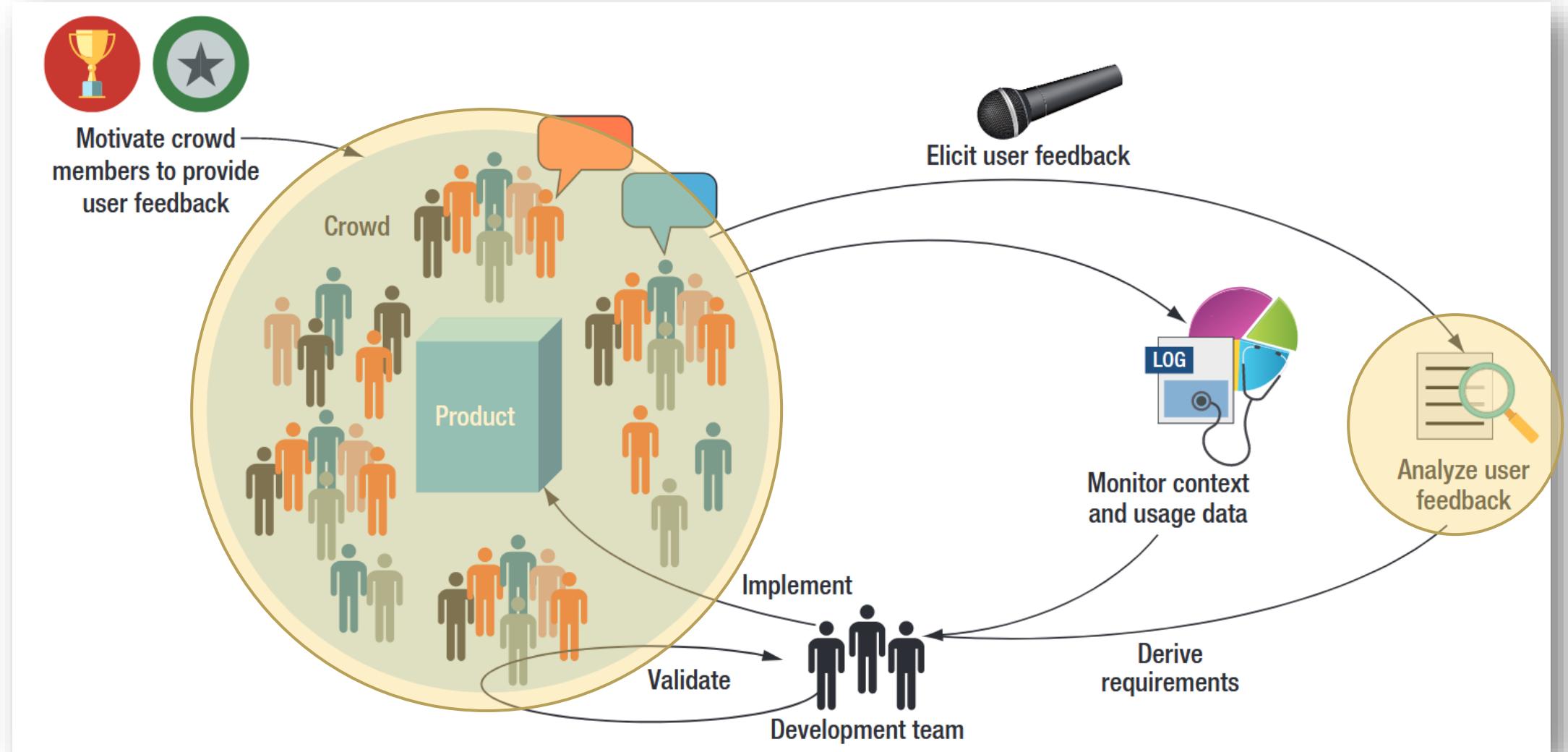
# Human effort in CrowdRE



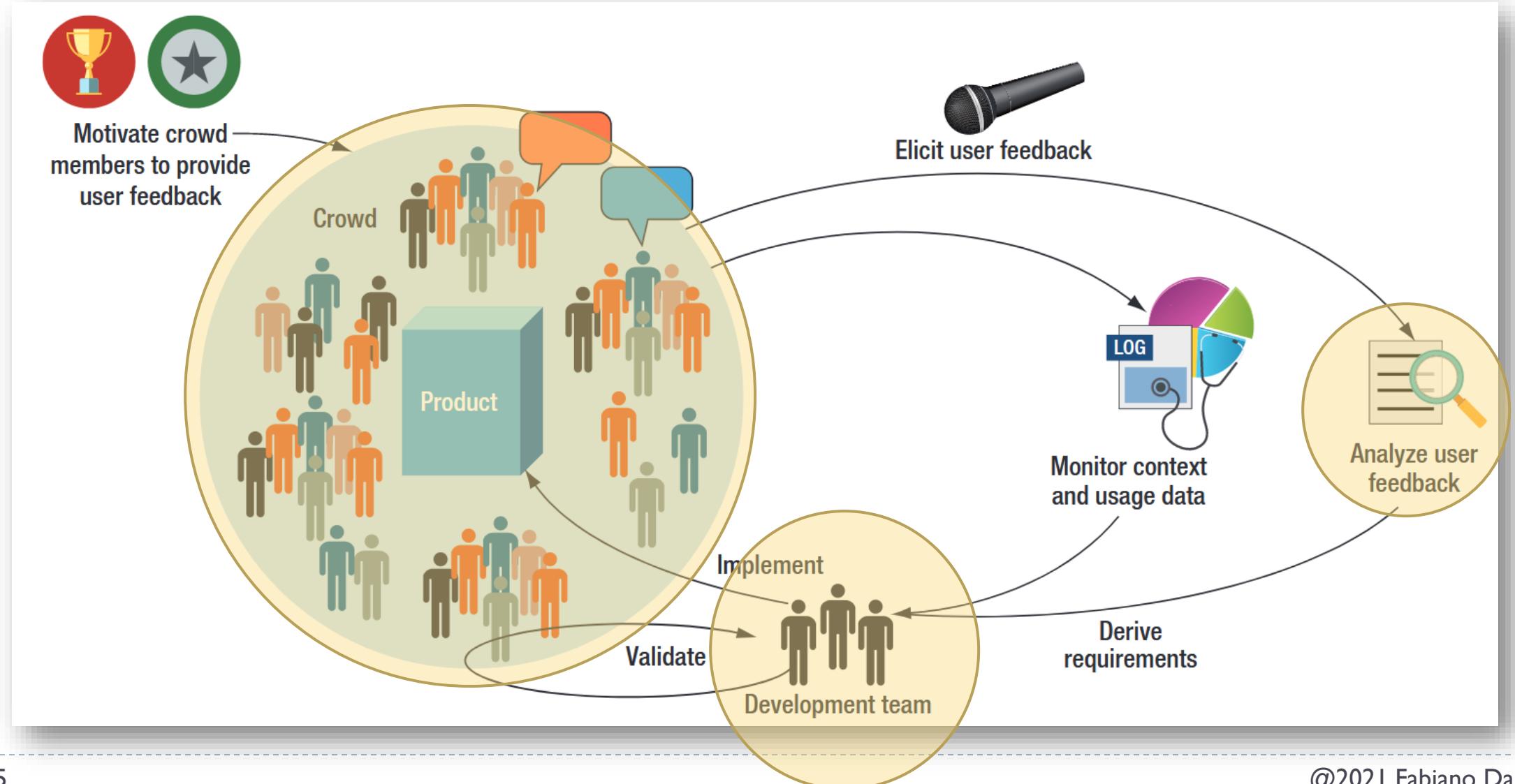
# Human effort in CrowdRE



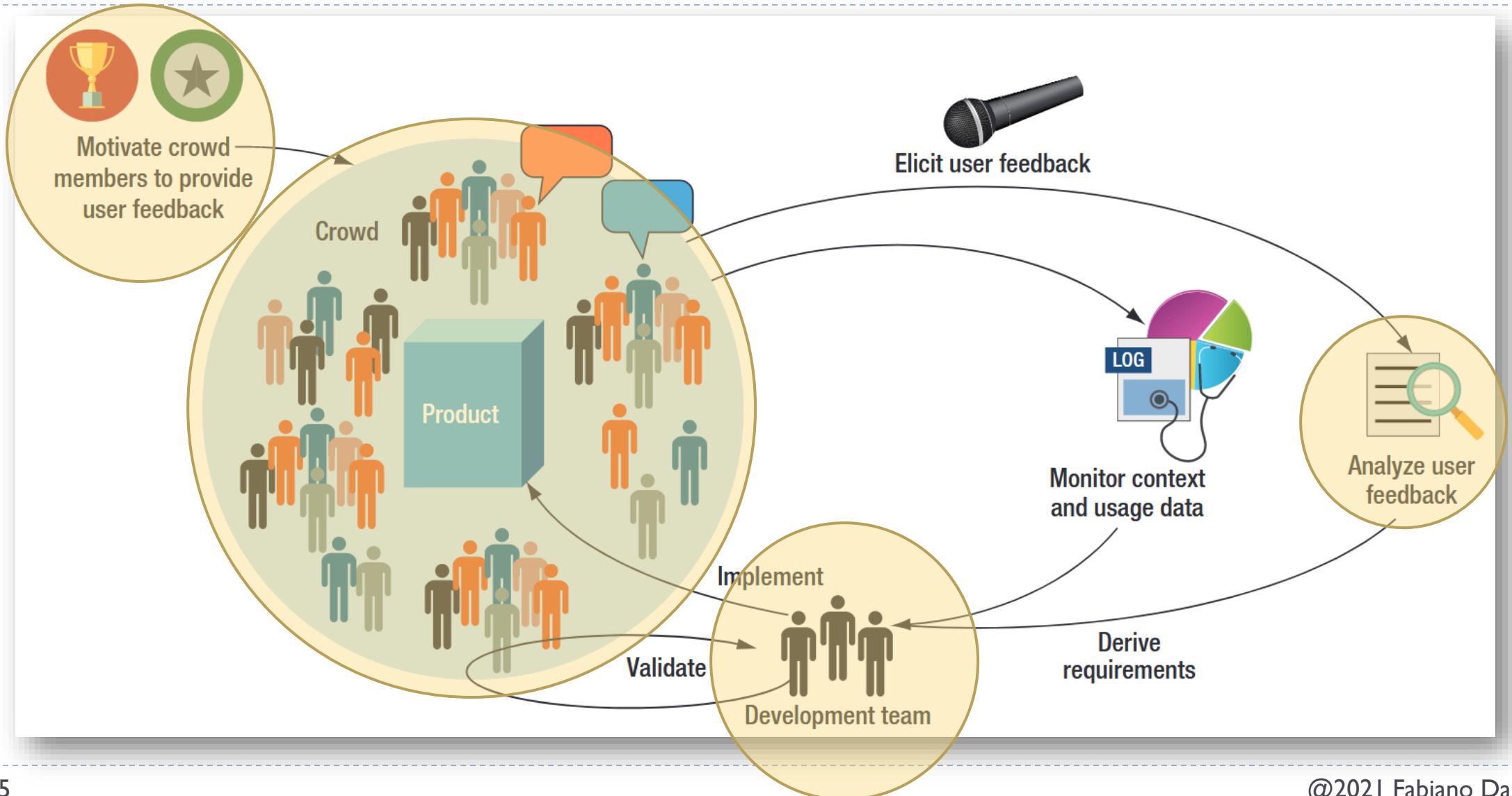
# Human effort in CrowdRE



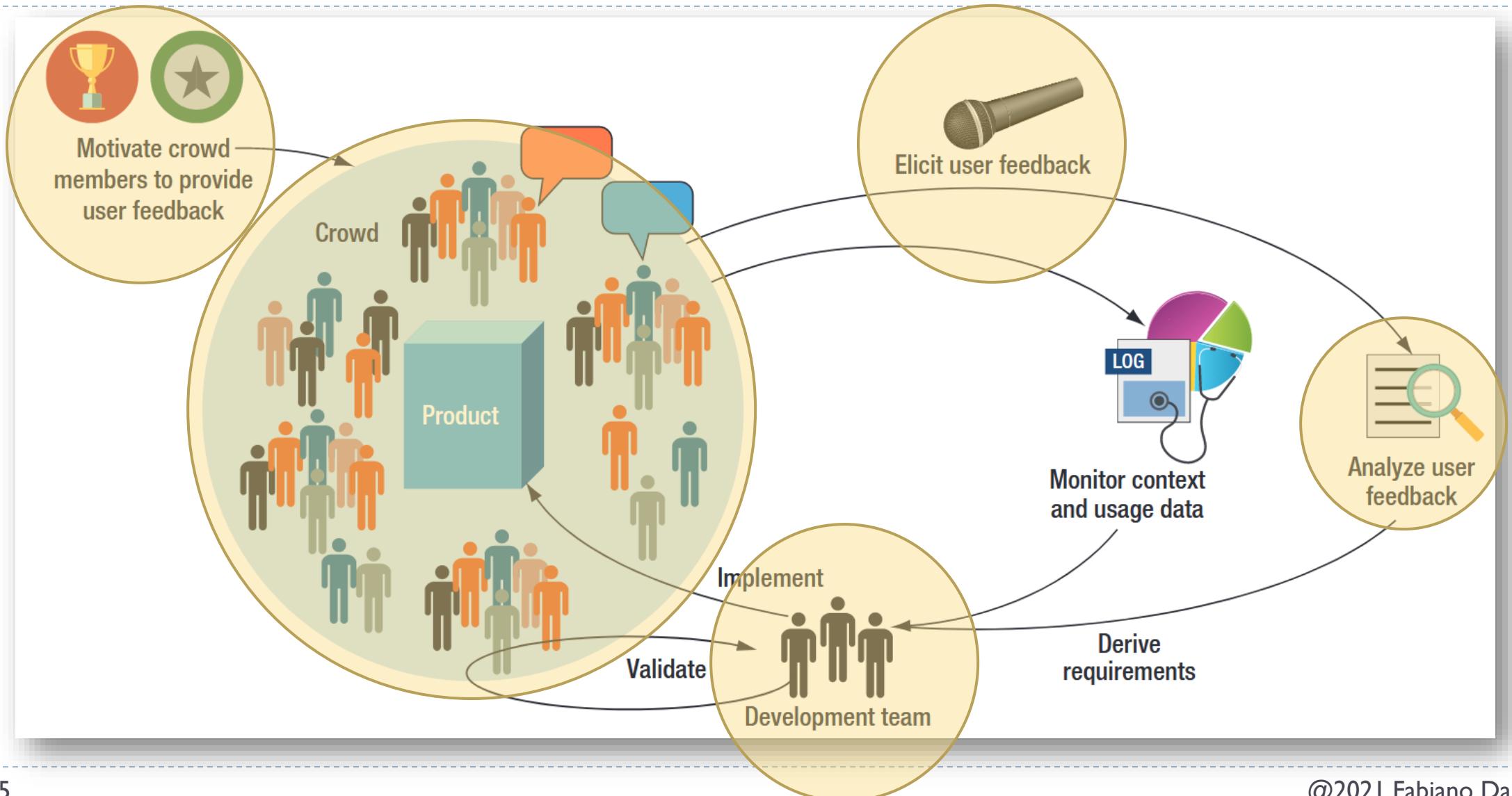
# Human effort in CrowdRE



# Human effort in CrowdRE



# Human effort in CrowdRE



# Research Question for the CrowdRE community

---

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



# Automation to the rescue

---

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice?**

# Automation to the rescue

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice?**

Requirements Eng (2016) 21:311–331  
DOI 10.1007/s00766-016-0251-9

RE 2015

## On the automatic classification of app reviews

Walid Maalej<sup>1</sup> · Zijad Kurtanović<sup>1</sup> · Hadeer Nabil<sup>2</sup> · Christoph Stanik<sup>1</sup>

ML outperforms  
keyword-based  
approaches

# Automation to the rescue

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice?**

Requirements Eng (2016) 21:311–331  
DOI 10.1007/s00766-016-0251-9

RE 2015

## On the automatic classification of app reviews

Walid Maalej<sup>1</sup> · Zijad Kurtanović<sup>1</sup> · Hadeer Nabil<sup>2</sup> · Christoph Stanik<sup>1</sup>

ML outperforms keyword-based approaches

Superior accuracy with less effort via continuous training

## App Review Analysis via Active Learning Reducing Supervision Effort without Compromising Classification Accuracy

Venkatesh T. Dhinakaran<sup>†</sup>, Raseshwari Pulle<sup>†</sup>, Nirav Ajmeri<sup>‡</sup>, Pradeep K. Murukannaiah<sup>†</sup>

# Automation to the rescue

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice?**

Requirements Eng (2016) 21:311–331  
DOI 10.1007/s00766-016-0251-9

RE 2015

## On the automatic classification of app reviews

Walid Maalej<sup>1</sup> · Zijad Kurtanović<sup>1</sup> · Hadeer Nabil<sup>2</sup> · Christoph Stanik<sup>1</sup>

ML outperforms keyword-based approaches

Superior accuracy with less effort via continuous training

## App Review Analysis via Active Learning Reducing Supervision Effort without Compromising Classification Accuracy

Venkatesh T. Dhinakaran<sup>†</sup>, Raseshwari Pulle<sup>†</sup>, Nirav Ajmeri<sup>‡</sup>, Pradeep K. Murukannaiah<sup>†</sup>

## How Do Users Like This Feature?

### A Fine Grained Sentiment Analysis of App Reviews

Emitza Guzman  
Technische Universität München  
Garching, Germany  
emitza.guzman@mytum.de

Walid Maalej  
University of Hamburg  
Hamburg, Germany  
maalej@informatik.uni-hamburg.de

What features are mentioned? With which sentiment?

# Automation = Value for Research

High citation  
figures

[How do users like this feature? a fine grained sentiment analysis of app reviews](#)

E Guzman, W Maalej

2014 IEEE 22nd international requirements engineering conference (RE), 153-162

541

2014

[Bug report, feature request, or simply praise? on automatically classifying app reviews](#)

W Maalej, H Nabil

2015 IEEE 23rd international requirements engineering conference (RE), 116-125

387

2015

[How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution](#)

S Panichella, AD Sorbo, E Guzman, A Visaggio, G Canfora, H Gall

Software Maintenance and Evolution (ICSME), 2015 IEEE International ...

356

2015

# Automation = Value for Research

High citation  
figures

How do users like this feature? a fine grained sentiment analysis of app reviews

E Guzman, W Maalej

2014 IEEE 22nd international requirements engineering conference (RE), 153-162

541

2014

Bug report, feature request, or simply praise? on automatically classifying app reviews

W Maalej, H Nabil

2015 IEEE 23rd international requirements engineering conference (RE), 116-125

387

2015

How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution

S Panichella, AD Sorbo, E Guzman, A Visaggio, G Canfora, H Gall

Software Maintenance and Evolution (ICSME), 2015 IEEE International ...

356

2015

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 43, NO. 9, SEPTEMBER 2017

A Survey of App Store Analysis  
for Software Engineering

William Martin, Federica Sarro, Yue Jia, Yuanyuan Zhang, and Mark Harman

An Overview of User Feedback Classification Approaches

Rubens Santos, Eduard C. Groen, Karina Villela  
Fraunhofer IESE, Kaiserslautern, Germany

Extensive number  
of papers

# Automation = Value for Research

High citation  
figures

How do users like this feature? a fine grained sentiment analysis of app reviews

E Guzman, W Maalej

2014 IEEE 22nd international requirements engineering conference (RE), 153-162

541

2014

Bug report, feature request, or simply praise? on automatically classifying app reviews

W Maalej, H Nabil

2015 IEEE 23rd international requirements engineering conference (RE), 116-125

387

2015

How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution

S Panichella, AD Sorbo, E Guzman, A Visaggio, G Canfora, H Gall

Software Maintenance and Evolution (ICSME), 2015 IEEE International ...

356

2015

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 43, NO. 9, SEPTEMBER 2017

A Survey of App Store Analysis  
for Software Engineering

William Martin, Federica Sarro, Yue Jia, Yuanyuan Zhang, and Mark Harman

An Overview of User Feedback Classification Approaches

Rubens Santos, Eduard C. Groen, Karina Villela  
Fraunhofer IESE, Kaiserslautern, Germany

Extensive number  
of papers

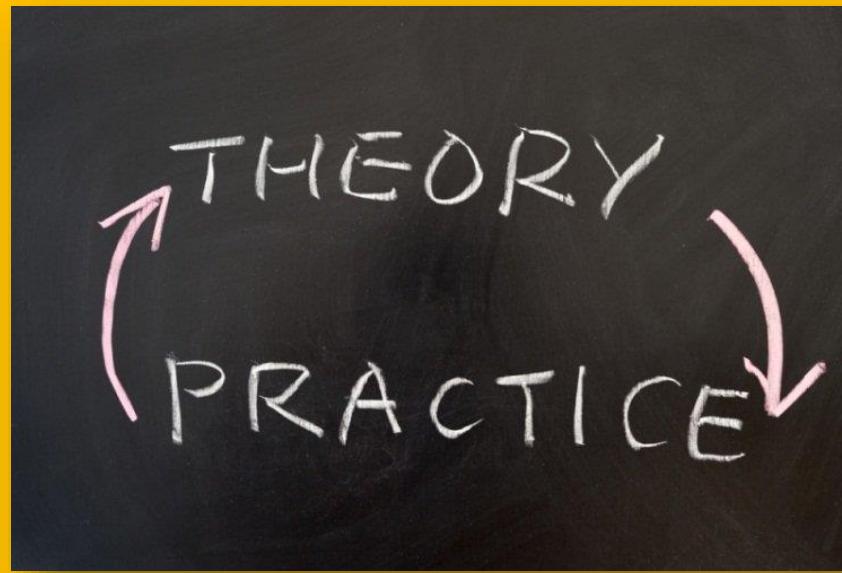
Standard metrics to  
show success

$$R = \frac{|fnd \cap rel|}{|rel|} = \frac{|TP|}{|TP| + |FN|}$$
$$P = \frac{|fnd \cap rel|}{|fnd|} = \frac{|TP|}{|FP| + |TP|} \quad (2)$$

posite of recall and precision that is often called  
' is captured by the  $F$ -measure:

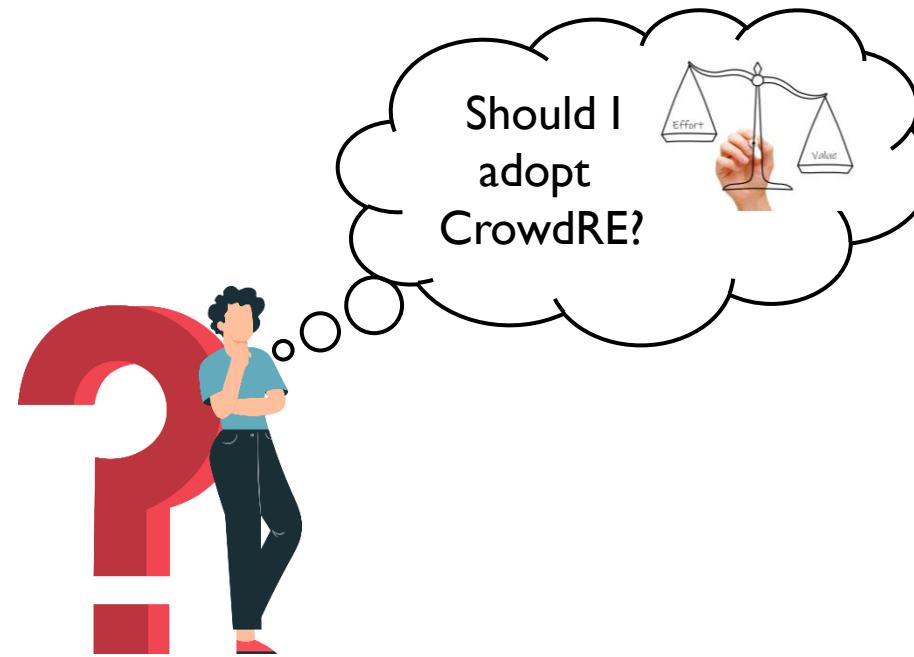
$$F = 2 \times \frac{P \times R}{P + R}, \quad (3)$$

## 2. How about value for practice?



# The practitioner's standpoint: meet **Alex**

---



# Value for practice in **push feedback**

Method	P1: Useless reviews			P1: Useful reviews		
	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
Crowdsourcing	0.93	0.84	0.88	0.83	0.93	0.88
SVM	0.90	0.79	0.84	0.83	0.92	0.87
Naïve Bayes	0.83	0.79	0.81	0.81	0.85	0.83
FastText	0.75	0.60	0.67	0.84	0.91	0.87
ELMo	0.83	0.80	0.82	0.81	0.84	0.82
BERT	<b>0.95</b>	<b>0.88</b>	<b>0.92</b>	<b>0.88</b>	<b>0.96</b>	<b>0.92</b>

Humans      {  
Automation    {

Thu 23 Sep

Displayed time zone: Eastern Time (US & Canada) [change](#)

[Edit session information](#)

08:00 30m ★ **Classifying User Requirements from Online Feedback in Small Dataset Environments using Deep Learning**  
Talk Research Papers [RESEARCH PAPER](#)  
Rohan Reddy Mekala Fraunhofer USA CESE,  
Asif Irfan Fraunhofer USA Center Mid-Atlantic, Eduard C. Groen Fraunhofer IESE,  
Adam Porter Fraunhofer USA CESE, Mikael Lindvall Fraunhofer USA CESE  
  
[Media Attached](#)

# Value for practice in **push feedback**

	Method	P1: Useless reviews			P1: Useful reviews		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
Humans	Crowdsourcing	0.93	0.84	0.88	0.83	0.93	0.88
	SVM	0.90	0.79	0.84	0.83	0.92	0.87
Automation	Naïve Bayes	0.83	0.79	0.81	0.81	0.85	0.83
	FastText	0.75	0.60	0.67	0.84	0.91	0.87
	ELMo	0.83	0.80	0.82	0.81	0.84	0.82
	BERT	<b>0.95</b>	<b>0.88</b>	<b>0.92</b>	<b>0.88</b>	<b>0.96</b>	<b>0.92</b>



# Value for practice in **push feedback**

Humans

Automation

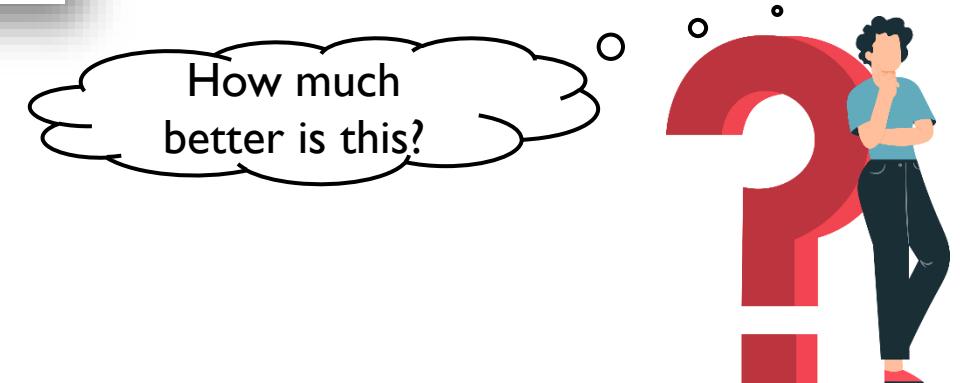
Method	P1: Useless reviews			P1: Useful reviews		
	P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
Crowdsourcing	0.93	0.84	0.88	0.83	0.93	0.88
SVM	0.90	0.79	0.84	0.83	0.92	0.87
Naïve Bayes	0.83	0.79	0.81	0.81	0.85	0.83
FastText	0.75	0.60	0.67	0.84	0.91	0.87
ELMo	0.83	0.80	0.82	0.81	0.84	0.82
BERT	<b>0.95</b>	<b>0.88</b>	<b>0.92</b>	<b>0.88</b>	<b>0.96</b>	<b>0.92</b>

But I need experts  
to train ML/DL!



# Value for practice in **push feedback**

	Method	P1: Useless reviews			P1: Useful reviews		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
Humans	Crowdsourcing	0.93	0.84	0.88	0.83	0.93	0.88
	SVM	0.90	0.79	0.84	0.83	0.92	0.87
Automation	Naïve Bayes	0.83	0.79	0.81	0.81	0.85	0.83
	FastText	0.75	0.60	0.67	0.84	0.91	0.87
	ELMo	0.83	0.80	0.82	0.81	0.84	0.82
	BERT	0.95	0.88	0.92	0.88	0.96	0.92



# Value for practice in **push feedback**

	Method	P1: Useless reviews			P1: Useful reviews		
		P	R	F <sub>1</sub>	P	R	F <sub>1</sub>
Humans	Crowdsourcing	0.93	0.84	0.88	0.83	0.93	0.88
	SVM	0.90	0.79	0.84	0.83	0.92	0.87
	Naïve Bayes	0.83	0.79	0.81	0.81	0.85	0.83
Automation	FastText	0.75	0.60	0.67	0.84	0.91	0.87
	ELMo	0.83	0.80	0.82	0.81	0.84	0.82
	BERT	<b>0.95</b>	<b>0.88</b>	<b>0.92</b>	<b>0.88</b>	<b>0.96</b>	<b>0.92</b>



# Value for practice in **pull feedback**

---

(i) *Amount of feedback data gathered and acceptance of the feedback tool.* SEnerCon received 81 feedback entries from 82 end-users in two weeks (seven feedback entries were excluded for further analysis due to invalid information, such as “test”). The feedback was mainly about (multiple-answers possible) ideas for improvements (37), computation error (17), others (17) and technical problems (13). “General error” was chosen in seven cases and “expert question” in four cases.

lack of knowledge on how software companies set up crowd involvement activities in practice. We observed a German software SME during their first steps in crowd-based feedback and monitoring data collection, and reported on their experiences during an initial implementation and evaluation phase.

We found that the end-users of the energy saving account communicate feedback very seldom and prefer a bilateral and non-public feedback communication with the software company. This could make it cumbersome for the software company

Stade et al., CrowdRE'19

# Value for practice in **pull feedback**

(i) *Amount of feedback data gathered and acceptance of the feedback tool.* SEnerCon received 81 feedback entries from 82 end-users in two weeks (seven feedback entries were excluded for further analysis due to invalid information, such as “test”). The feedback was mainly about (multiple-answers possible) ideas for improvements (37), computation error (17), others (17) and technical problems (13). “General error” was chosen in seven cases and “expert question” in four cases.

lack of knowledge on how software companies set up crowd involvement activities in practice. We observed a German software SME during their first steps in crowd-based feedback and monitoring data collection, and reported on their experiences during an initial implementation and evaluation phase.

We found that the end-users of the energy saving account communicate feedback very seldom and prefer a bilateral and non-public feedback communication with the software company. This could make it cumbersome for the software company

**Table 6** Overview of key figures

	Total	Avg./User
Platform visitors (see Table 7 and Fig. 10)	726	
Potential stakeholders (see Sect. 6.4.2)	244	
Stakeholders (see Sect. 6.4.3)	63	
Active stakeholders (see Sect. 6.4.3, Table 8, and Figs. 10 and 11)	32	
Users per expertise level (finally reached level; see Table 10)	I: 19; II: 3; III: 6; IV: 2; V: 2	
User activities	Total	Avg./User
Duration of participation (days from registration to last activity; see Table 9)	418	13.1
Number of logins (see Table 9)	177	5.5
Viewed posts (see Table 9)	468	14.6
Shared posts (see Table 10)	37	1.2
Shared sub-posts (see Table 10)	19	1.5
Ratings (see Table 10)	288	9
Votes on posts (see Table 10)	141	10.9
Votes on sub-posts (see Table 10)	19	1.9

Stade et al., CrowdRE'19

Kolpondinos & Glinz, RE Journal 2020

# Value for practice in **pull feedback**

(i) Amount of feedback data gathered and acceptance of the feedback tool. SEnerCon received 81 feedback entries from 82 end-users in two weeks (seven feedback entries were excluded for further analysis due to incomplete responses). The feedback was mainly ideas for improvements (36), suggestions (17) and technical problems (17) and technical problem reports (17) in seven cases and “expert opinions” in one case.

Lack of knowledge on how to involve end-users in involvement activities in the software SME during their initial monitoring and monitoring data collections during an initial implementation.

We found that the end-users communicate feedback very seldom and prefer a bilateral and non-public feedback communication with the software company. This could make it cumbersome for the software company.

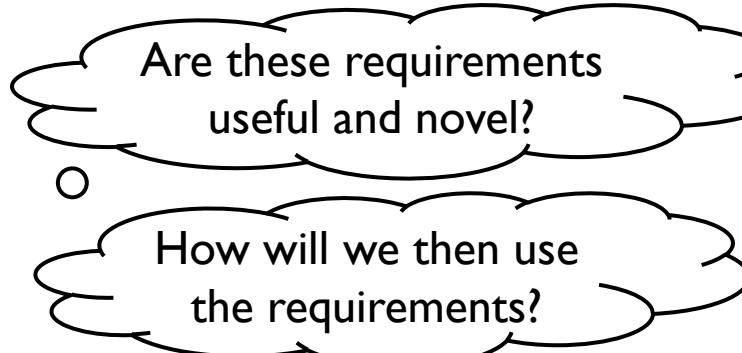


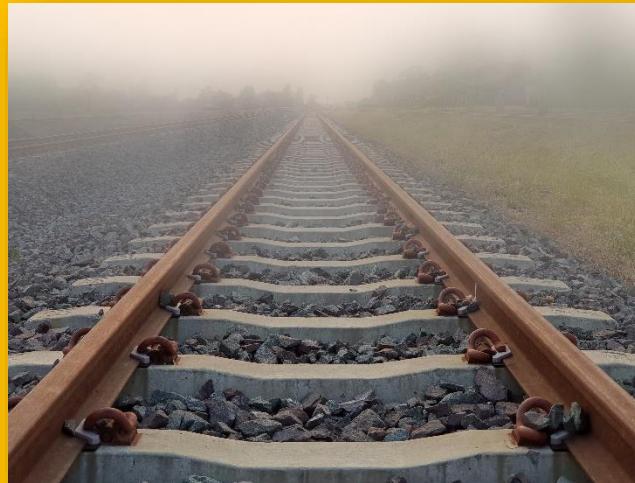
Table 6 Overview of key figures

	Total	Avg./User
Platform visitors (see Table 7 and Fig. 10)	726	
Potential stakeholders (see Sect. 6.4.2)	244	
Stakeholders (see Sect. 6.4.2)	63	
Stakeholders (see Sect. 6.4.2)	32	
I: 19; II: 3; III: 6; IV: 2; V: 2		
Total	418	13.1
Shared sub-posts (see Table 10)	177	5.5
Ratings (see Table 10)	468	14.6
Votes on posts (see Table 10)	37	1.2
Votes on sub-posts (see Table 10)	19	1.5
Votes on sub-posts (see Table 10)	288	9
Votes on sub-posts (see Table 10)	141	10.9
Votes on sub-posts (see Table 10)	19	1.9

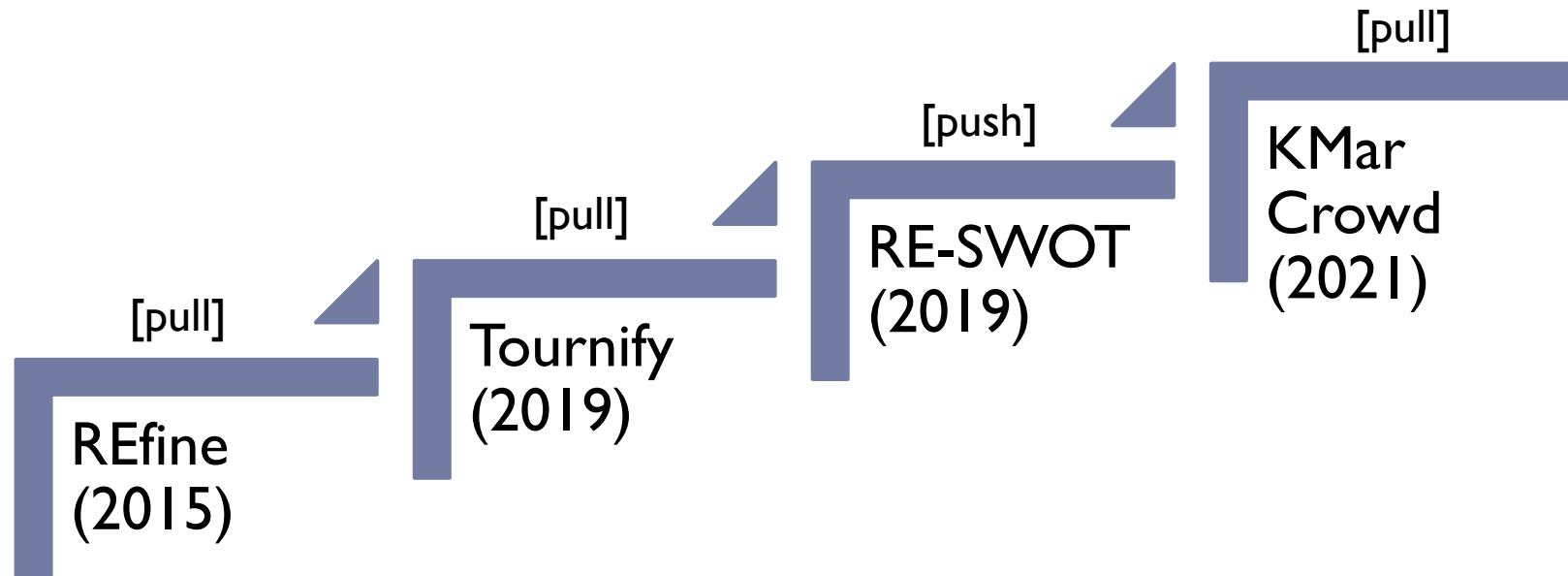
Stade et al., CrowdRE'19

Kolpondinos & Glinz, RE Journal 2020

### **3. A personal journey across research and practice**



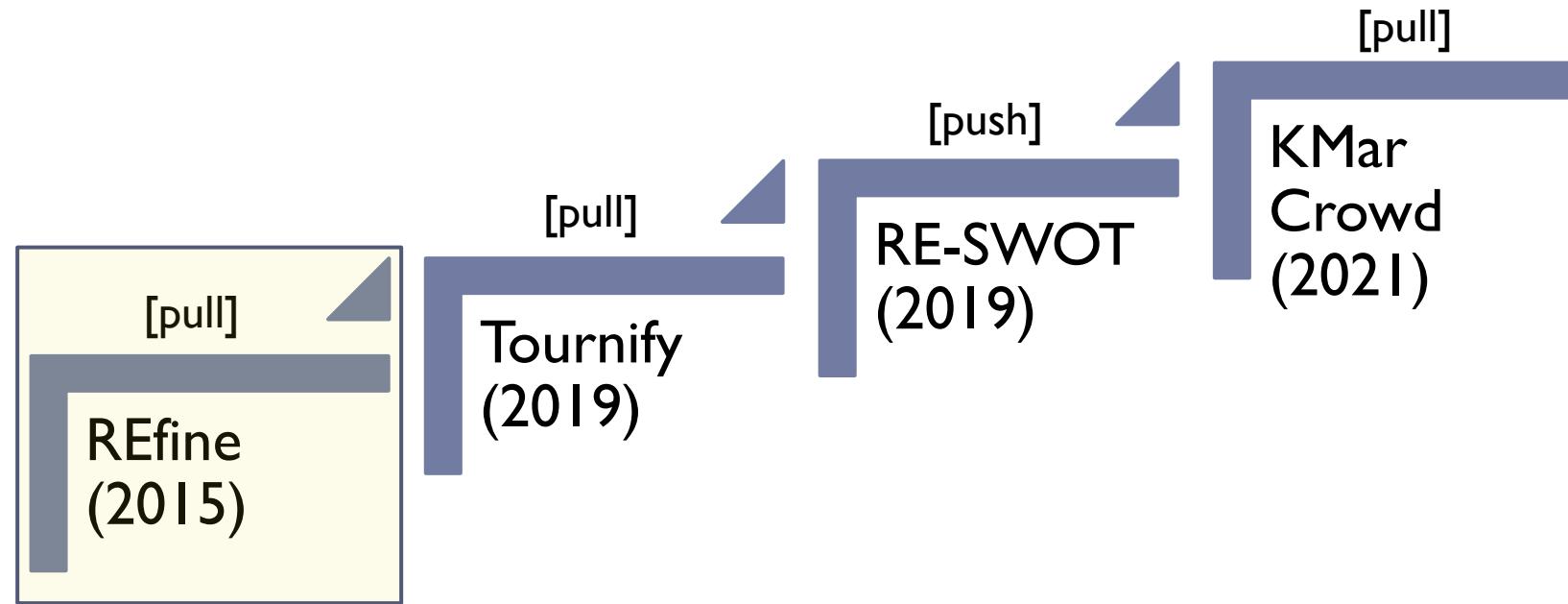
# Personal Journey: Overview



Remco Snijders, Fabiano Dalpiaz, Sjaak Brinkkemper, Mahmood Hosseini, Raian Ali, Atilla Ozum:  
REFine: A gamified platform for participatory requirements engineering. CrowdRE@RE 2015: 1-6

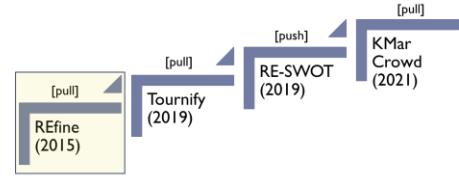


# Personal Journey: Overview

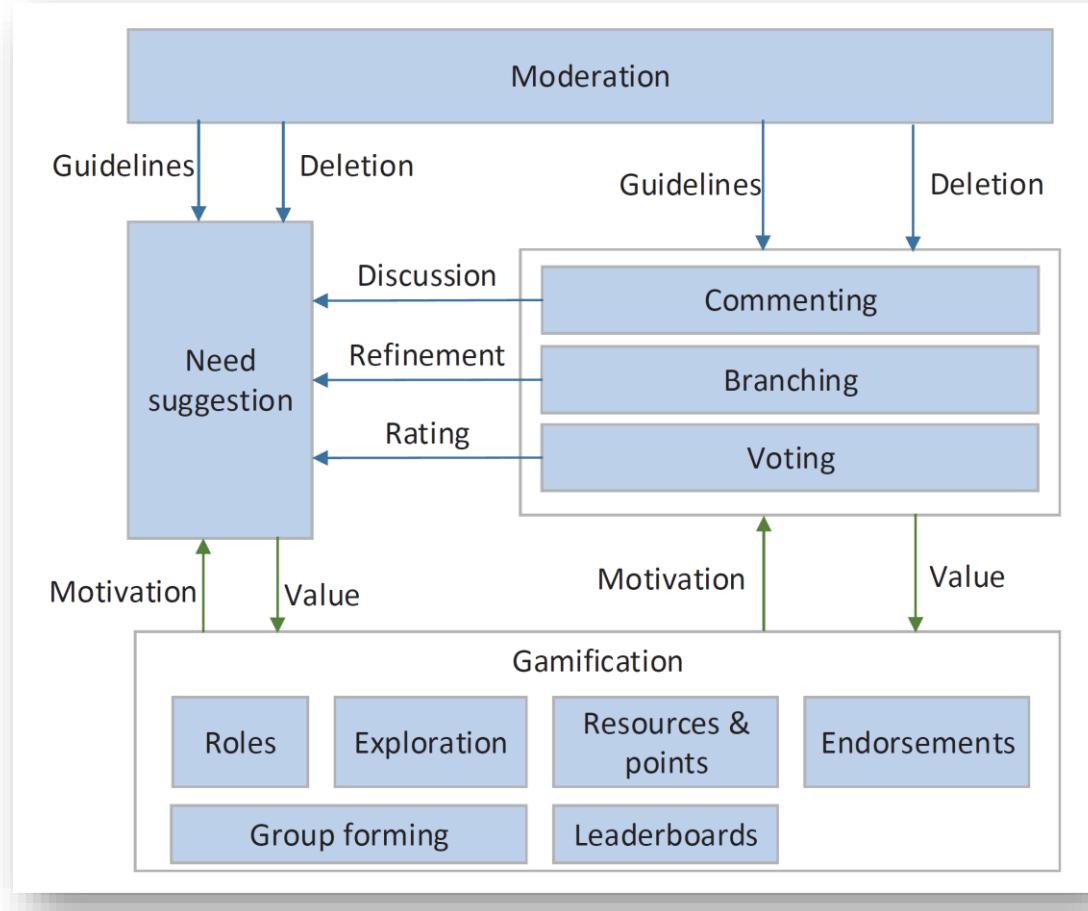


Remco Snijders, Fabiano Dalpiaz, Sjaak Brinkkemper, Mahmood Hosseini, Raian Ali, Atilla Ozum:  
REFine: A gamified platform for participatory requirements engineering. CrowdRE@RE 2015: 1-6

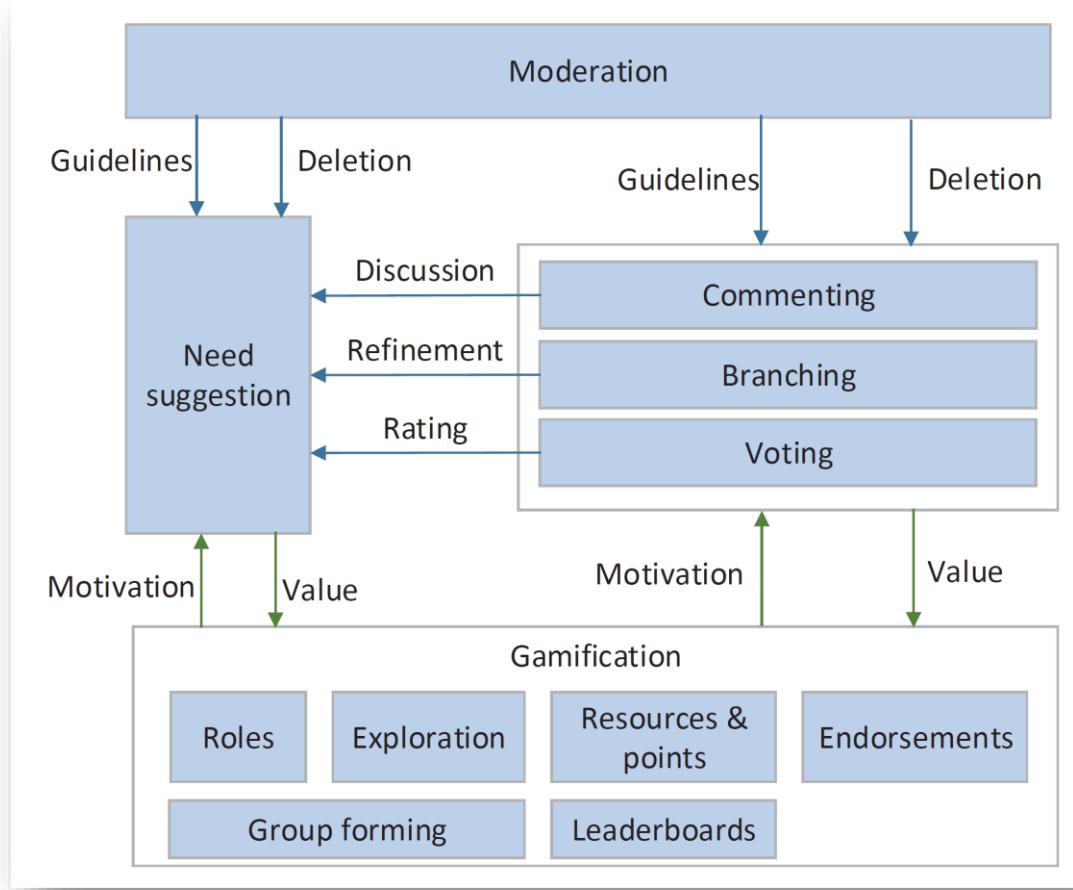
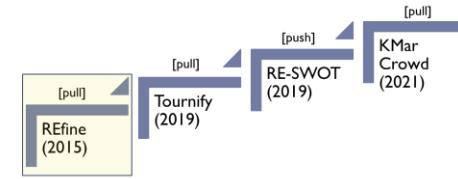




# REFine: the start of a journey into CrowdRE



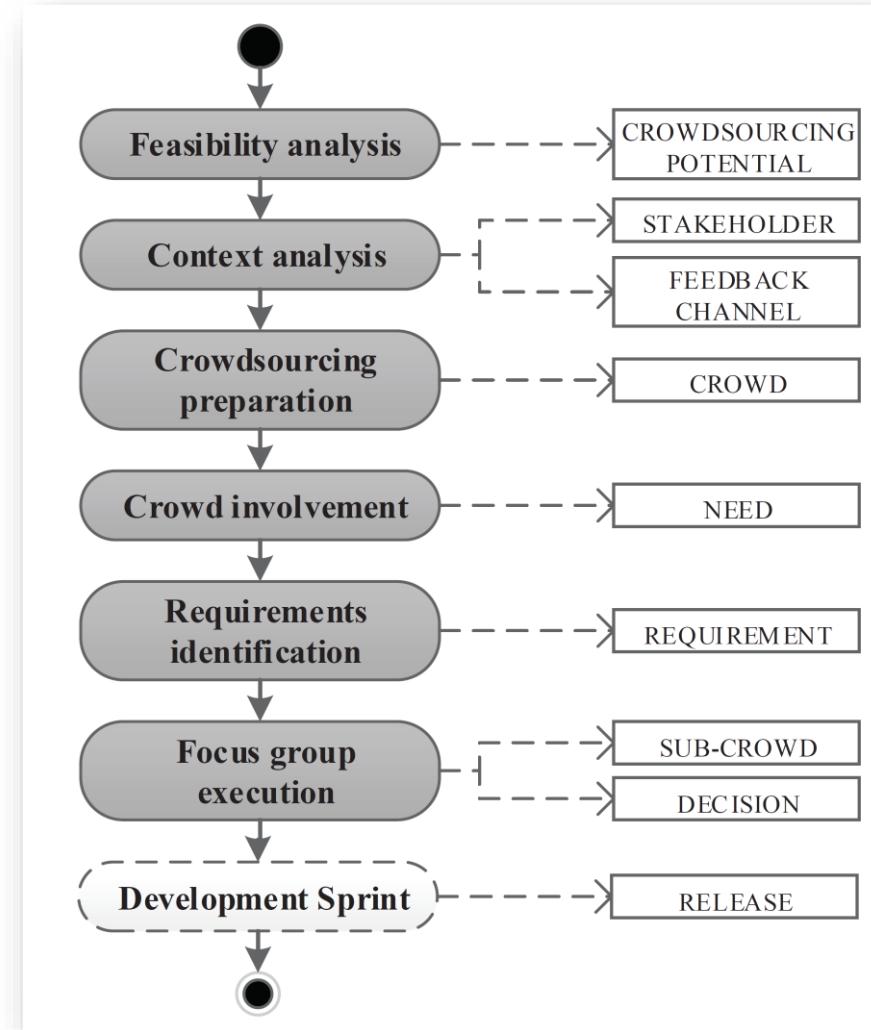
# REFine: the start of a journey into CrowdRE



A screenshot of the 'Refine' platform interface. The top navigation bar includes 'Refine', 'Home', 'About', 'Needs', 'Leaderboards', and 'Contact'. A user profile is shown on the right with 'Test user' and various icons. Below the navigation is a main content area with the heading 'How can Qubus 7 help you in being more efficient?'. A text block states: 'Qubus 7 is brand new. Compared to the previous versions, it is an improvement, but we're sure that we can improve it much further. In order to address the right needs, we want to hear the opinion of the people who use Qubus. We want to hear you!'. Below this are several user posts with 'Sort by: most recent | most popular' and a search bar. Posts include:

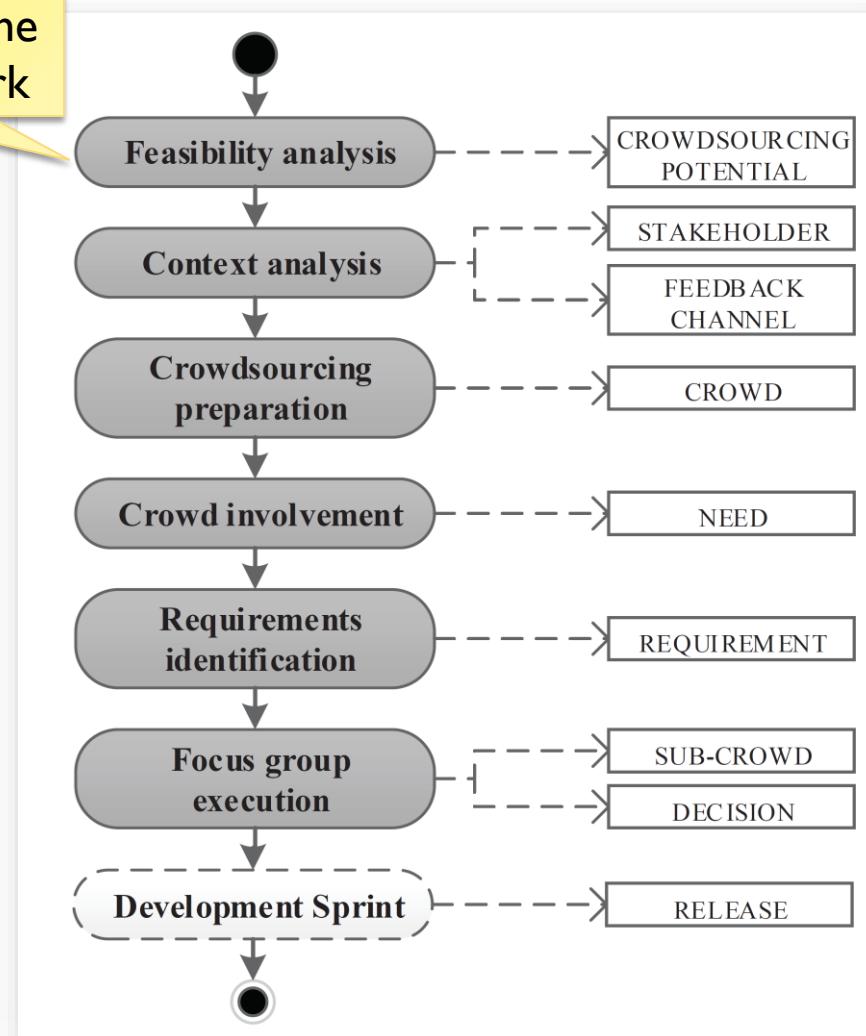
- Style-less qubus7**: 'Qubus7 is known for its refactored frontend (among other features of course). Wouldn't it be a good idea to also make sure the interface is somewhat browsable / usable if a simple layout is used? (For text browsers / older mobile devices / braille browsers and the likes?)' by Remy Alidaro on 30 Dec 2014 | 2 comments. Reactions: Agree (6), Disagree (0).
- Visualizator of question tree**: 'The visualisation of the question tree does not represent the hierarchy of the question tree good enough to understand the hierarchy quickly.' by Pieter Buitenhuis on 06 Jan 2015 | 7 comments. Reactions: Agree (6), Disagree (2).
- Adding a wizard or tutorial**: 'After logging in it is quite unclear what to do next. Maybe a wizard or tutorial could be handy for the first time.' by Sinned on 16 Jan 2015 | 1 comment. Reactions: Agree (6), Disagree (0).

# REfine and the **CCRE** method



# REfine and the **CCRE** method

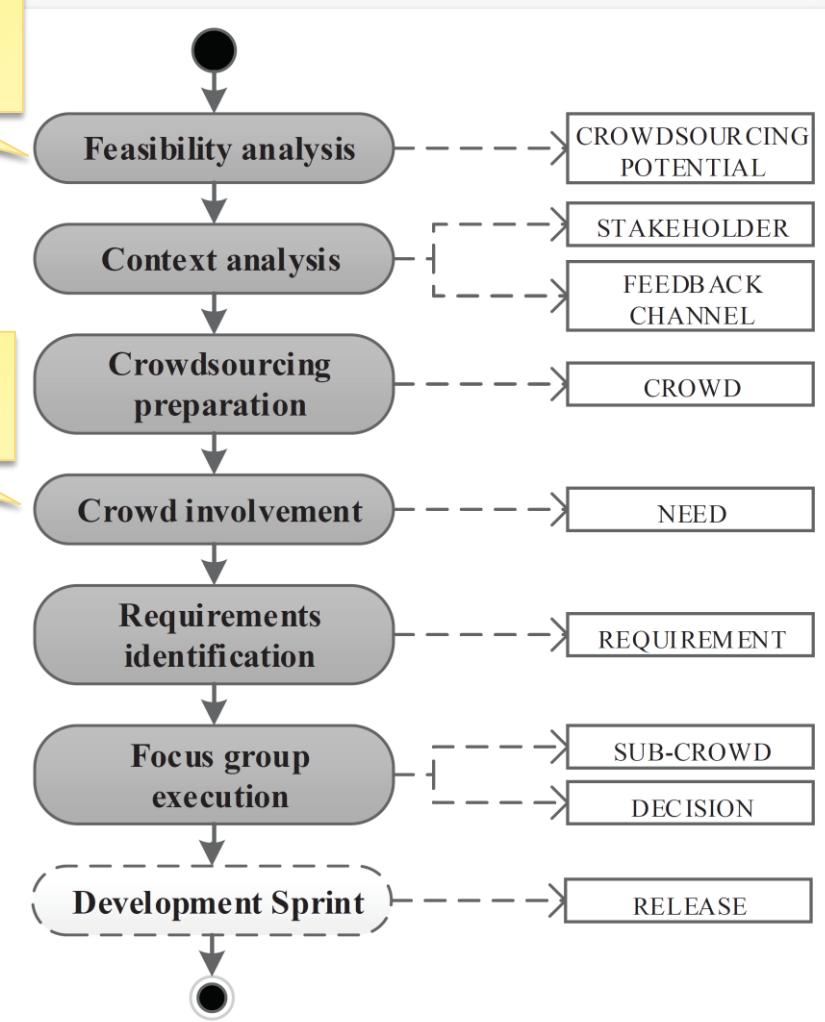
Don't always assume  
CrowdRE will work



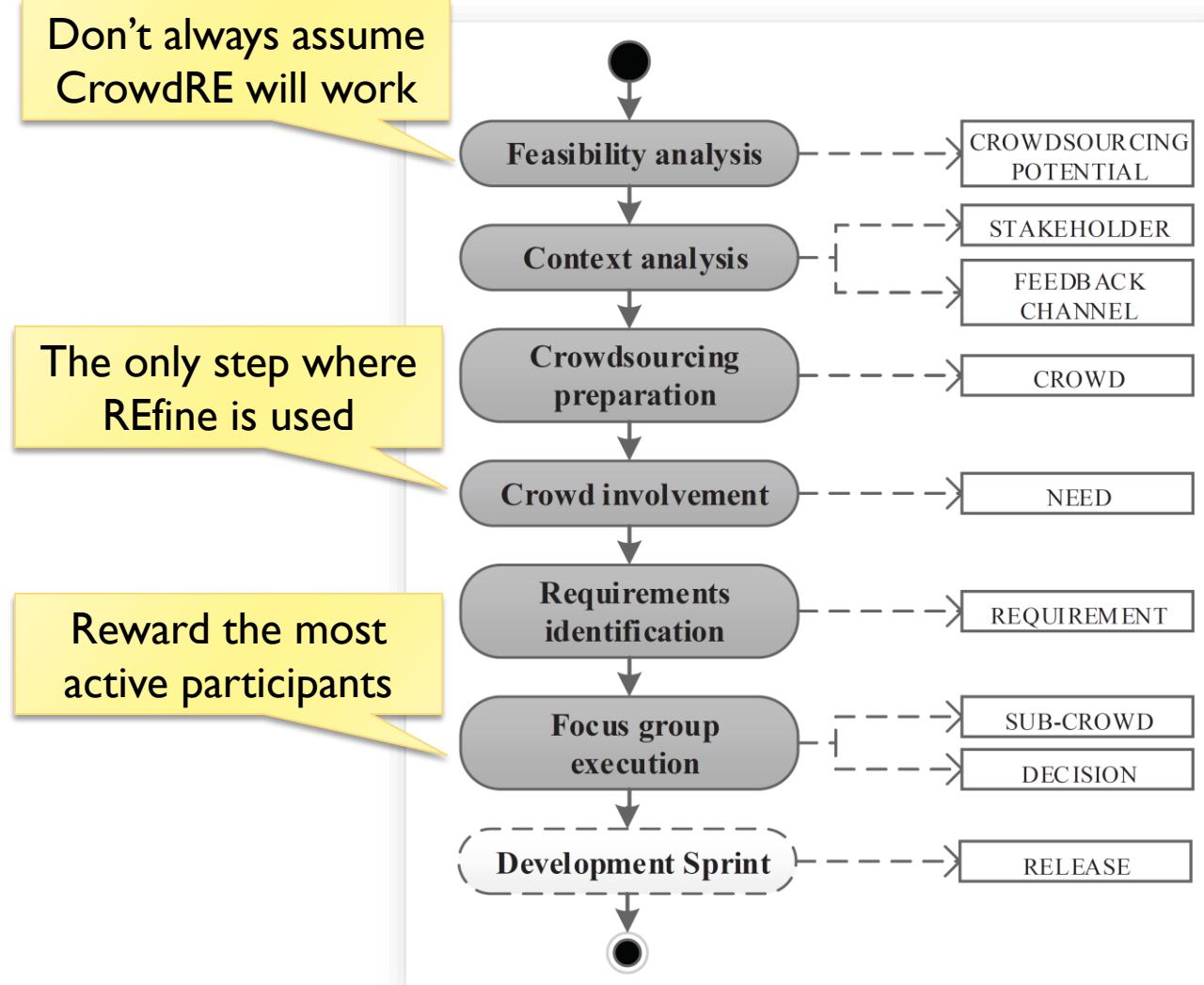
# REFINE and the CCRE method

Don't always assume CrowdRE will work

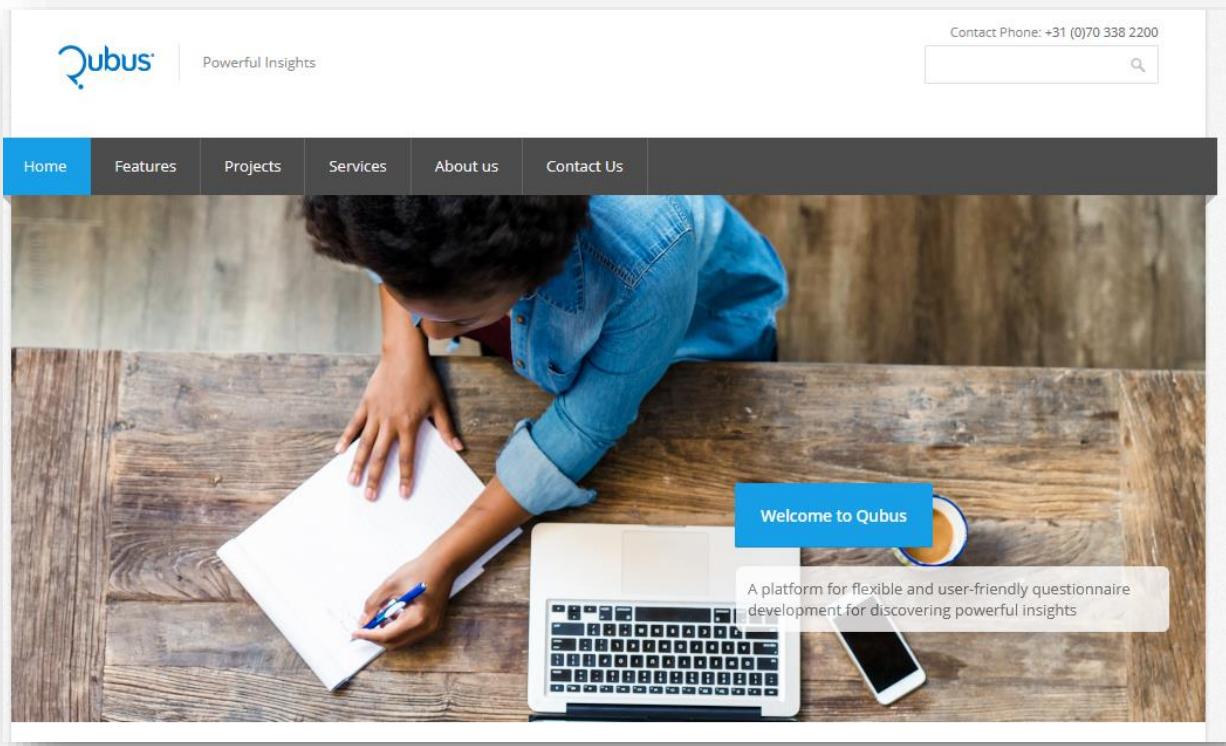
The only step where REfine is used



# REFINE and the CCRE method



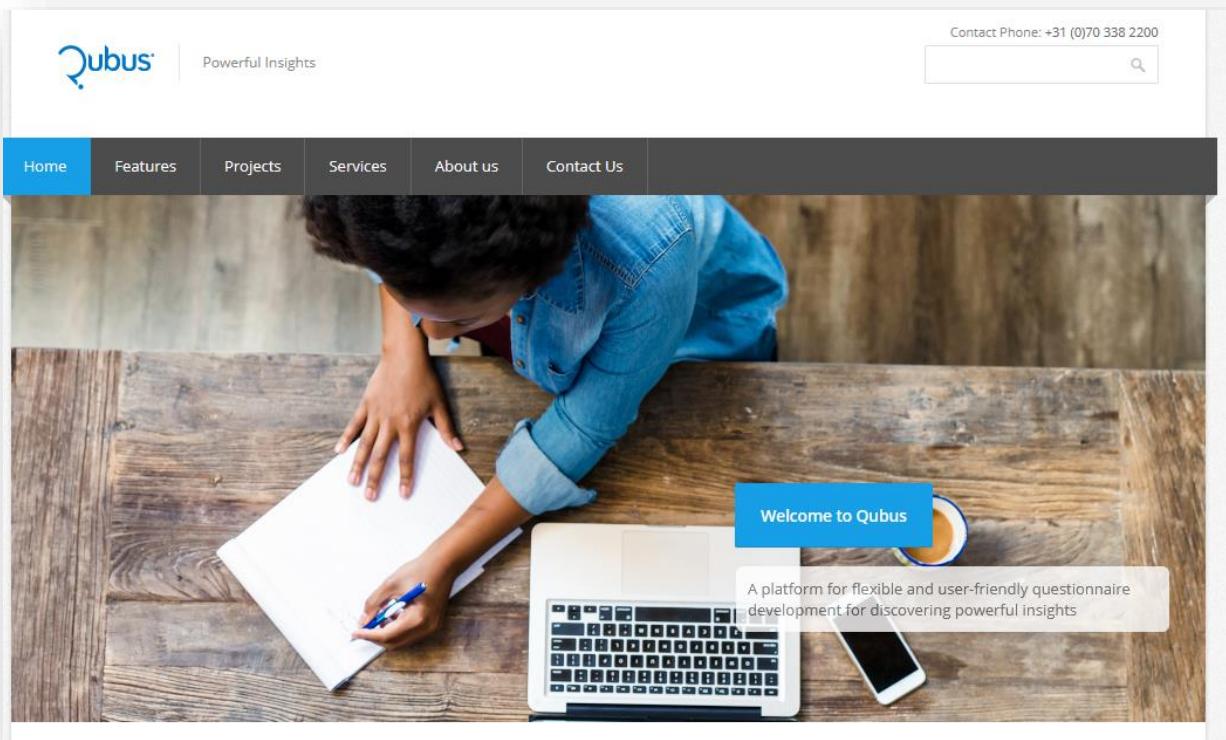
# Case study: **KPMG's Qubus**



- ▶ Platform for compliance auditors
  - ▶ Questionnaire-based
  - ▶ Customizable workflow



# Case study: **KPMG's Qubus**



- ▶ Platform for compliance auditors
  - ▶ Questionnaire-based
  - ▶ Customizable workflow
- ▶ Focus of CrowdRE elicitation:  
Improved usability for mobile devices



# Case study: KPMG's Qubus – the **users' perspective**

Stakeholder type	N	Needs		Comments		Votes		Points	
		avg.	tot.	avg.	tot.	avg.	tot.	avg.	tot.
Community manager	1	2	2	8	8	21	21	41	41
Product management	2	0	0	3	6	6	12	13.5	27
Development team	4	0.8	3	3.8	15	14.8	59	26.8	107
Expert	4	1	4	1.3	5	2.5	10	16	64
Client	1	0	0	0	0	3	3	3	3
End-User	1	1	1	0	0	3	3	11	11
Off-stage actor	6	1.8	11	0.5	3	4.2	25	20	120
<i>Total</i>	<b>19</b>	<b>1.1</b>	<b>21</b>	<b>1.9</b>	<b>37</b>	<b>6.8</b>	<b>130</b>	<b>19.6</b>	<b>373</b>



# Case study: KPMG's Qubus – the **users' perspective**

Stakeholder type	N	Needs		Comments		Votes		Points	
		avg.	tot.	avg.	tot.	avg.	tot.	avg.	tot.
Community manager	1	2	2	8	8	21	21	41	41
Product management	2	0	0	3	6	6	12	13.5	27
Development team	4	0.8	3	3.8	15	14.8	59	26.8	107
Expert	4	1	4	1.3	5	2.5	10	16	64
Client	1	0	0	0	0	3	3	3	3
End-User	1	1	1	0	0	3	3	11	11
Off-stage actor	6	1.8	11	0.5	3	4.2	25	20	120
<i>Total</i>	<b>19</b>	<b>1.1</b>	<b>21</b>	<b>1.9</b>	<b>37</b>	<b>6.8</b>	<b>130</b>	<b>19.6</b>	<b>373</b>

Not that large, some members highly involved



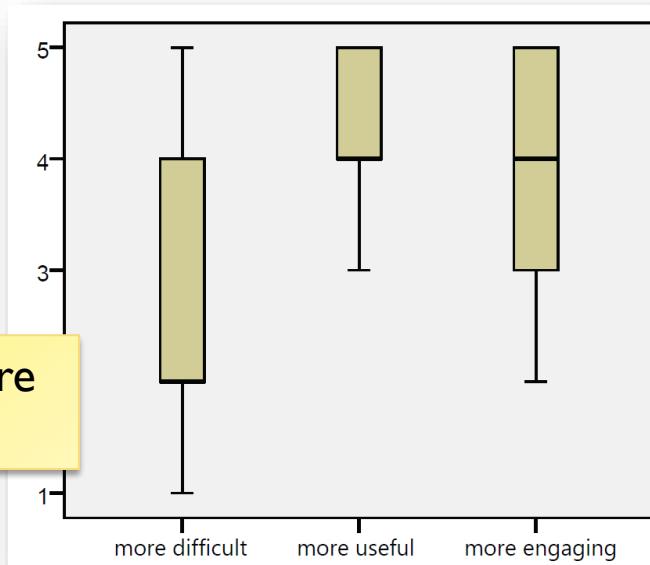
# Case study: KPMG's Qubus – the **users' perspective**

Stakeholder type	N	Needs		Comments		Votes		Points	
		avg.	tot.	avg.	tot.	avg.	tot.	avg.	tot.
Community manager	1	2	2	8	8	21	21	41	41
Product management	2	0	0	3	6	6	12	13.5	27
Development team	4	0.8	3	3.8	15	14.8	59	26.8	107
Expert	4	1	4	1.3	5	2.5	10	16	64
Client	1	0	0	0	0	3	3	3	3
End-User	1	1	1	0	0	3	3	11	11
Off-stage actor	6	1.8	11	0.5	3	4.2	25	20	120
<i>Total</i>	<b>19</b>	<b>1.1</b>	<b>21</b>	<b>1.9</b>	<b>37</b>	<b>6.8</b>	<b>130</b>	<b>19.6</b>	<b>373</b>

Not that large, some members highly involved



Yet, the experience was more useful and more engaging



# Case study: KPMG's Qubus – the **experts perspective**

---

- + Approachable method
- + Mostly useful for elicitation
- + Expected user adoption



Product managers  
(2 Qubus, 3 external)



# Case study: KPMG's Qubus – the **experts perspective**

---

- + Approachable method
- + Mostly useful for elicitation
- + Expected user adoption

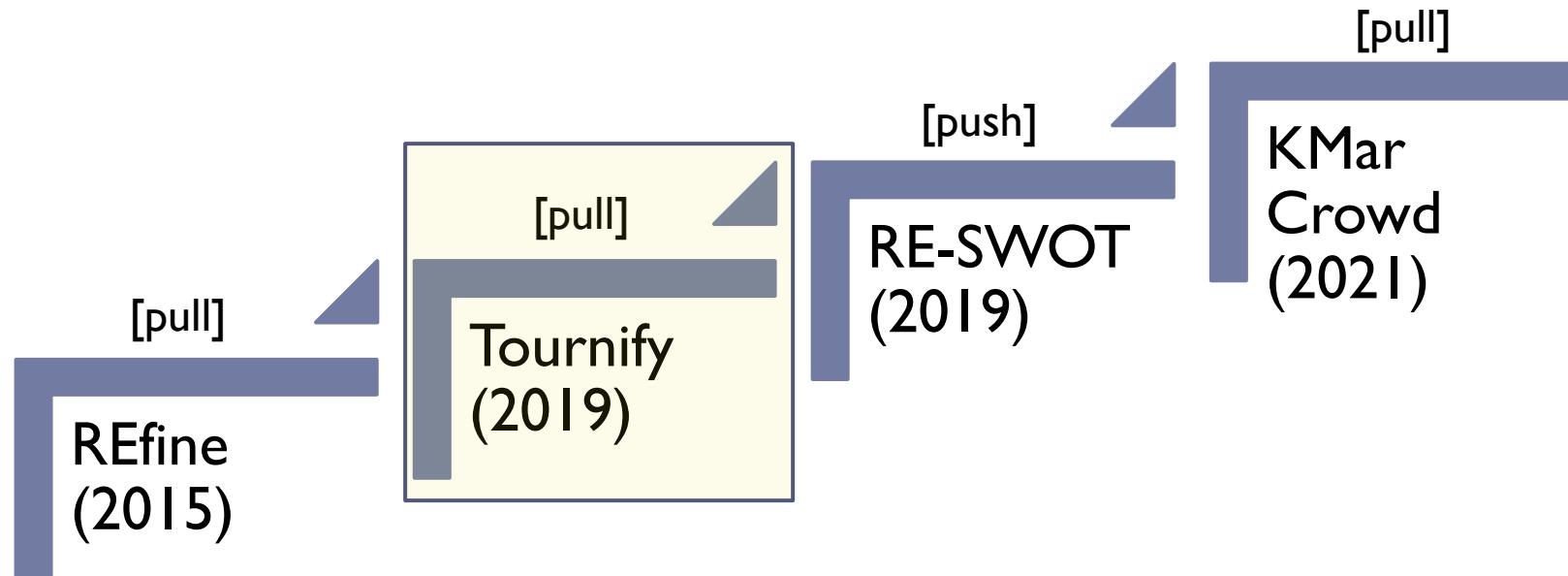


Product managers  
(2 Qubus, 3 external)

- Little incentive to return
- Possibly trivial needs
- Needs inadequate for backlog



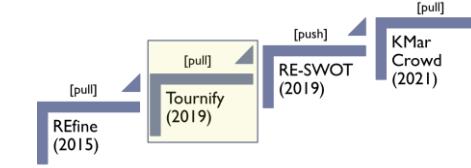
# Personal Journey: Overview



Abel Menkeld, Sjaak Brinkkemper, Fabiano Dalpiaz: User Story Writing in Crowd Requirements Engineering:  
The Case of a Web Application for Sports Tournament Planning. RE Workshops 2019: 174-179



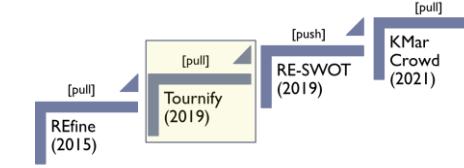
# Tournify – User Story Writing



Product managers  
(2 Qubus, 3 external)

- Little incentive to return
- Possibly trivial needs
- Needs inadequate for backlog

# Tournify – User Story Writing

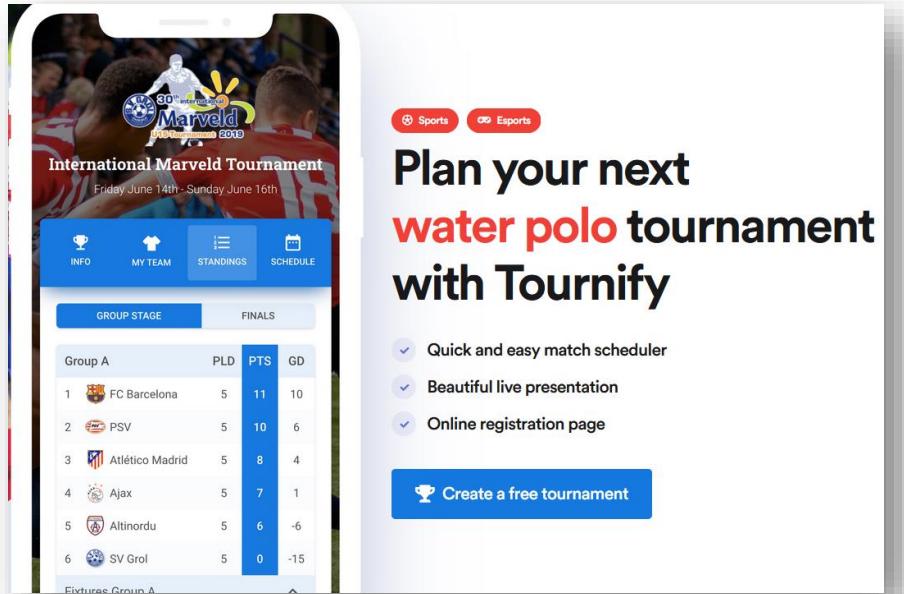


- Little incentive to return
- Possibly trivial needs
- Needs inadequate for backlog



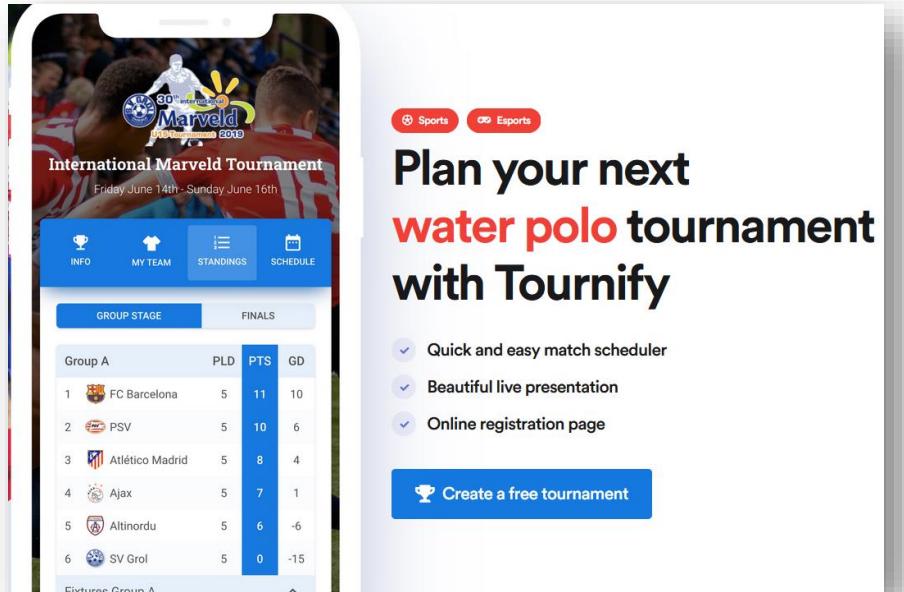
Can we obtain **more specific needs** by asking the crowd to write user stories?

# Tournify – Current Situation



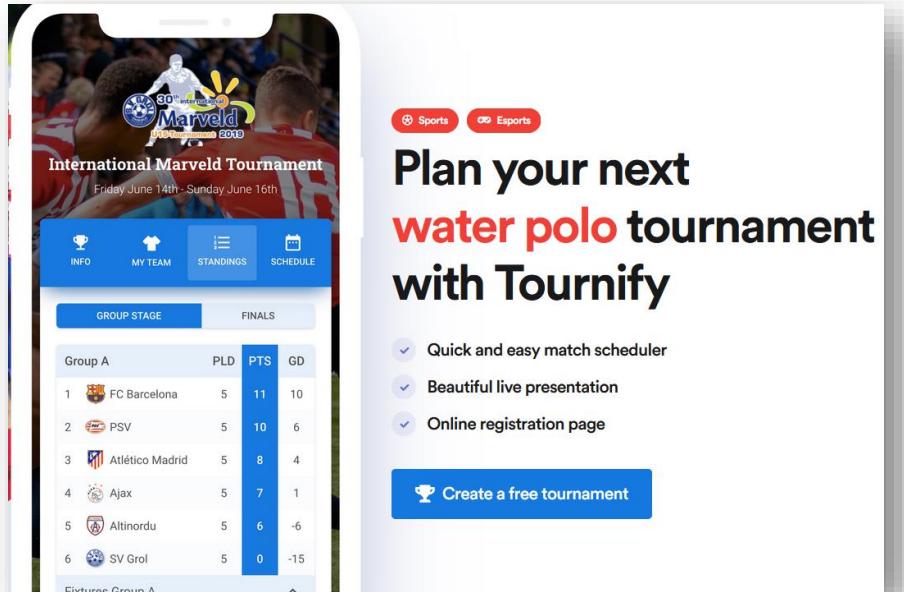
- ▶ Previously, user feedback was collected via e-mail and a support chat

# Tournify – Current Situation



- ▶ Previously, user feedback was collected via e-mail and a support chat
- ▶ The company adopts agile dev with user stories
  - ▶ For ~70% of the feedback, converting user feedback to user stories was easy/possible

# Tournify – Current Situation



- ▶ Previously, user feedback was collected via e-mail and a support chat
- ▶ The company adopts agile dev with user stories
  - ▶ For ~70% of the feedback, converting user feedback to user stories was easy/possible
  - ▶ **However, the reason behind feature requests is often omitted**

# Tournify – A CrowdRE Elicitation Platform

The screenshot shows a web-based feature request management system. At the top, there's a blue header bar with the title "Feature Requests". Below the header, there are five cards, each representing a user suggestion:

- Suggestion 1:** "Als organisator wil ik de mogelijkheid om de ploegen zelf als scheidrechter te laten fungeren en dus mee te laten nemen in de berekening van de schema's, zodat er geen wedstrijden doorgaan waarvan de ploegen op dezelfde moment moeten arbitreren".  
Created: 26 February 2019 - General | 1 comment | 3 likes.
- Suggestion 2:** "Als organisator wil ik dat bij de sheets voor de schermpresentaties start en eindtijden worden toegevoegd, zodat je kunt zorgen dat een sheet alleen op de gewenste tijdstippen wordt getoond zoals bijvoorbeeld voor het toernooi de welkomstboodschap of bijv tijdens lunchtijd".  
Created: 26 February 2019 - Presentation | 1 comment | 3 likes.
- Suggestion 3:** "Als organisator wil ik dat wedstrijden per veld, zodat de wedstrijdsecretaris het schema alleen hoeft op te lezen zonder nadenken".  
Created: 26 February 2019 - Presentation | 1 comment | 3 likes.
- Suggestion 4:** "Als organisator wil ik graag de mogelijkheid om logo's toe te voegen aan de presentatie. Dit zoals het vlaggetje voor de deelnemersnaam maar dan hun eigen logo, zodat de logo's ook mooi kunnen worden getoond in de wedstrijdschema's en standen".  
Status: In development  
Created: 26 February 2019 - Presentation | 2 comments | 9 likes.
- Suggestion 5:** "Als organisator wil ik dat fase 1 niet meer getoond wordt zodra de volgende fase begint, zodat deelnemers niet gedwongen worden om zelf te klikken op de volgende".  
Created: 26 February 2019 - Presentation | 1 comment | 3 likes.

# Tournify – A CrowdRE Elicitation Platform

The image shows a screenshot of the Tournify platform. On the left, there is a sidebar titled "Feature Requests" with four items listed:

- Als organisator wil ik de mogelijkheid om de ploegen zelf als scheidrechter te laten fungeren en dus mee te laten nemen in de berekening van de schema's, zodat er geen wedstrijden doorgaan waarvan de ploegen op dezelfde moment moeten arbitreren  
26 February 2019 - General 1 like 3
- Als organisator wil ik dat bij de sheets voor de schermpresentaties start en eindtijden worden toegevoegd, zodat je kunt zorgen dat een sheet alleen op de gewenste tijdstippen wordt getoond zoals bijvoorbeeld voor het toernooi de welkomstbericht of bijv tijdens lunchtijd.  
26 February 2019 - Presentation 1 like 3
- Als organisator wil ik dat de wedstrijden per veld, zodat de wedstrijdsecretaris het schema alleen hoeft op te lezen zonder nadenken  
26 February 2019 - Presentation 1 like 3
- Als organisator wil ik dat fase 1 niet meer getoond wordt zodra de volgende fase begint, zodat deelnemers niet gedwongen worden om zelf te klikken op de volgende start aan altijd aan de rechterkant.

The main area of the screen shows a "New idea" process in four steps:

- Step 1 of 4**: How do you use Tournify?
  - I organize a tournament
  - I participate in a tournament
  - I support a team in a tournament[TO STEP 2](#)
- Step 2 of 4**: What would you like to do with Tournify?

I want to export the match schedule to PDF

[STEP BACK](#) [TO STEP 3](#)
- Step 3 of 4**: Why do you want to export the match schedule to PDF?

So that I can print it out and have it as a backup in case the connection is lost.

[STEP BACK](#) [TO STEP 4](#)
- Step 4 of 4**: Verify your idea, choose a category and submit it.

As an organizer I want to export the match schedule to PDF, so that I can print it out and have it as a backup in case the connection is lost.

[STEP BACK](#) [SUBMIT IDEA](#)

At the bottom of the main area, there are buttons for "General", "Participants", "Format", "Schedule", "Presentation", "Results", and "Other".

# Tournify – A CrowdRE Elicitation Platform

The image shows a screenshot of the Tournify platform. On the left, there is a sidebar titled "Feature Requests" with several user posts. The top post discusses allowing organizers to have the schema always start at the right side. The second post discusses adding sleep functions to presentations. The third post discusses adding time sheets for presentations. The fourth post discusses exporting match schedules to PDF.

The main area of the screen shows a "New idea" step in a crowdRE elicitation process. A yellow callout bubble points to the "Pre-defined roles" section. The question asks how users use Tournify, with three options: "I organize a tournament" (selected), "I participate in a tournament", and "I support a team in a tournament". Below this is a "TO STEP 2" button.

The next step, "Step 2 of 4", asks what users want to do with Tournify. The selected option is "I want to export the match schedule to PDF". Buttons for "STEP BACK" and "TO STEP 3" are available.

The final step, "Step 4 of 4", asks users to verify their idea and choose a category. The text states: "As an organizer I want to export the match schedule to PDF, so that I can print it out and have it as a backup in case the connection is lost." Categories listed include General, Participants, Format, Schedule, Presentation, Results, and Other. Buttons for "STEP BACK" and "SUBMIT IDEA" are present.

# Tournify – A CrowdRE Elicitation Platform

The image shows a screenshot of the Tournify platform. On the left, there's a sidebar titled "Feature Requests" with several user posts. The top post discusses wanting to add a feature where organizers can start sheets from the right side. The second post discusses adding sleep functions to presentation sheets. The third post discusses wanting to export match schedules to PDF.

The main part of the screen shows a "New idea" process:

- Step 1 of 4: Pre-defined roles**

How do you use Tournify?

I organize a tournament  
 I participate in a tournament  
 I support a team in a tournament
- Step 2 of 4: What feature?**

What would you like to do with Tournify?

I want to export the match schedule to PDF
- Step 3 of 4: Why?**

Why do you want to export the match schedule to PDF?

So that I can print it out and have it as a backup in case the connection is lost.
- Step 4 of 4: Verify your idea, choose a category and submit it.**

Verify your idea, choose a category and submit it.

As an organizer I want to export the match schedule to PDF, so that I can print it out and have it as a backup in case the connection is lost.

General Participants Format Schedule Presentation Results  
Other

# Tournify – A CrowdRE Elicitation Platform

The screenshot displays the Tournify platform interface, featuring a sidebar of feature requests and a main area for idea elicitation.

**Feature Requests:**

- Als organisator wil ik de mogelijkheid om de ploegen zelf als scheidrechter te laten fungeren en dus mee te laten nemen in de berekening van de schema's, zodat er geen wedstrijden doorgaan waarvan de ploegen op dezelfde moment moeten arbitreren  
26 February 2019 - General 1 like
- Als organisator wil ik dat bij de sheets voor de schermpresentaties start en eindtijden worden toegevoegd, zodat je kunt zorgen dat een sheet alleen op de gewenste tijdstippen wordt getoond zoals bijvoorbeeld voor het toernooi de welkomstbericht of bijv tijdens lunchtijd.  
26 February 2019 - Presentation 1 like
- wedstrijden per veld, zodat de wedstrijdsecretaris het schema alleen hoeft op te lezen zonder nadenken  
26 February 2019 - Presentation 1 like 3
- Als organisator wil ik de dia's met een sleepfunctie of wijsbare volgordenummering kunnen laten wijzigen van volgorde, zodat je de presentatie niet volledig opnieuw hoeft te maken als er een extra dia tussenkomt.  
26 February 2019 - Presentation 5 likes
- Als organisator wil ik dat fase 1 niet meer getoond wordt zodra de volgende fase begint, zodat deelnemers niet gedwongen worden om zelf te klikken op de volgende

**Idea Elicitation Process:**

- Step 1: New idea**

How do you use Tournify?

I organize a tournament  
 I participate in a tournament  
 I support a team in a tournament

**Pre-defined roles**
- Step 2 of 4: New idea**

What would you like to do with Tournify?

I want to export the match schedule to PDF

**What feature?**
- Step 3 of 4: New idea**

Why do you want to export the match schedule to PDF?

So that I can print it out and have it as a backup in case the connection is lost.

**Why?**
- Step 4 of 4: New idea**

Verify your idea, choose a category and submit it.

As an organizer I want to export the match schedule to PDF, so that I can print it out and have it as a backup in case the connection is lost.

**Tags as metadata**

General Participants Format Schedule Presentation Results  
Other

STEP BACK TO STEP 3 SUBMIT IDEA

# Tournify – Case Study Execution

---

- ▶ **337 invited participants**
  - ▶ Previous feature requesters
  - ▶ Newsletter subscribers
  - ▶ Recent paying clients
- ▶ **Data collection for 5 weeks**
  - ▶ Same duration as REfine, but users outside the organization
- ▶ **Users informed via a snack bar message on the Tournify page**

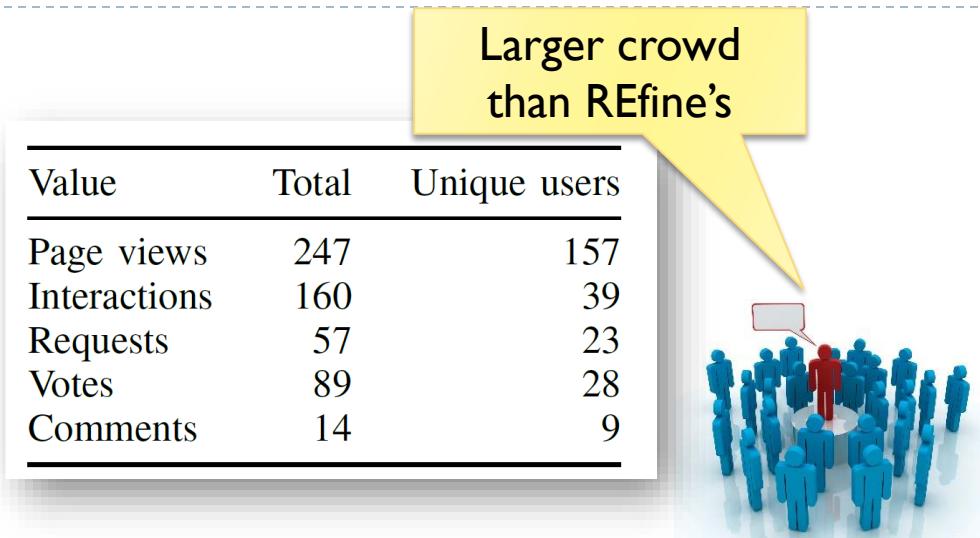
# Tournify – Results

---

Value	Total	Unique users
Page views	247	157
Interactions	160	39
Requests	57	23
Votes	89	28
Comments	14	9



# Tournify – Results



# Tournify – Results

Value	Total	Unique users
Page views	247	157
Interactions	160	39
Requests	57	23
Votes	89	28
Comments	14	9

Larger crowd  
than REfine's

These are not  
perfectly written  
requirements...



criterion	# USs with defect	% USs with defect
Well-formed	3	5.4
Atomic	5	8.9
Minimal	24	42.9
Conceptual	5	8.9
Problem-oriented	8	14.3
Unambiguous	9	16.0
Full sentence	19	33.9
Estimatable	3	5.4

# Tournify – Results

Value	Total	Unique users
Page views	247	157
Interactions	160	39
Requests	57	23
Votes	89	28
Comments	14	9

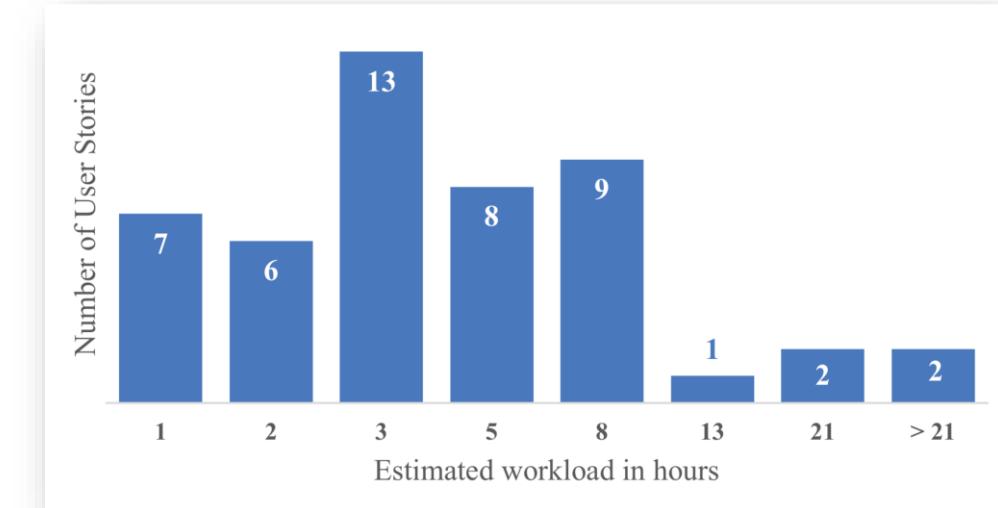
Larger crowd than REfine's

These are not perfectly written requirements...

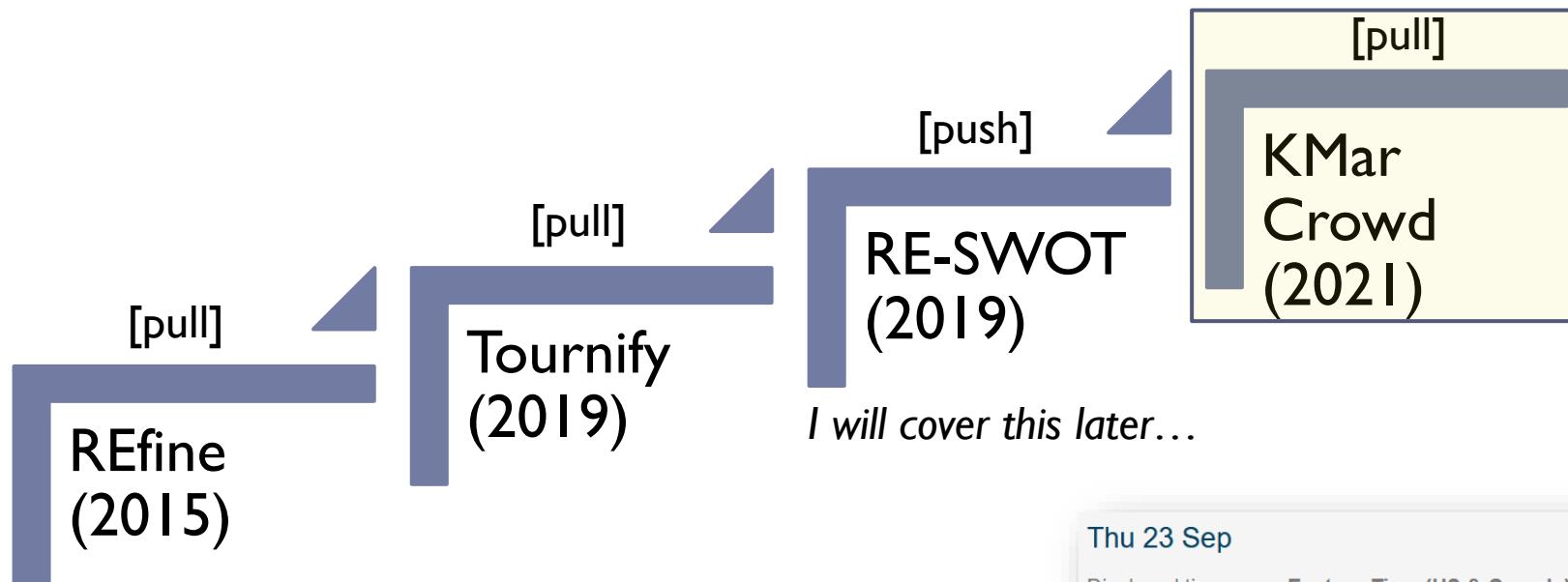
Yet, most are sufficiently specific



criterion	# USs with defect	% USs with defect
Well-formed	3	5.4
Atomic	5	8.9
Minimal	24	42.9
Conceptual	5	8.9
Problem-oriented	8	14.3
Unambiguous	9	16.0
Full sentence	19	33.9
Estimatable	3	5.4



# Personal Journey: Overview



Thu 23 Sep

Displayed time zone: Eastern Time (US & Canada) [change](#)

10:00 30m ★ **CrowdRE in a Governmental Setting: Lessons from Two Case Studies** INDUSTRIAL INNOVATION PAPER

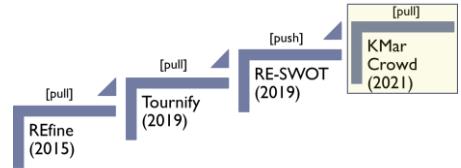
Talk Industrial Innovation Papers

Jelle Wouters Royal Netherlands Marechaussee, Rob Janssen Royal Netherlands Marechaussee, Bas van Hulst Royal Netherlands Marechaussee, John van Veenhuizen Royal Netherlands Marechaussee, Fabiano Dalpiaz Utrecht University, Sjaak Brinkkemper Utrecht University

Pre-print



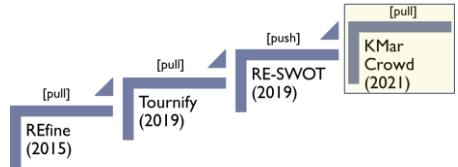
# KMar Crowd – Premises



Measurement	Menkfeld	REFine	GARUSO
Duration in days	35	35	92
Participants:			
- Invited	unk.	37	unk.
- Accessed	157	19	726
- Active	39	19	32
Ideas	57	21	56
Logins	247	unk.	unk.
Votes	89	130	160
Comments	14	37	unk.
Ideas / Accessed	0.36	1.11	0.08



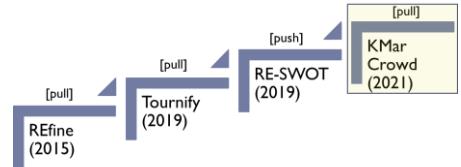
# KMar Crowd – Premises



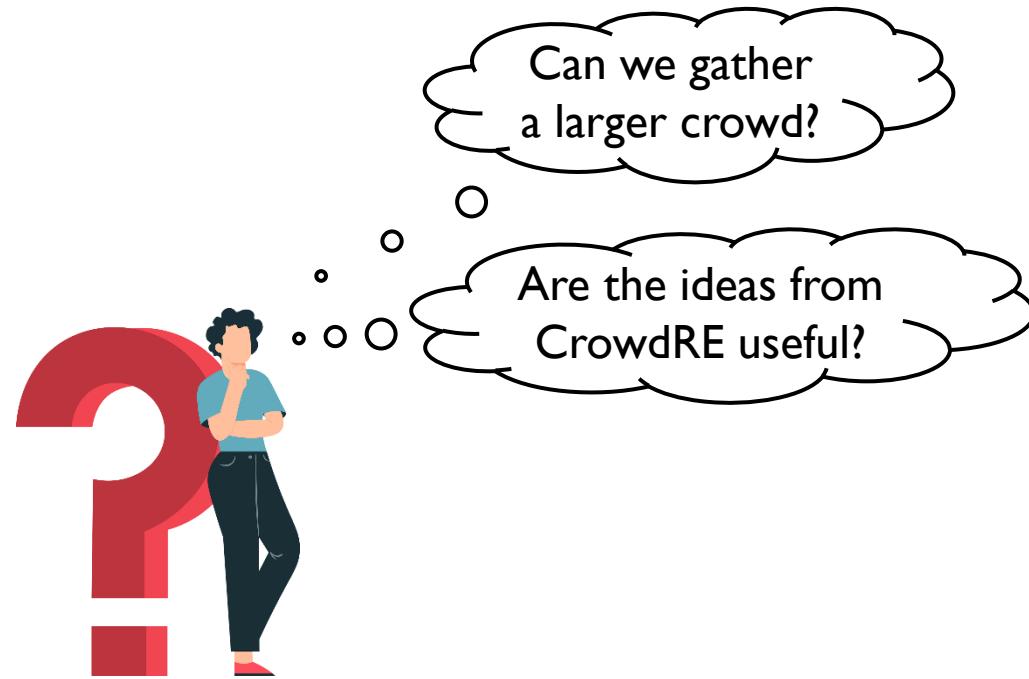
Measurement	Menkvelt	REFine	GARUSO
Duration in days	35	35	92
Participants:			
- Invited	unk.	37	unk.
- Accessed	157	19	726
- Active	39	19	32
Ideas	57	21	56
Logins	247	unk.	unk.
Votes	89	130	160
Comments	14	37	unk.
Ideas / Accessed	0.36	1.11	0.08



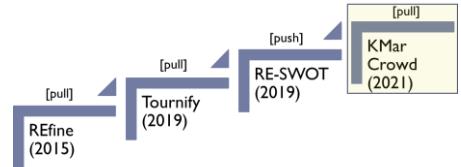
# KMar Crowd – Premises



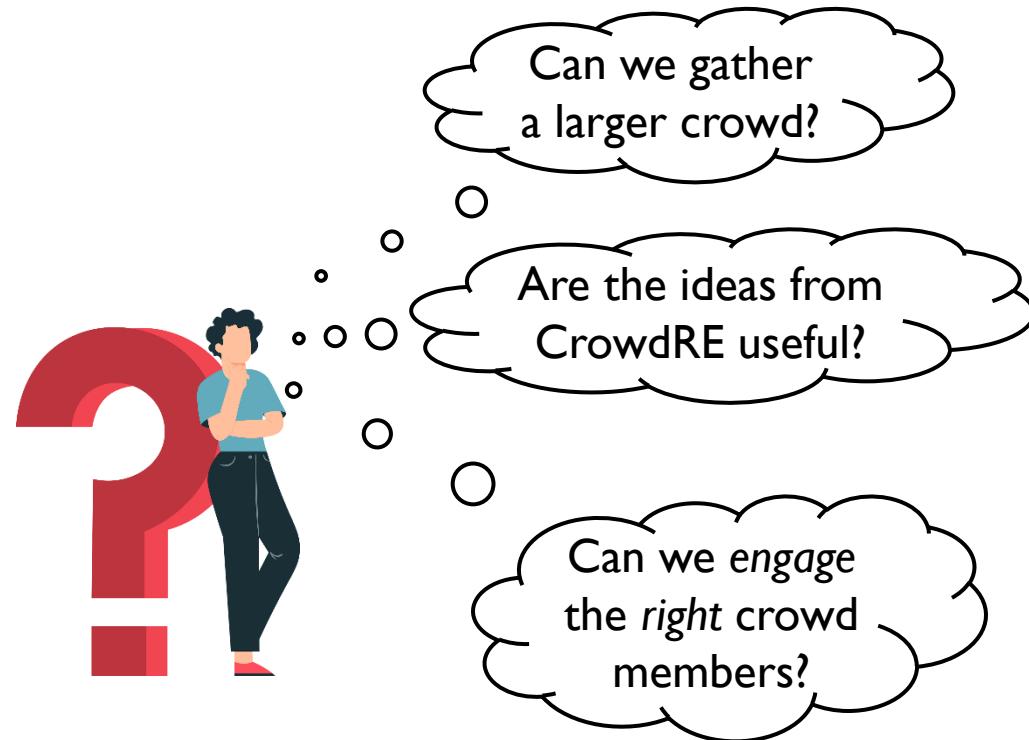
Measurement	Menkfeld	REFine	GARUSO
Duration in days	35	35	92
Participants:			
- Invited	unk.	37	unk.
- Accessed	157	19	726
- Active	39	19	32
Ideas	57	21	56
Logins	247	unk.	unk.
Votes	89	130	160
Comments	14	37	unk.
Ideas / Accessed	0.36	1.11	0.08



# KMar Crowd – Premises



Measurement	Menkfeld	REFine	GARUSO
Duration in days	35	35	92
Participants:			
- Invited	unk.	37	unk.
- Accessed	157	19	726
- Active	39	19	32
Ideas	57	21	56
Logins	247	unk.	unk.
Votes	89	130	160
Comments	14	37	unk.
Ideas / Accessed	0.36	1.11	0.08



# KMar Crowd – Case Studies

- ▶ Research conducted at the Royal Netherlands Marechaussee
  - ▶ 7,000 employees, split across 24 brigades
  - ▶ Canonical action research



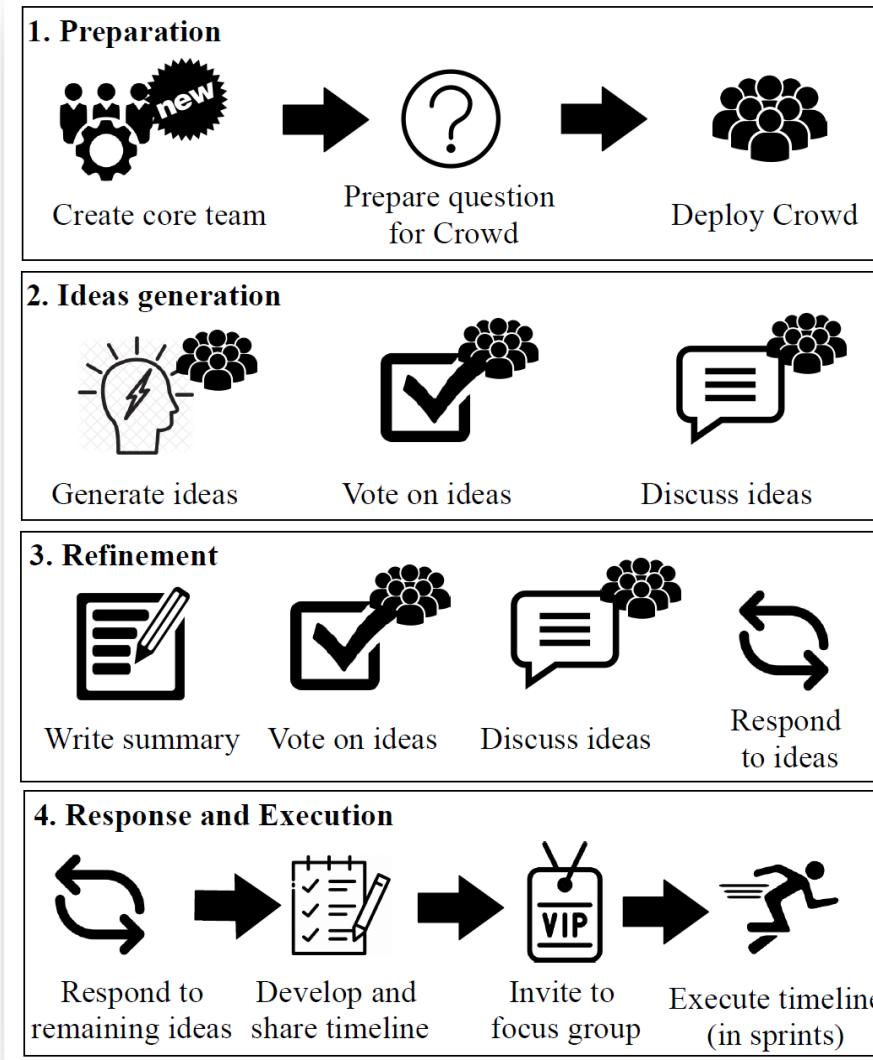
# KMar Crowd – Case Studies

- ▶ Research conducted at the Royal Netherlands Marechaussee
  - ▶ 7,000 employees, split across 24 brigades
  - ▶ Canonical action research
- ▶ Three research cycles
  1. Pilot to test the Kmar-Crowd platform
  2. S-Sys: 1 brigade, **comparison to existing requirements**
  3. V-Sys: **13 brigades**, no prior requirements



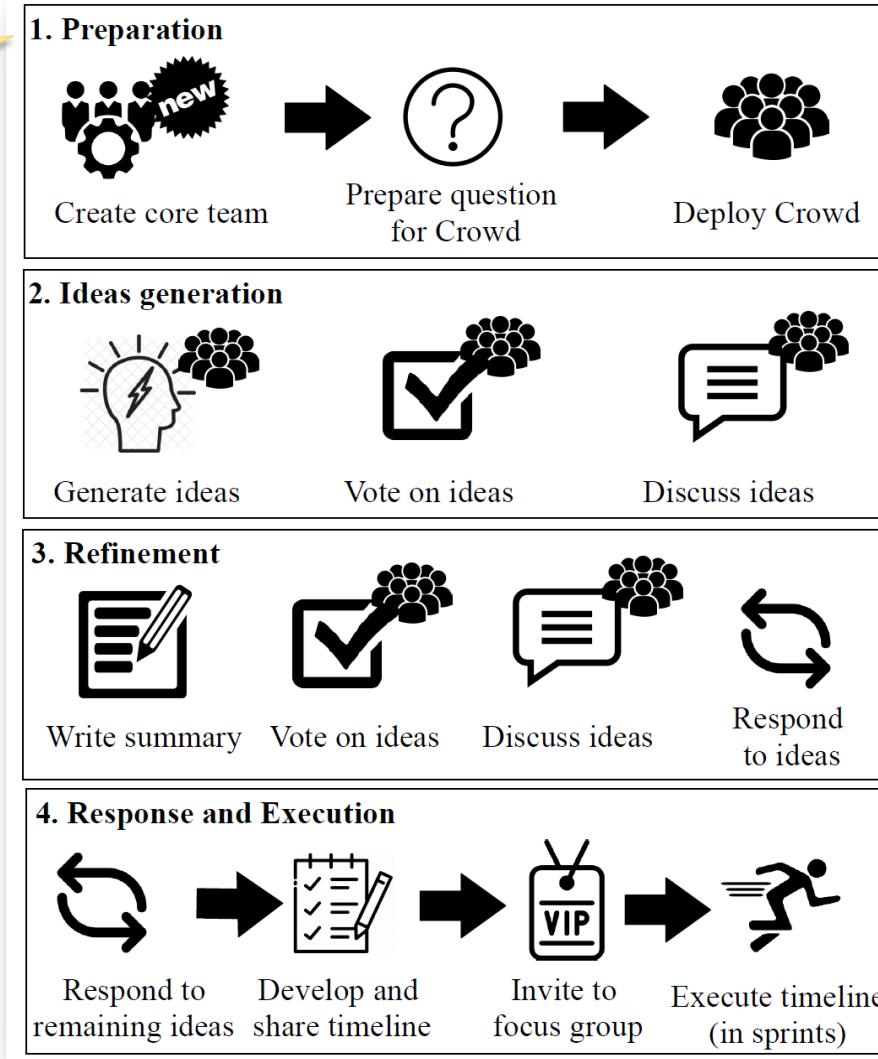
Royal Netherlands Marechaussee

# KMar Crowd – **CrowdRE** method



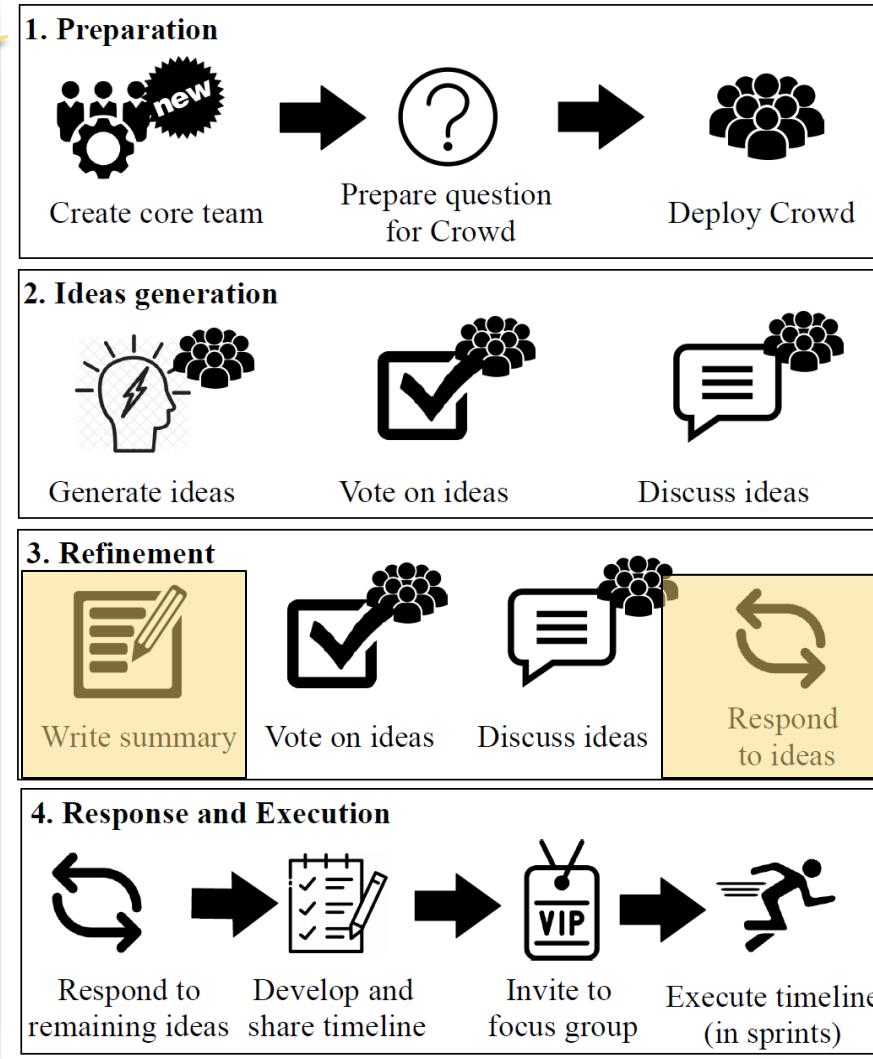
# KMar Crowd – CrowdRE method

Extends the ideas  
of REfine



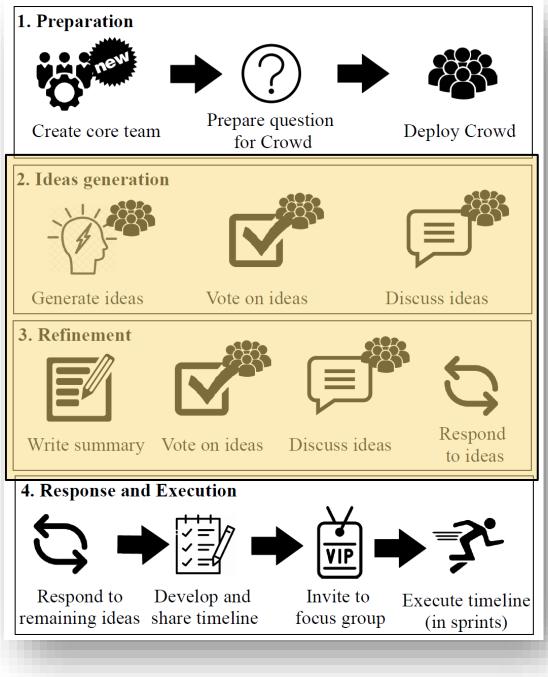
# KMar Crowd – CrowdRE method

Extends the ideas  
of REfine

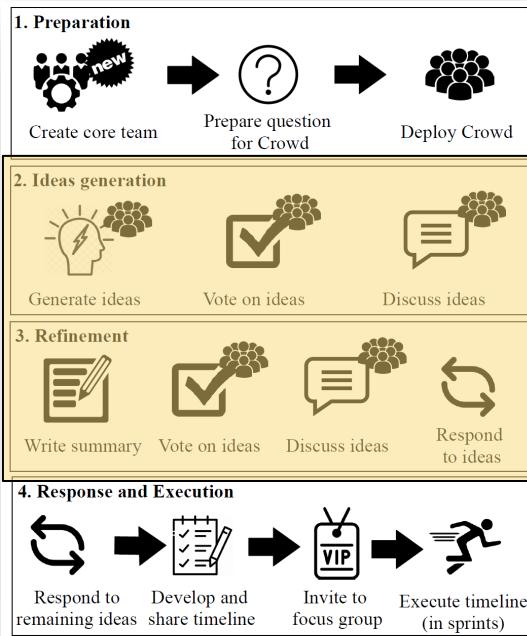


The moderator's  
role is crucial

# KMar Crowd – CrowdRE platform



# KMar Crowd – CrowdRE platform



**De nieuwe Module Vreemdelingen**

Laat van je horen! Wat wil JIJ in de nieuwe Module Vreemdelingen hebben? Vul je eigen ideeën in... of stem op die van anderen.

GA NAAR IDEEENBORD

**Wat kan ik winnen?**

**1 ster**

50 Punt  
Je hebt nu een ster. We kiezen ervoor om je meteen een ster te geven. Immers, als je meedenkt over de toekomst van belangrijke IT-applicaties binnen de KMar, dan is dat zeker een ster waard!

Je kunt echter nog veel hoger komen! Wie weet verdien jij binnenkort wel een tweede ster!

**2 sterren**

50 Punt  
Je bent al goed op weg: je hebt nu twee sterren. Je hebt een idee geformuleerd en heb je ook al op twee andere ideeën gestemd. Daarnaast heb je een idee van een van je collega van commentaar voorzien.

Ga vooral zo door, want vanaf vier sterren zijn er prijzen te winnen!

**3 sterren**

50 Punt  
Goed bezig, je hebt vier sterren! Je hebt twee ideeën aangeleverd, en er op vijf gestemd. Daarnaast heb je ook twee ideeën van commentaar voorzien, en ben je regelmatig op ons platform te vinden.

Jelle  
Wijng | Uitgaven

Idee toevoegen

Vul een pakkende titel in...

Wat is je idee?

Welke rol past het beste bij je?

Verbalist

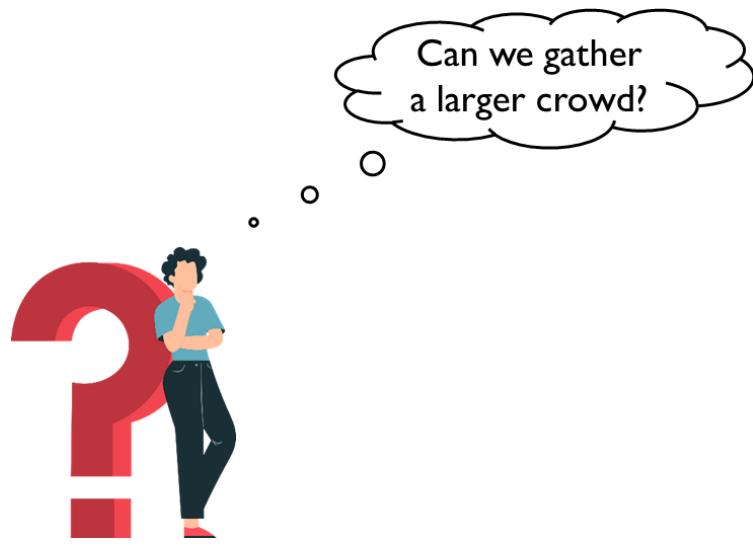
Waarom wil je dat?

Hoe doe je dat nu in BPS?

Browse...

VERSTUUR →

# KMar Crowd – Results – largeness



Measurement	S-Sys	V-Sys	Menkveld	REFine	GARUSO
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32
Ideas	32	78	57	21	56
Logins	240	623	247	unk.	unk.
Votes	316	531	89	130	160
Comments	28	78	14	37	unk.
Ideas / Accessed	0.24	0.20	0.36	1.11	0.08

# KMar Crowd – Results – largeness



Measurement	S-Sys	V-Sys	Menkveld	REFine	GARUSO
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32
Ideas	32	78	57	21	56
Logins	240	623	247	unk.	unk.
Votes	316	531	89	130	160
Comments	28	78	14	37	unk.
Ideas / Accessed	0.24	0.20	0.36	1.11	0.08

# KMar Crowd – Results – largeness



Measurement	S-Sys	V-Sys	Menkveld	REfine	GARUSO
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32
Ideas	32	78	57	21	56
Logins	240	623	247	unk.	unk.
Votes	316	531	89	130	160
Comments	28	78	14	37	unk.
Ideas / Accessed	0.24	0.20	0.36	1.11	0.08



# KMar Crowd – Results – origin

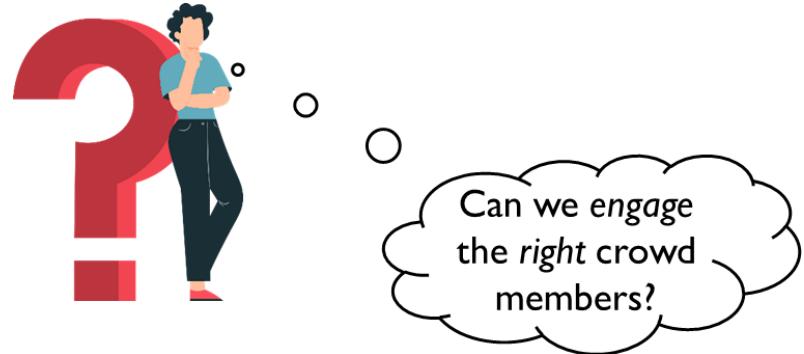


TABLE IV  
ACTIVITY PER USER TYPE IN THE S-SYS CASE STUDY (N=135).

Origin	% of total	Per user activity		
		Ideas	Votes	Logins
Operational employee	58.52%	0.23	2.89	1.84
Middle management	8.15%	0.82	4.00	2.55
Non-targeted employee	33.34%	0.11	0.88	1.55

TABLE VII  
ACTIVITY PER USER RANK IN THE V-SYS CASE STUDY (N=385).

Origin	% of total	Per user activity		
		Ideas	Votes	Logins
Soldiers	0.00%	-	-	-
Corporals	3.64%	0.36	3.43	2.43
NCOs	76.62%	0.22	1.41	1.63
Sub-altern officers	11.43%	0.10	1.45	1.68
Head officers	2.34%	0.00	0.00	1.22
Citizens	5.97%	0.09	0.13	1.30

# KMar Crowd – Results – origin

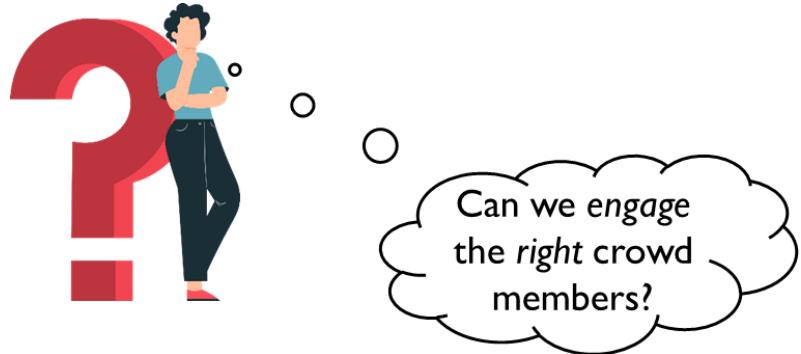


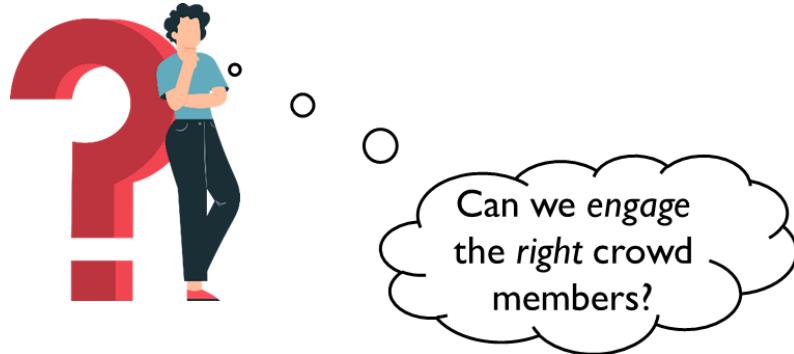
TABLE IV  
ACTIVITY PER USER TYPE IN THE S-SYS CASE STUDY (N=135).

Origin	% of total	Per user activity		
		Ideas	Votes	Logins
Operational employee	58.52%	0.23	2.89	1.84
Middle management	8.15%	0.82	4.00	2.55
Non-targeted employee	33.34%	0.11	0.88	1.55

TABLE VII  
ACTIVITY PER USER RANK IN THE V-SYS CASE STUDY (N=385).

Origin	% of total	Per user activity		
		Ideas	Votes	Logins
Soldiers	0.00%	-	-	-
Corporals	3.64%	0.36	3.43	2.43
NCOs	76.62%	0.22	1.41	1.63
Sub-altern officers	11.43%	0.10	1.45	1.68
Head officers	2.34%	0.00	0.00	1.22
Citizens	5.97%	0.09	0.13	1.30

# KMar Crowd – Results – origin



Operational employees were reached

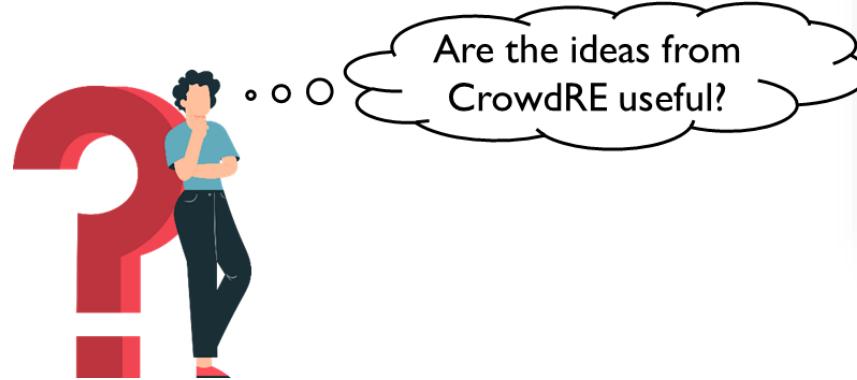
TABLE IV  
ACTIVITY PER USER TYPE IN THE S-SYS CASE STUDY (N=135).

Origin	% of total	Per user activity		
		Ideas	Votes	Logins
Operational employee	58.52%	0.23	2.89	1.84
Middle management	8.15%	0.82	4.00	2.55
Non-targeted employee	33.34%	0.11	0.88	1.55

TABLE VII  
ACTIVITY PER USER RANK IN THE V-SYS CASE STUDY (N=385).

Origin	% of total	Per user activity		
		Ideas	Votes	Logins
Soldiers	0.00%	-	-	-
Corporals	3.64%	0.36	3.43	2.43
NCO	76.62%	0.22	1.41	1.63
Officers	11.43%	0.10	1.45	1.68
	2.34%	0.00	0.00	1.22
	5.97%	0.09	0.13	1.30

# KMar Crowd – Results – usefulness



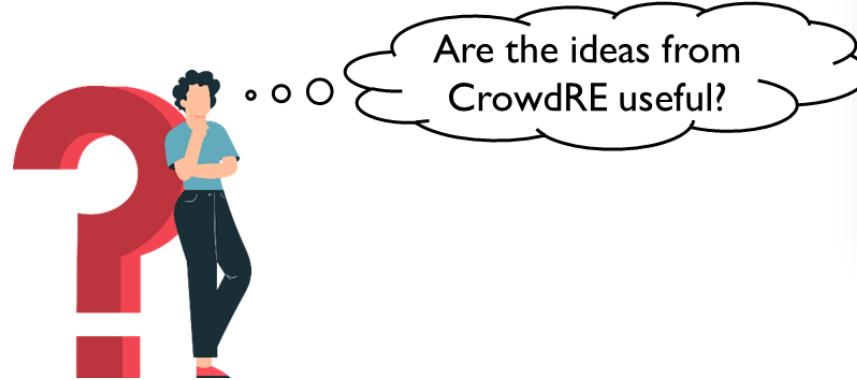
USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP		59.5
		27.8
		12.7
Granularity	Epic	40.5
	User Story	54.4
	Not applicable	5.1

# KMar Crowd – Results – usefulness



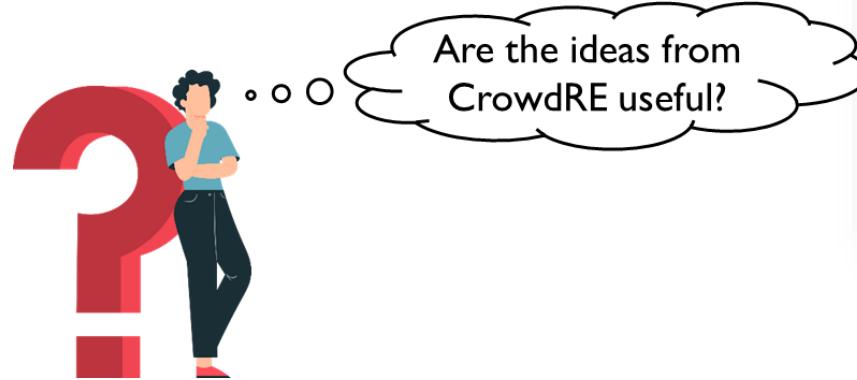
USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Granularity	Enough for MVP	59.5
	Enough for product	27.8
	Epic	40.5
Granularity	User Story	54.4
	Not applicable	5.1

# KMar Crowd – Results – usefulness



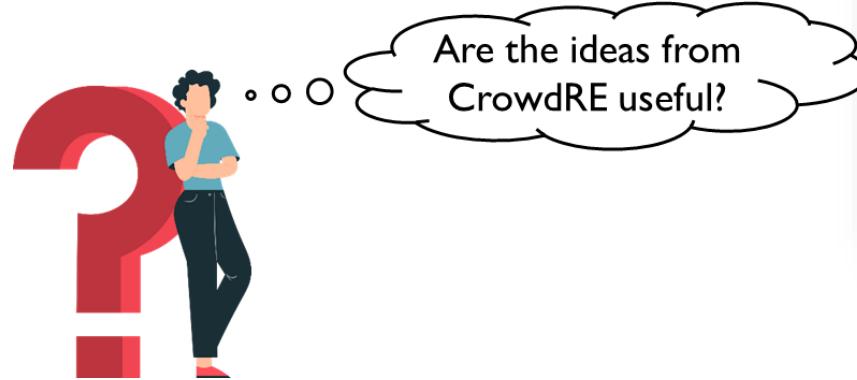
USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Granularity	Enough for MVP	59.5
	Enough for product	27.8
	Epic	40.5
Granularity	User Story	54.4
	Not applicable	5.1

# KMar Crowd – Results – usefulness



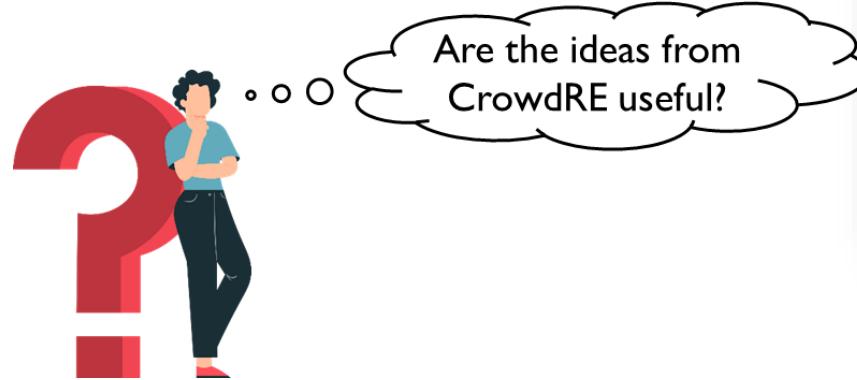
USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP	Enough for MVP	59.5
	Enough for product	27.8
	Epic	40.5
Granularity	User Story	54.4
	Not applicable	5.1

# KMar Crowd – Results – usefulness



USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP	Enough for MVP	59.5
	Enough for product	27.8
Granularity	Epic	40.5
	User Story	54.4
	Not applicable	5.1

# KMar Crowd – Results – usefulness



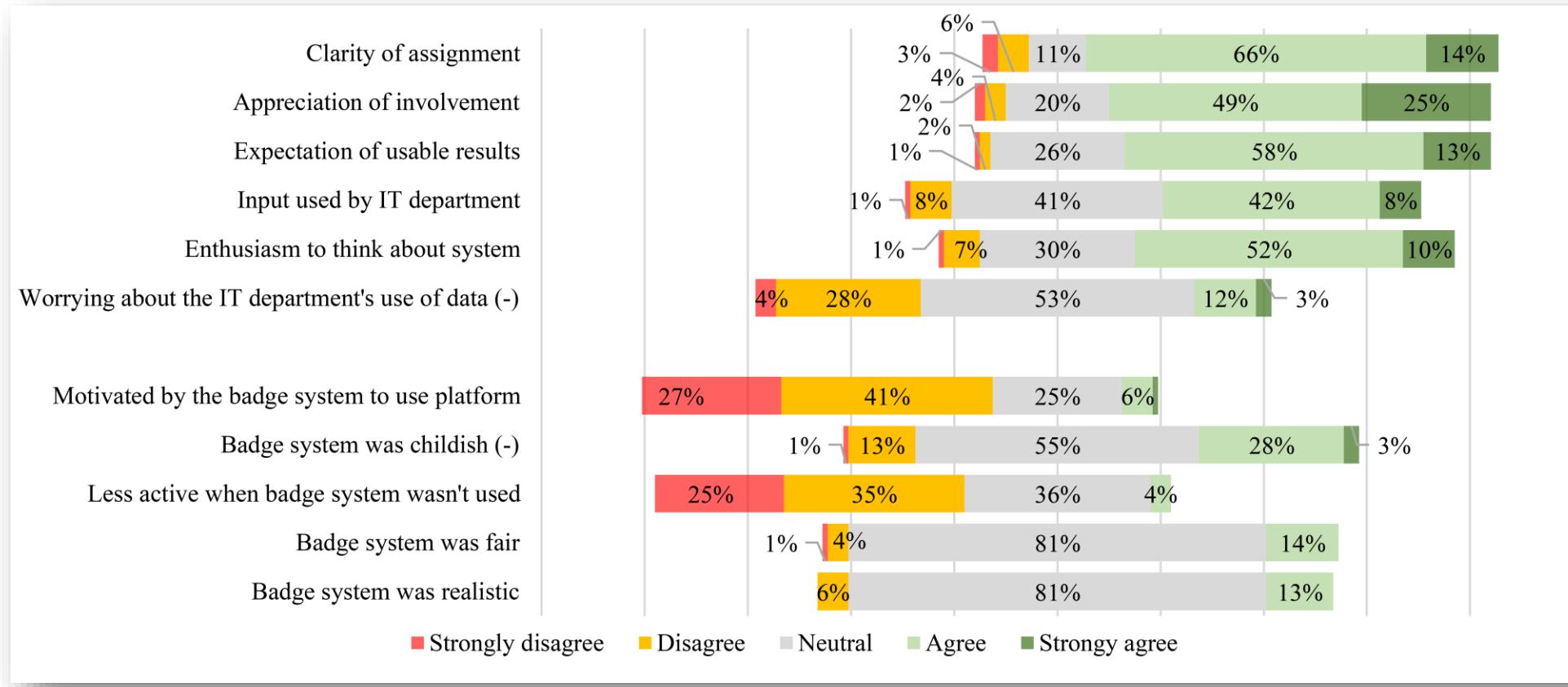
USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

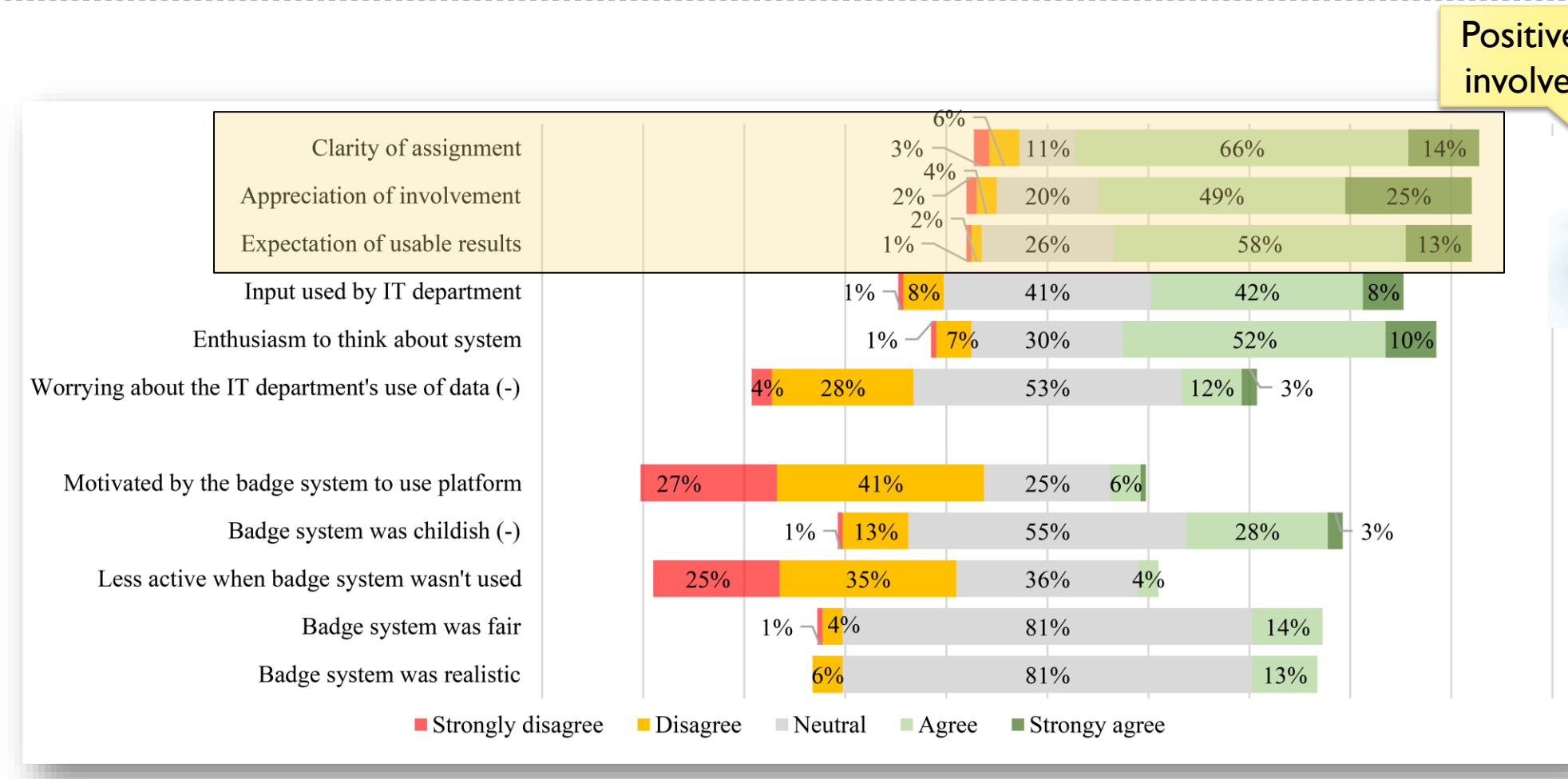
V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP	Enough for MVP	59.5
	Enough for product	27.8
Granularity	Epic	40.5
	User Story	54.4
	Not applicable	5.1

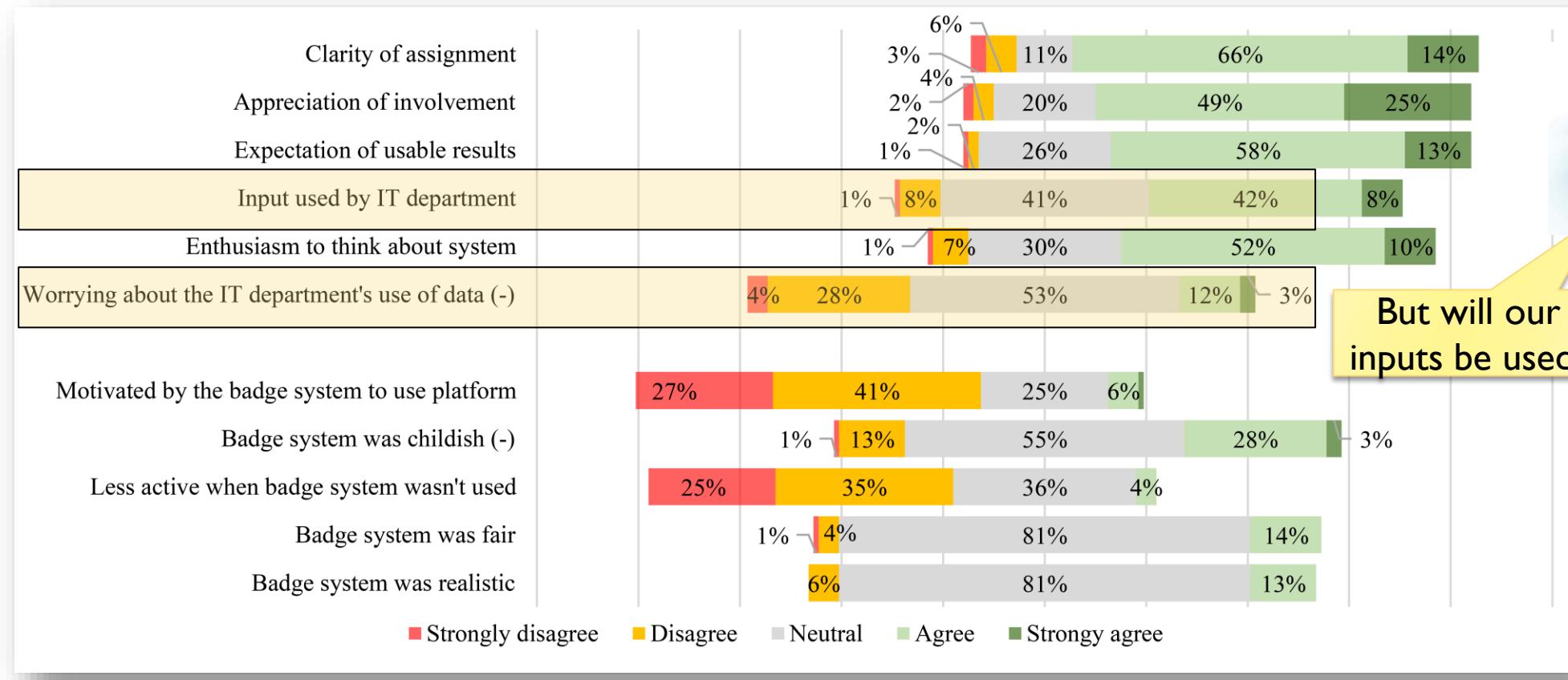
# KMar Crowd – Results – appreciation



# KMar Crowd – Results – appreciation

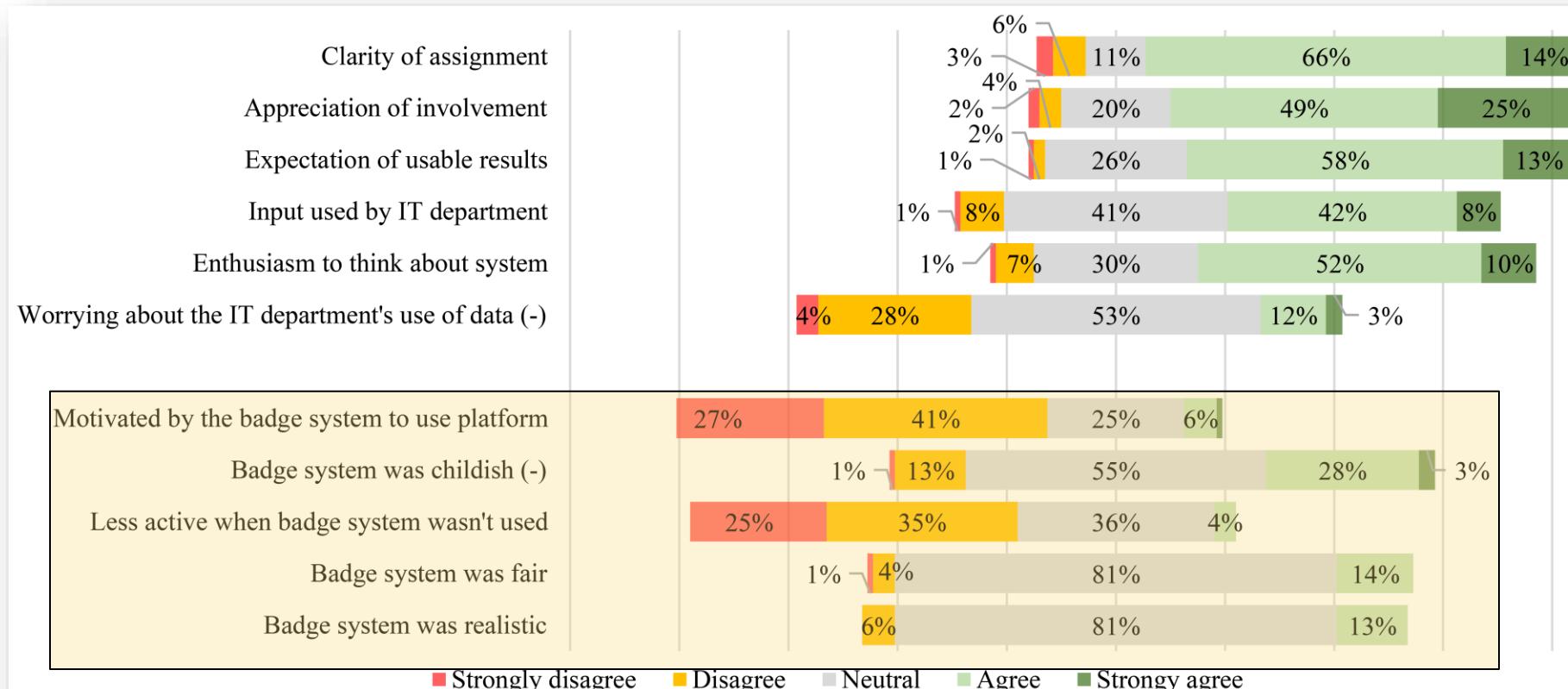


# KMar Crowd – Results – appreciation



But will our  
inputs be used?

# KMar Crowd – Results – appreciation



Gamification?  
Not for us!

## 4. Implications and Outlook

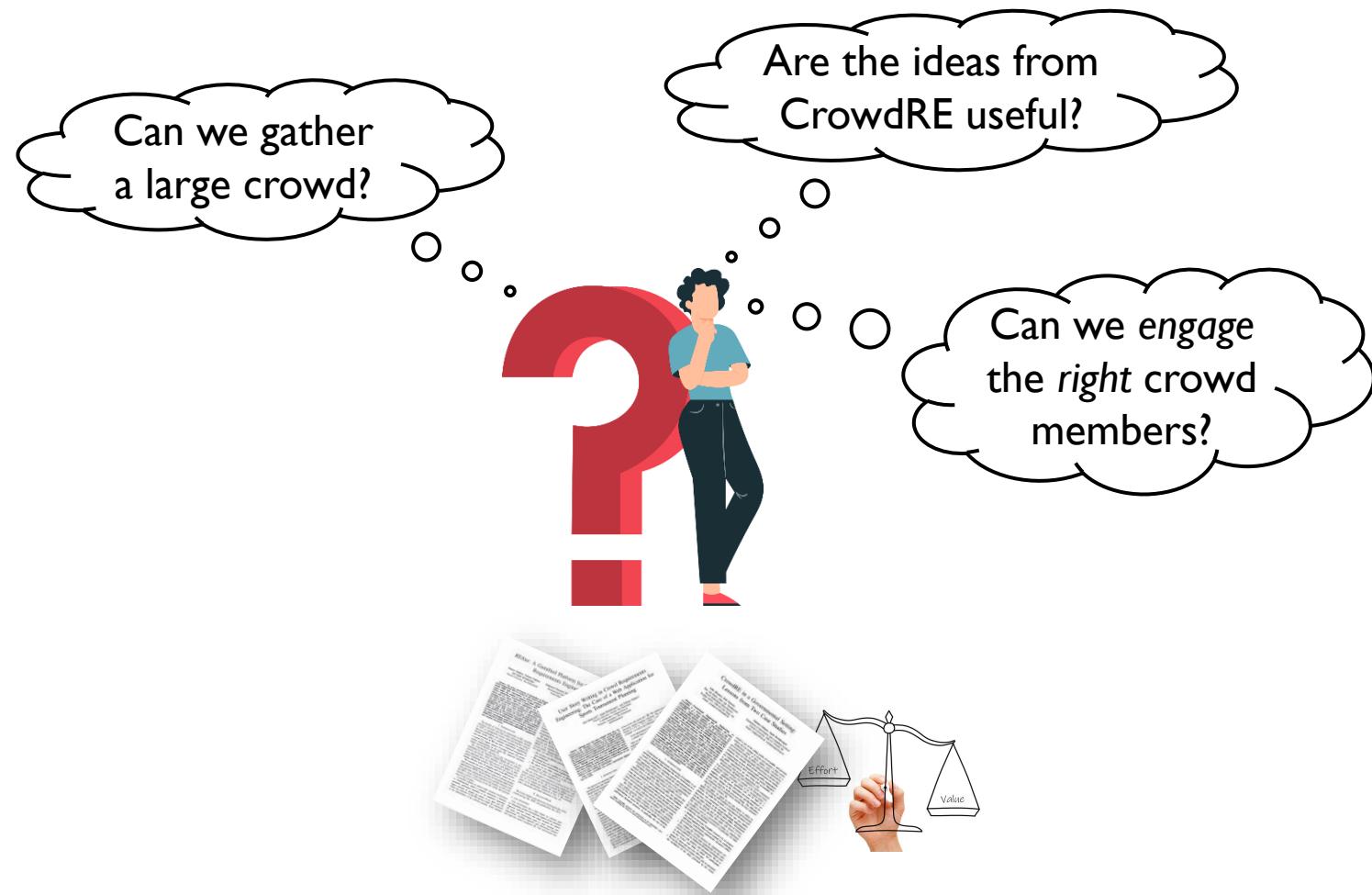


# Implications for Practice – **pull** feedback

---



# Implications for Practice – **pull** feedback



# Pull feedback – **crowd largeness**

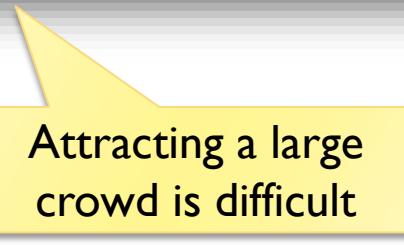
---

<b>Measurement</b>	<b>S-Sys</b>	<b>V-Sys</b>	<b>Menkveld</b>	<b>REFine</b>	<b>GARUSO</b>
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32

# Pull feedback – **crowd largeness**

---

<b>Measurement</b>	<b>S-Sys</b>	<b>V-Sys</b>	<b>Menkveld</b>	<b>REFine</b>	<b>GARUSO</b>
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32



Attracting a large crowd is difficult

# Pull feedback – **crowd largeness**

---

Do we even know how to  
measure a crowd's size ?

Measurement	S-Sys	V-Sys	Menkveld	REFine	GARUSO
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32

Attracting a large  
crowd is difficult

# Pull feedback – **crowd largeness**

Do we even know how to  
measure a crowd's size ?

Measurement	S-Sys	V-Sys	Menkveld	REFine	GARUSO
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32



↑people → ↑effort  
...but...  
↑people → ↑value ???

Attracting a large  
crowd is difficult

# Implications for Practice – **ideas' usefulness**

USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP		59.5
		27.8
		40.5
Granularity	Epic	54.4
	User Story	5.1
	Not applicable	

# Implications for Practice – ideas' usefulness

USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

Useful as **novel**:  
Difficult to validate, we need  
existing specifications

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP		59.5
		27.8
		40.5
Granularity	Epic	54.4
	User Story	5.1
	Not applicable	

# Implications for Practice – ideas' usefulness

USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP		59.5
		27.8
		40.5
Granularity	Epic	54.4
	User Story	5.1
	Not applicable	

Useful as **delighters**:  
some ideas are likely to be  
attractive requirements

# Implications for Practice – ideas' usefulness

USEFULNESS OF THE IDEAS IN THE S-SYS CASE STUDY, ASSESSED BY THE ANALYSTS WHO CONDUCTED ELICITATION WITHOUT *KMar-Crowd*.

Measurement	Value	# Ideas
KANO model	Must-be	13
	One-dimensional	10
	Attractive	7
Gathered earlier	Completely	19
	Partly	6
	Not at all	5
Complete for dev teams	Yes	11
	No	19

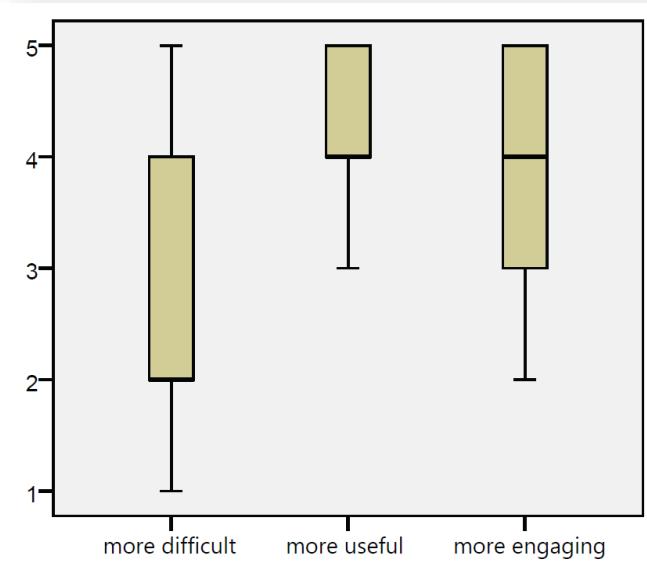
V-SYS: USEFULNESS OF THE IDEAS, ASSESSED BY A POOL OF ANALYSTS.

Measurement	Value	% Ideas
KANO model	Must-be	50.6
	One-dimensional	36.7
	Attractive	12.7
Enough for MVP	Enough for MVP	59.5
	Enough for product	27.8
	Epic	40.5
Granularity	User Story	54.4
	Not applicable	5.1

changes developing the right software. Furthermore, we have found that 95% of the crowdsourced USs are both easy to estimate and easily implementable based on our quality analysis and hour estimation as done by the main developer. Almost 90% of the feature requests can even be implemented within one workday.

Useful as **specific for dev teams**:  
many ideas will be, not all

# Implications for Practice – user engagement

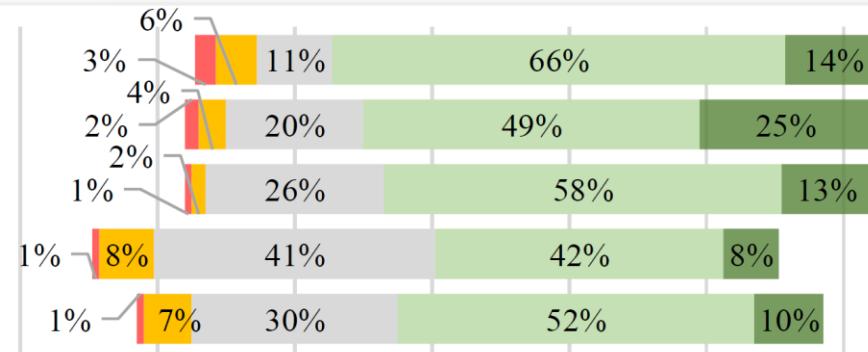


a feature themselves, while the other three respondents only voted for a feature. They perceived the platform as very useful, regarding all four possible interactions when rated on a five-point Likert scale: requesting ( $M = 4.9$ ;  $SD = 0.28$ ), viewing ( $M = 4.8$ ;  $SD = 0.38$ ), voting ( $M = 4.5$ ;  $SD = 0.88$ ), and commenting ( $M = 4.5$ ;  $SD = 0.66$ ). One user who requested a feature, voted for and commented on an idea and had previous experience in writing USs commented:

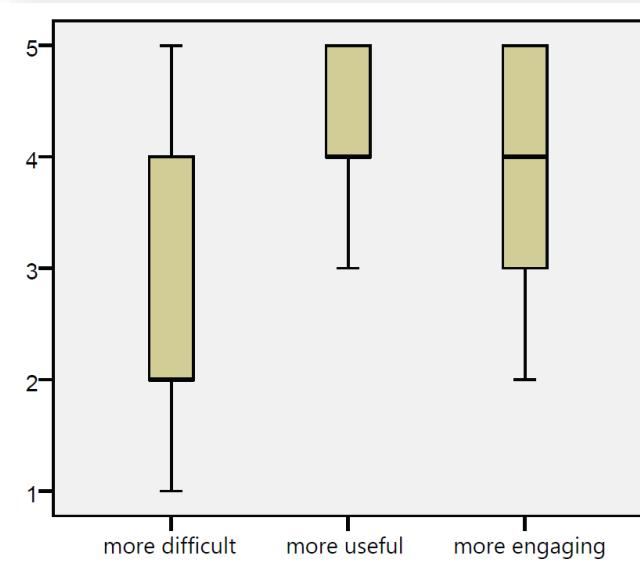
*"You implemented the agile methodology in a very fun way. In such a manner the users get involved better and at least have the feeling their opinion matters"*

Others found it "a fantastic way to improve the application" "very useful to allow users to submit ideas" and see it

- Clarity of assignment
- Appreciation of involvement
- Expectation of usable results
- Input used by IT department
- Enthusiasm to think about system



# Implications for Practice – user engagement



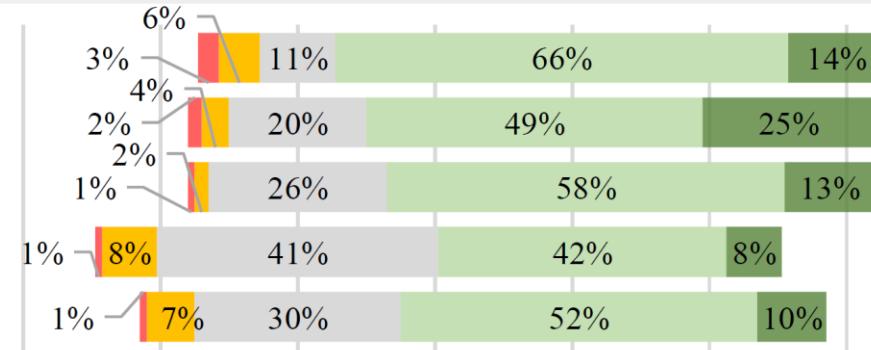
**Perceived engagement  
is generally high**

a feature themselves, while the other three respondents only voted for a feature. They perceived the platform as very useful, regarding all four possible interactions when rated on a five-point Likert scale: requesting ( $M = 4.9$ ;  $SD = 0.28$ ), viewing ( $M = 4.8$ ;  $SD = 0.38$ ), voting ( $M = 4.5$ ;  $SD = 0.88$ ), and commenting ( $M = 4.5$ ;  $SD = 0.66$ ). One user who requested a feature, voted for and commented on an idea and had previous experience in writing USs commented:

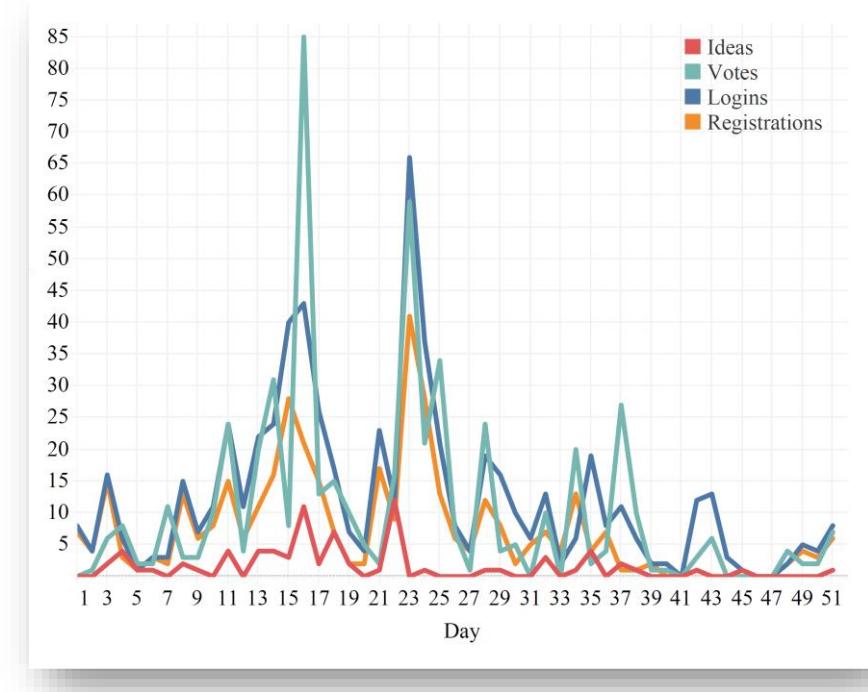
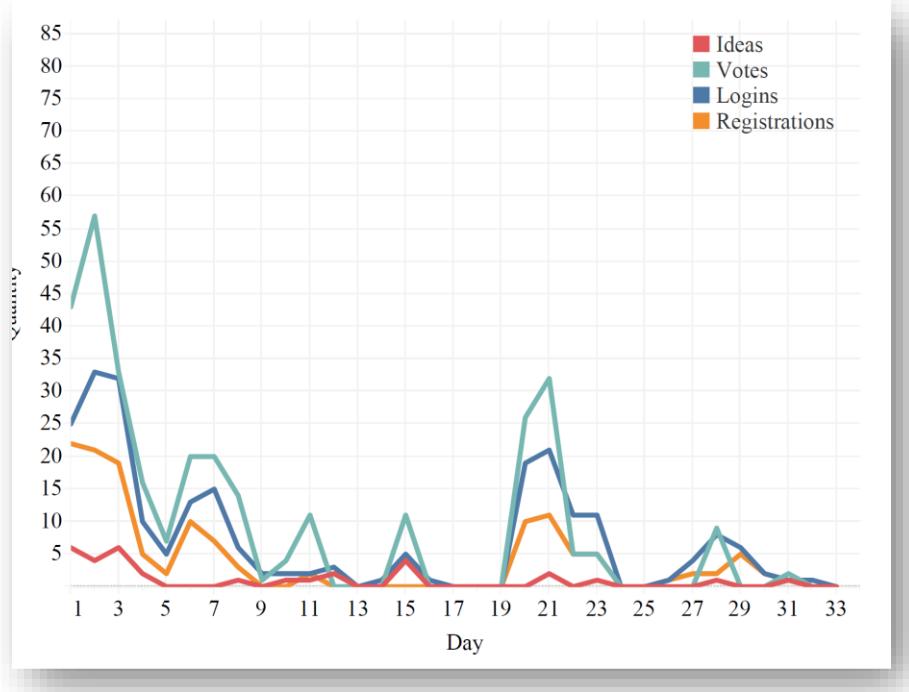
*"You implemented the agile methodology in a very fun way. In such a manner the users get involved better and at least have the feeling their opinion matters"*

Others found it "a fantastic way to improve the application" "very useful to allow users to submit ideas" and see it

- Clarity of assignment
- Appreciation of involvement
- Expectation of usable results
- Input used by IT department
- Enthusiasm to think about system

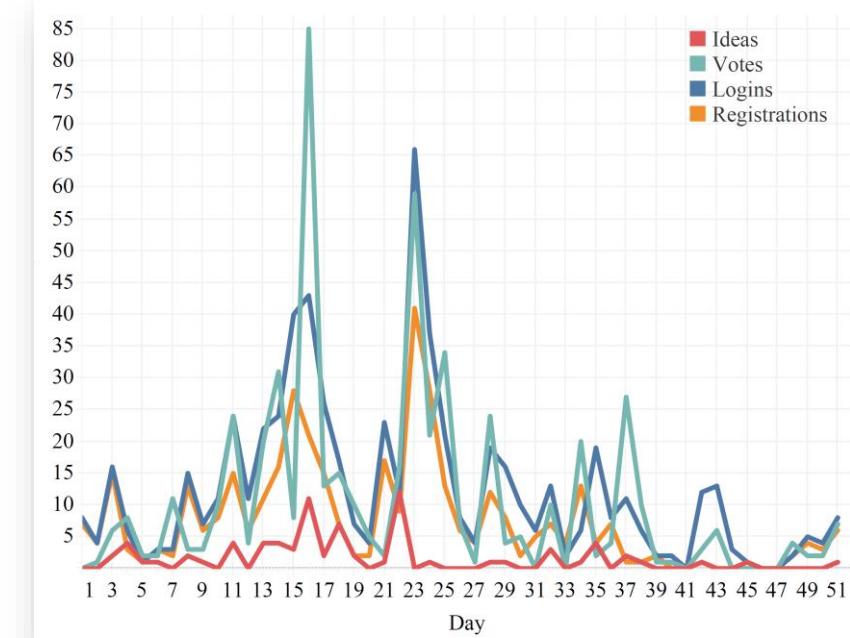
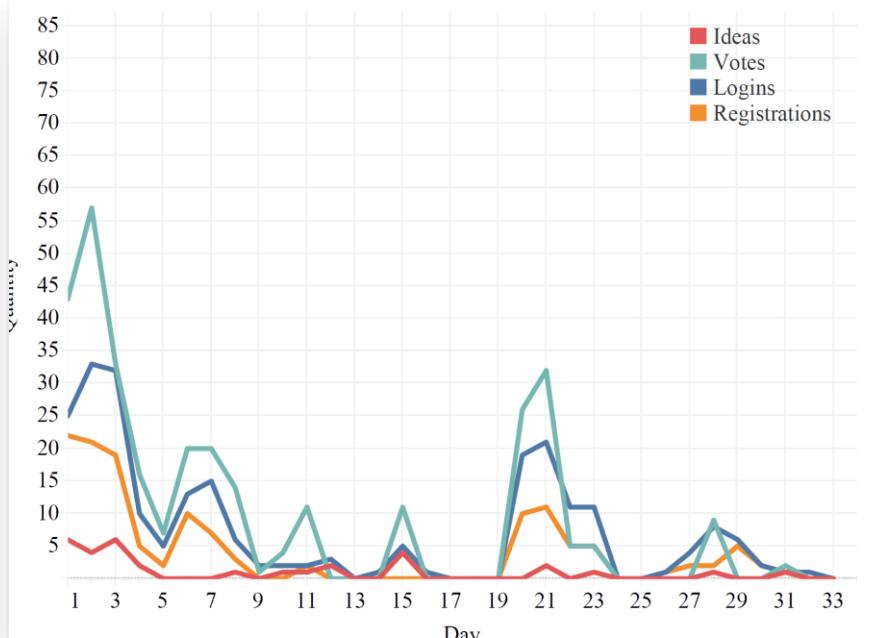


# Implications for Practice – **user engagement**



# Implications for Practice – **user engagement**

But engagement is not  
self-sustained



# Implications for Practice – **user engagement**

---

<b>Measurement</b>	<b>S-Sys</b>	<b>V-Sys</b>	<b>Menkveld</b>	<b>RRefine</b>	<b>GARUSO</b>
Duration in days	33	56	35	35	92
Participants:					
- Invited	478	2,393	unk.	37	unk.
- Accessed	135	385	157	19	726
- Active	60	130	39	19	32
Ideas	32	78	57	21	56
Logins	240	623	247	unk.	unk.
Votes	316	531	89	130	160
Comments	28	78	14	37	unk.
Ideas / Accessed	0.24	0.20	0.36	1.11	0.08



**And expect many more users than ideas!**

# Research Question – pull feedback answer

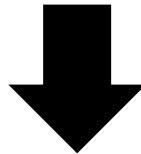
---

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



# Research Question – pull feedback answer

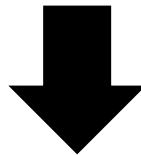
RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



- Extensive human effort is **unavoidable** with pull feedback

# Research Question – pull feedback answer

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



- Extensive human effort is **unavoidable** with pull feedback
- Rather, focus on **increasing value**
  - Make use of the ideas
  - Update the users on how the ideas are used
  - Reward the most active participants

# Research Question – push feedback answer

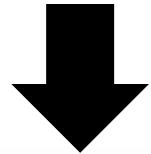
---

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



# Research Question – push feedback answer

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



**Automation = Value for Research**

High citation figures

How do users like this feature? a fine grained sentiment analysis of app reviews  
E Guzman, W Maalej  
2014 IEEE 22nd international requirements engineering conference (RE), 153-162

Bug report, feature request, or simply praise? on automatically classifying app reviews  
W Maalej, H Nabil  
2015 IEEE 23rd international requirements engineering conference (RE), 116-125

How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution  
S Panichella, AD Sorbo, E Guzman, A Visaggio, G Canfora, H Gall  
Software Maintenance and Evolution (ICSME), 2015 IEEE International ...

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 45, NO. 8, SEPTEMBER 2017

A Survey of App Store Analysis for Software Engineering  
William Martin, Federica Sacco, Yue Jia, Yunyanan Zhang, and Mark Harman

An Overview of User Feedback Classification Approaches  
Ruthens Santos, Eduard C. Gómez, Karina Villegas  
Fraunhofer ISE, Kaiserslautern, Germany

Extensive number of papers

Standard metrics to show success

$$R = \frac{|Ind \cap rel|}{|rel|} = \frac{|TP|}{|TP| + |FN|}$$
 (1)

$$P = \frac{|Ind \cap rel|}{|Ind|} = \frac{|TP|}{|TP| + |FP|}$$
 (2)

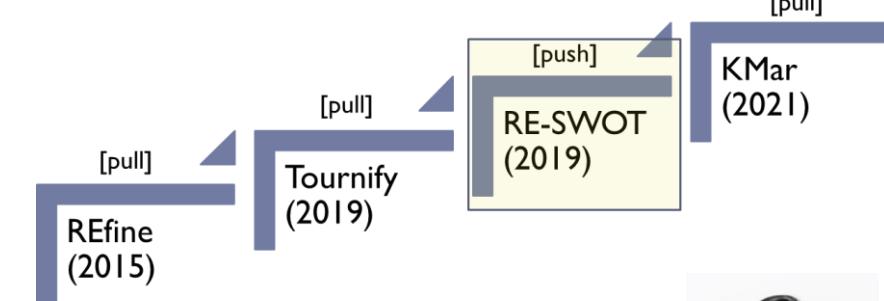
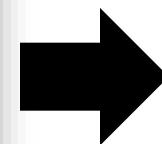
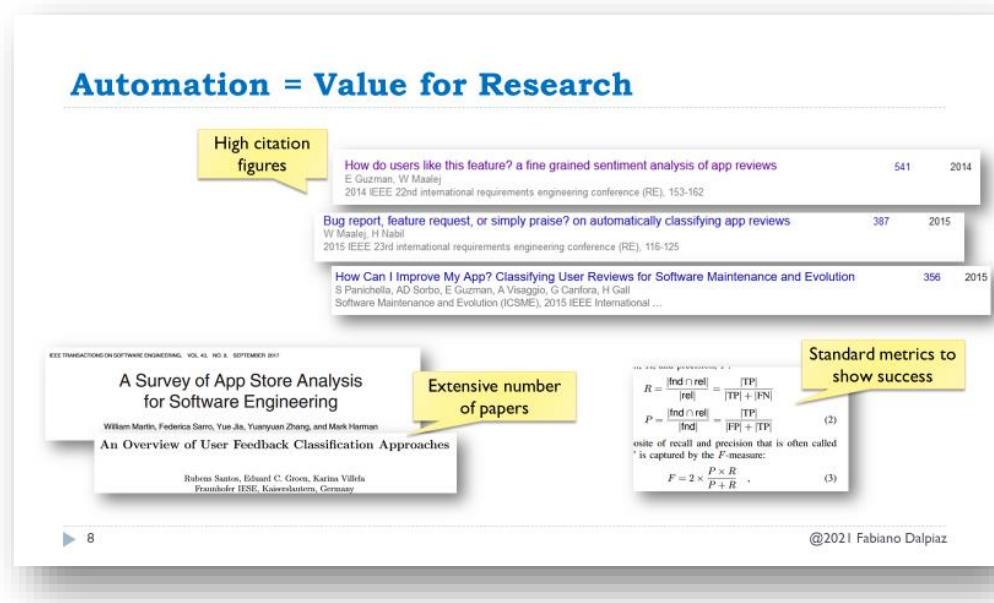
osite of recall and precision that is often called F-measure:

$$F = 2 \times \frac{P \times R}{P + R}$$
 (3)

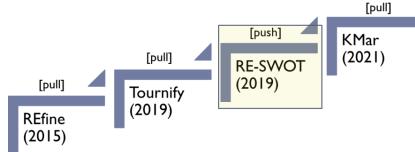
8 @2021 Fabiano Dalpiaz

# Research Question – push feedback answer

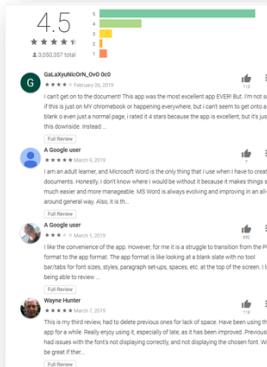
RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



Fabiano Dalpiaz, Micaela Parente: RE-SWOT:  
From User Feedback to Requirements via  
Competitor Analysis. REFSQ 2019: 55-70



# Which value in **push** feedback?



NLP + ML  
→  
gears

Topic	Senti. score
upload_photo, load_photo, photo_take, photo_want, move_photo, keep_upload, keep_try	1.51 Positive
file_name, folder_file, rename_file, file_add, chan	1.49 Positive
Sentence	Category
They just need to update the layout I fill like everything is hidden I want a better task bar.	feature request
Please restore a way to open pin links in external browser or let us save photos.	feature request
App crashes when new power up notice pops up.	problem discovery
Please fix the syncing issues with the iPad app.	problem discovery
It's already possible to rearrange boards why not the pins on a single board?	Information seeking
Overall it is fun and provides a lot of good info.	Information giving
This app runs so smoothly and I rarely have issues with it anymore.	Information giving



... ,  $R$ , and  $P$  precision,  $F$ :

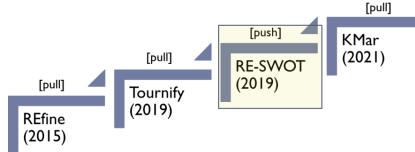
$$R = \frac{|fnd \cap rel|}{|rel|} = \frac{|TP|}{|TP| + |FN|}$$

$$P = \frac{|fnd \cap rel|}{|fnd|} = \frac{|TP|}{|FP| + |TP|} \quad (2)$$

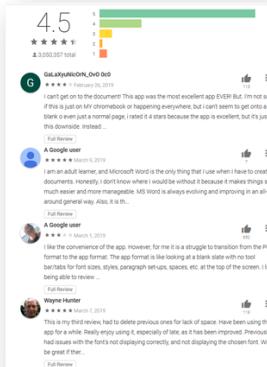
osite of recall and precision that is often called ' is captured by the  $F$ -measure:

$$F = 2 \times \frac{P \times R}{P + R}, \quad (3)$$

Standard metrics to show success



# Which value in **push** feedback?



NLP + ML  
→  
Gears

Topic	Senti. score
upload_photo, load_photo, photo_take, photo_want, move_photo, keep_upload, keep_try	1.51 Positive
file_name, folder_file, rename_file, file_add, chan	1.49 Positive
Sentence	Category
They just need to update the layout I fill like everything is hidden I want a better task bar.	feature request
Please restore a way to open pin links in external browser or let us save photos.	feature request
App crashes when new power up notice pops up.	problem discovery
Please fix the syncing issues with the iPad app.	problem discovery
It's already possible to rearrange boards why not the pins on a single board?	Information seeking
Overall it is fun and provides a lot of good info.	Information giving
This app runs so smoothly and I rarely have issues with it anymore.	Information giving

... ,  $R$ , and  $P$  are:

$$R = \frac{|fnd \cap rel|}{|rel|} = \frac{|TP|}{|TP| + |FN|}$$

$$P = \frac{|fnd \cap rel|}{|fnd|} = \frac{|TP|}{|FP| + |TP|} \quad (2)$$

Composite of recall and precision that is often called 'F' is captured by the  $F$ -measure:

$$F = 2 \times \frac{P \times R}{P + R}, \quad (3)$$

Standard metrics to show success



# Pillars of RE-SWOT

---



- ▶ Practitioner-focused setting
  - ▶ A **product manager** wants to analyze their **app's reviews**
  - ▶ Idea: analyze with respect to **competitors**
  - ▶ Through an **automatically created dashboard**

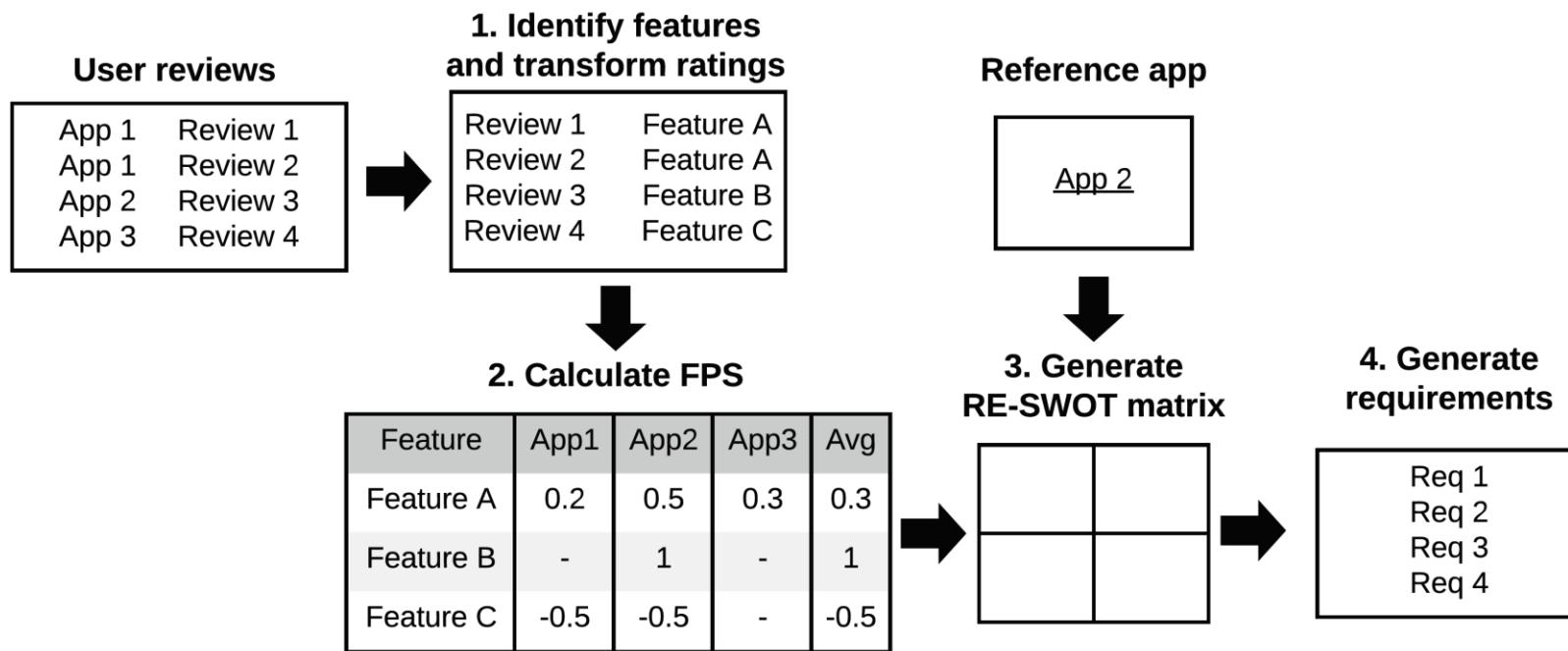
# Pillars of RE-SWOT

---

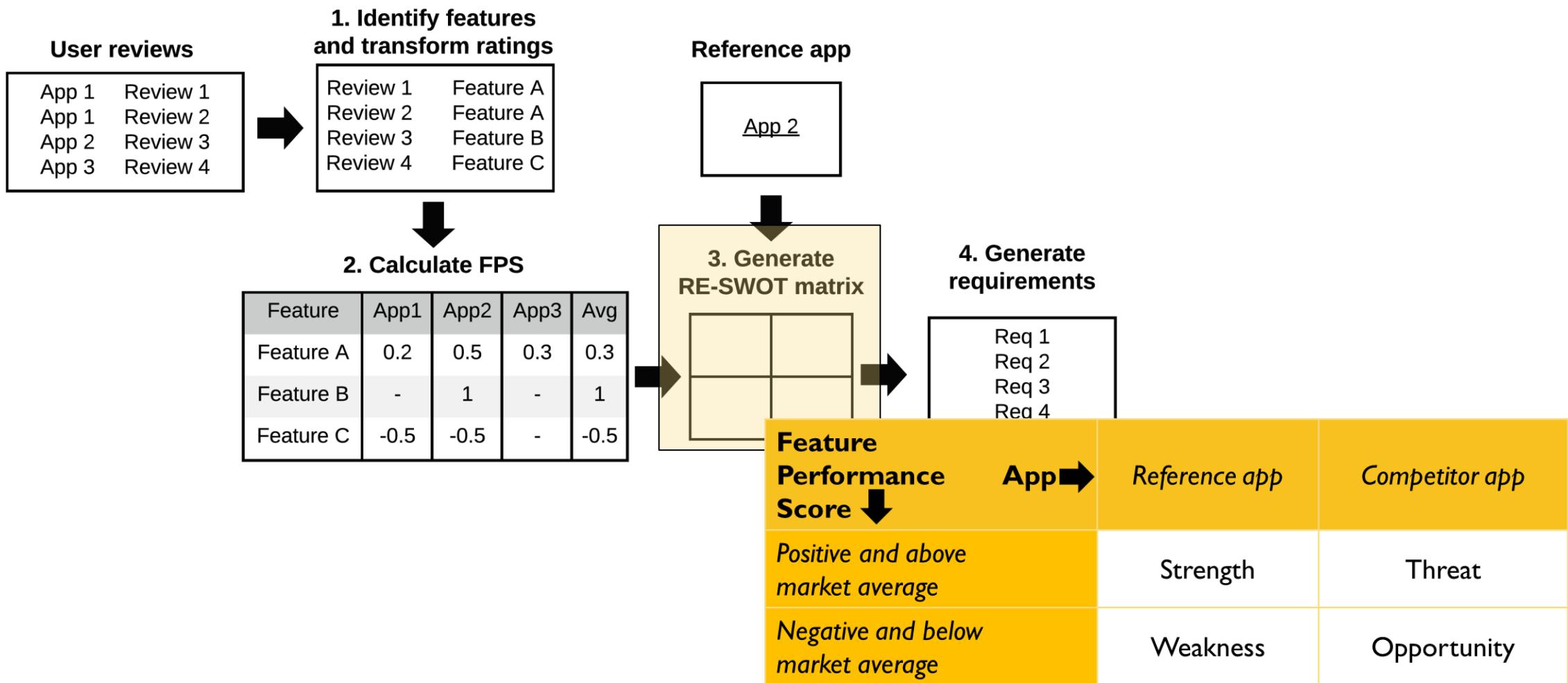


- ▶ Practitioner-focused setting
  - ▶ A **product manager** wants to analyze their **app's reviews**
  - ▶ Idea: analyze with respect to **competitors**
  - ▶ Through an **automatically created dashboard**
- ▶ In-vivo validation
  - ▶ **Ignore precision, recall, F1**
  - ▶ Consider product manager's inputs

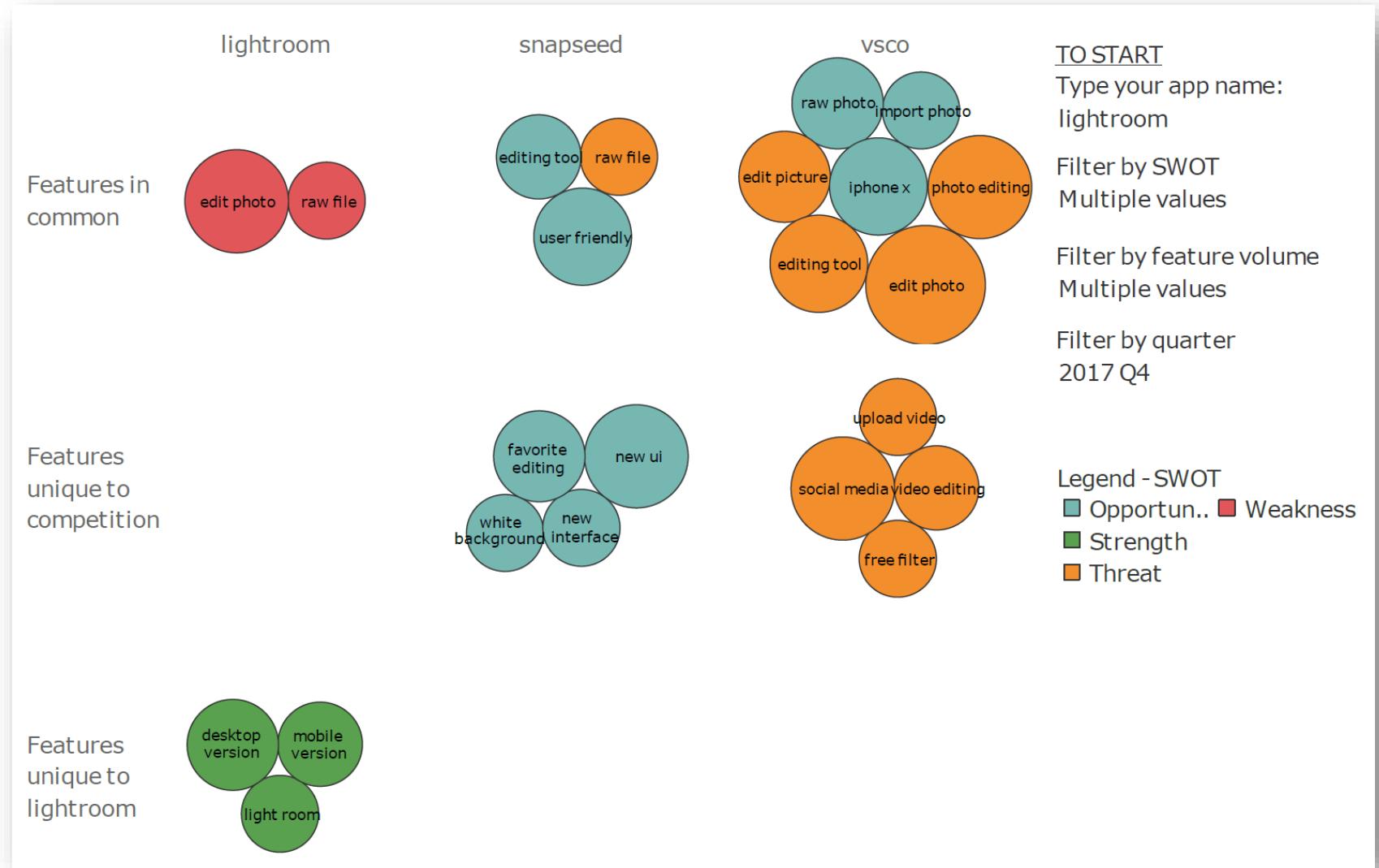
# RE-SWOT illustrated



# RE-SWOT illustrated



# The RE-SWOT Matrix, in practice



# Three case studies

---

## ▶ **Dating apps**

- ▶ Canadian company with 150M users
- ▶ Competitors: market leader, fast growing company

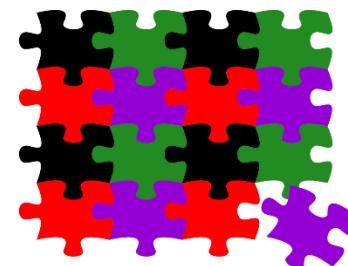


## ▶ **Travel apps for tourists**

- ▶ 400 employees company
- ▶ Competitors: app for Asia, market leader

## ▶ **Puzzle games**

- ▶ App with 135 puzzle types, small company
- ▶ Competitors: similar apps



# Three case studies

---

## ▶ **Dating apps**

- ▶ Canadian company with 150M users
- ▶ Competitors: market leader, fast growing company



### **Interviews** with product managers

1. Introduction
2. Your current practices
3. Demo
4. Tool use
5. Questions

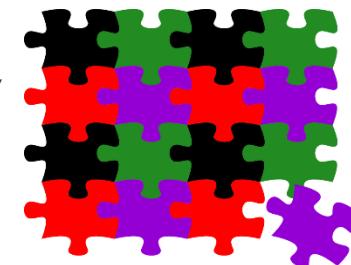
sts

/

sia, market leader

## ▶ **Puzzle games**

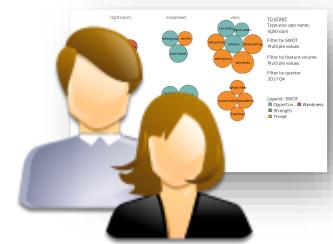
- ▶ App with 135 puzzle types, small company
- ▶ Competitors: similar apps



# Main results from RE-SWOT

---

- + Visual/interactive aspect
- + Allows to handle many reviews
- + Some competitors' features could be found



Product managers



# Main results from RE-SWOT

- + Visual/interactive aspect
- + Allows to handle many reviews
- + Some competitors' features could be found



Product managers

- Feature extraction is inaccurate
- SWOT classification too
- We need trend analysis



# Research Question – push feedback answer

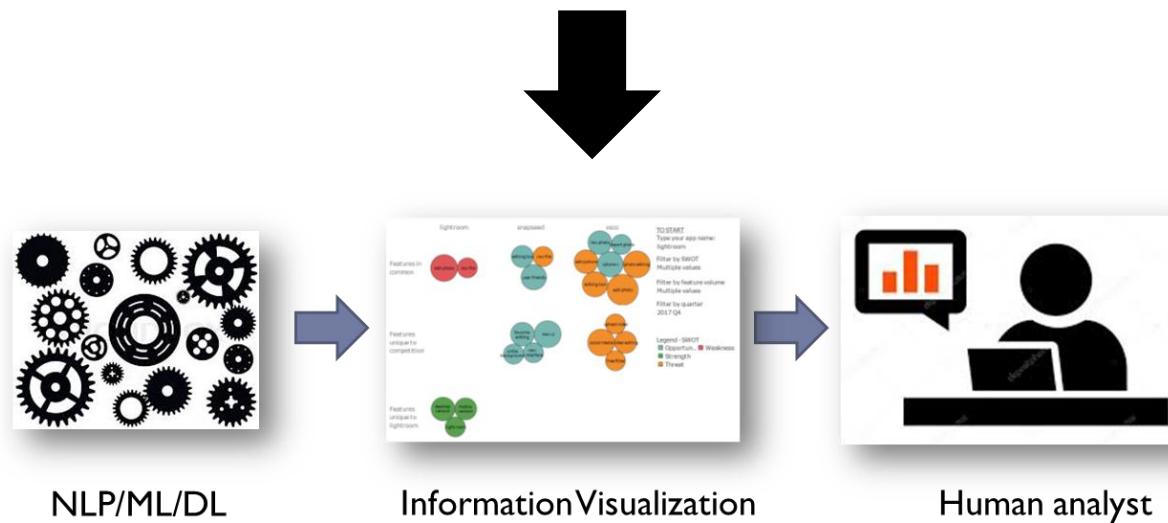
---

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



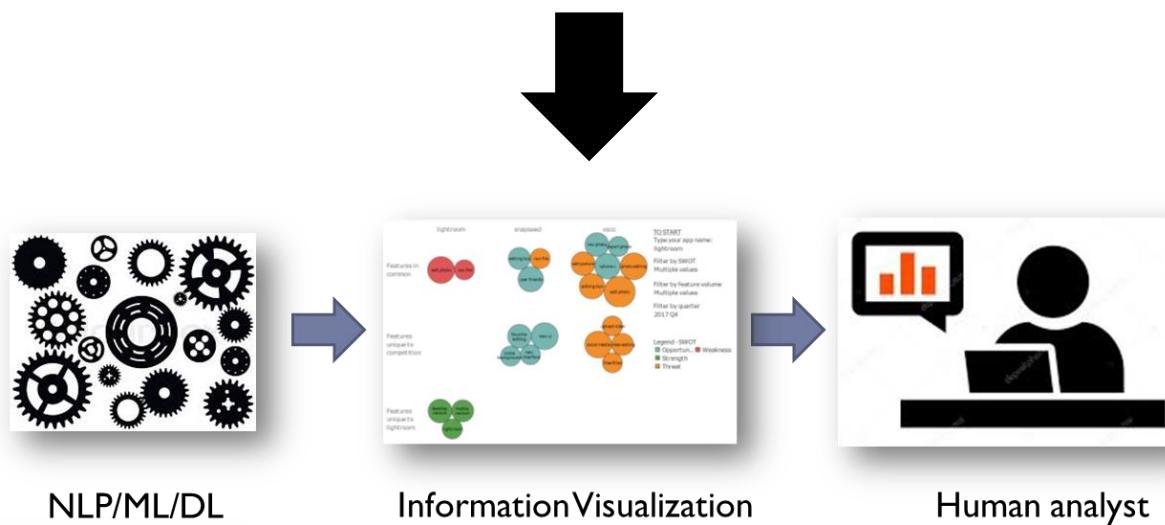
# Research Question – push feedback answer

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



# Research Question – push feedback answer

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



Requirements Eng (2014) 19:257–279  
DOI 10.1007/s00766-013-0194-3

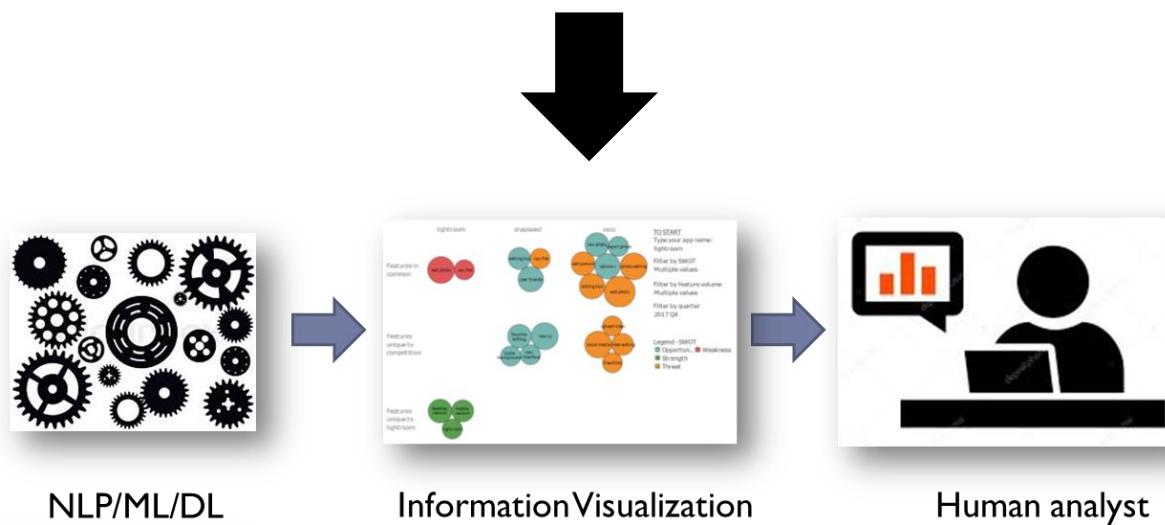
RE 2013

Visual requirements analytics: a framework and case study

Sandeep Reddivari · Shirin Rad · Tanmay Bhowmik ·  
Nisreen Cain · Nan Niu

# Research Question – push feedback answer

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



Requirements Eng (2014) 19:257–279  
DOI 10.1007/s00766-013-0194-3

RE 2013

Visual requirements analytics: a framework and case study

Sandeep Reddivari · Shirin Rad · Tanmay Bhowmik ·  
Nisreen Cain · Nan Niu

- Don't focus solely on ready-to-use metrics
- Yet, the metrics are useful to improve the algorithms
- Involve practitioners!

# Conclusion

---

- ▶ It is time for CrowdRE to assess its **value in vivo**
  - ▶ Collaboration with industry is crucial
  - ▶ Don't focus only on academic metrics of success



# Conclusion

---

- ▶ It is time for CrowdRE to assess its **value in vivo**
  - ▶ Collaboration with industry is crucial
  - ▶ Don't focus only on academic metrics of success
- ▶ **Human effort is unavoidable**
  - ▶ AI and visual analytics can help provide a summary
  - ▶ But you can't replace the human



# A new research question for the CrowdRE community?

---

RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



# A new research question for the CrowdRE community?

---

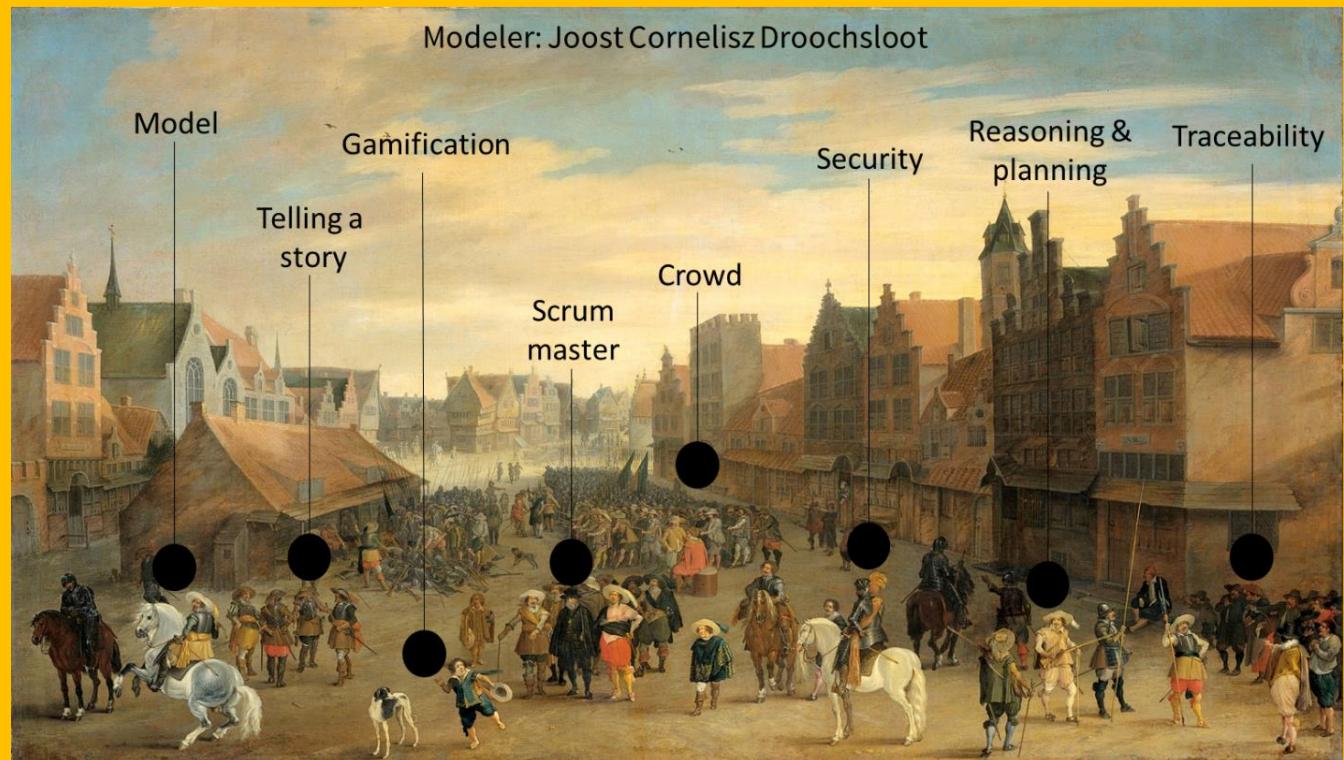
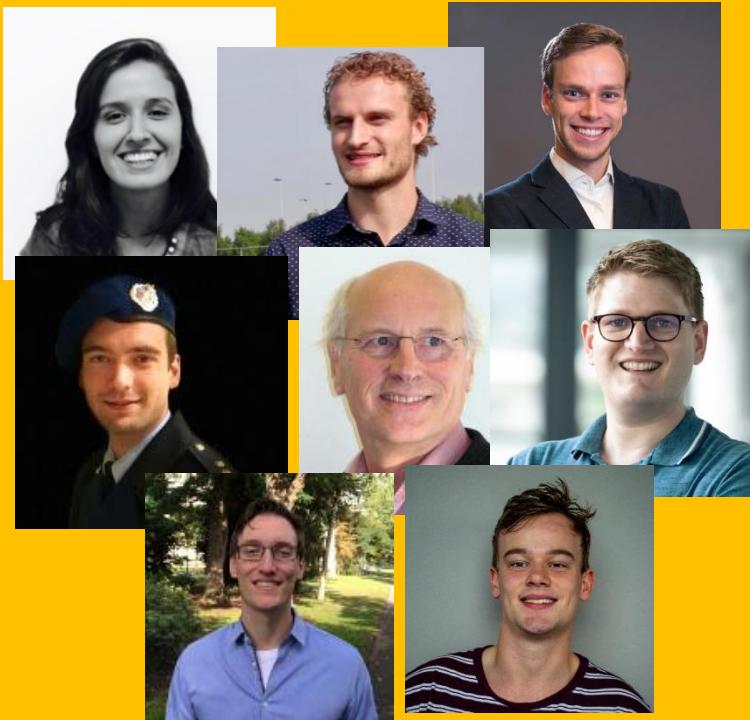
RQ: How to **reduce** the required **human effort** for managing user feedback so to **increase** CrowdRE's **value for practice**?



**RQ': How to assess the delivered value of CrowdRE so to increase CrowdRE's adoption in practice?**

# Thanks from the Requirements Engineering Lab at Utrecht University!

## Our CrowdRE crowd ☺



f.dalpiaz@uu.nl



@FabianoDalpiaz

