LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

TIB

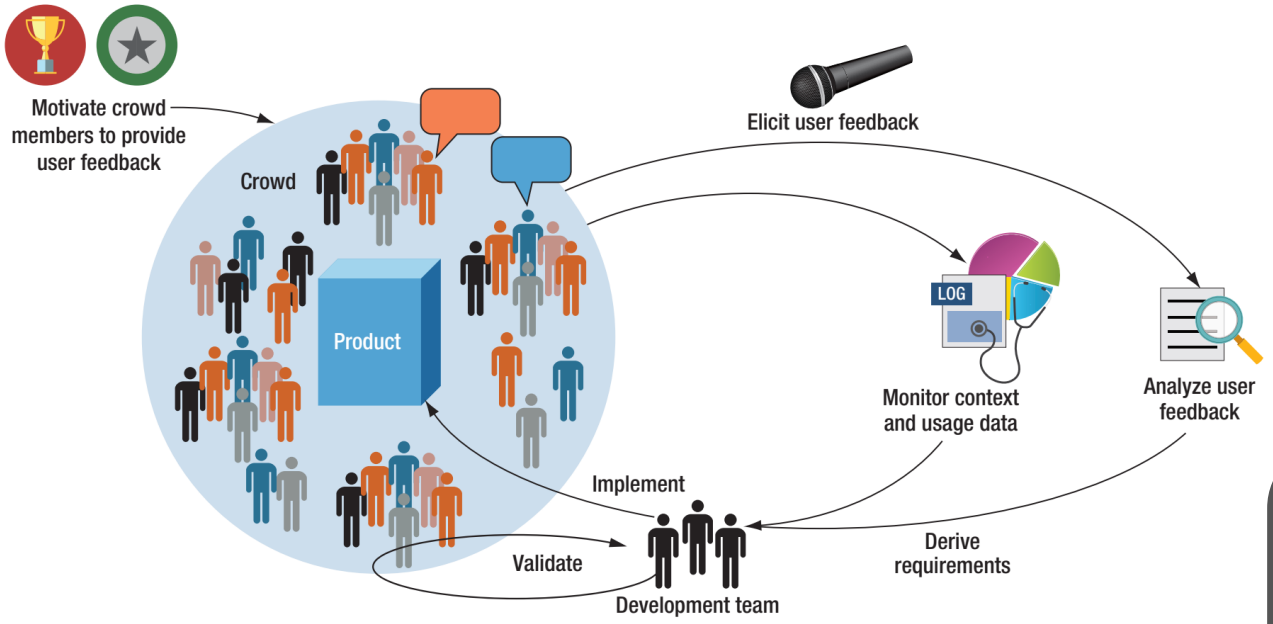
The Potential of Using Vision Videos for CrowdRE: Video Comments as a Source of Feedback

Oliver Karras, Eklekta Kristo, and Jil Klünder

Dr. rer. nat. Oliver Karras,
Notre Dame, South Bend, USA 13. September 2021
5th International Workshop on Crowd-Based Requirements Engineering

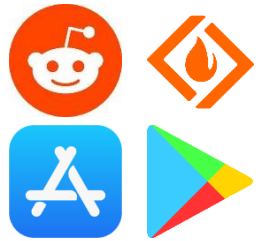
Leibniz Association

CrowdRE[1] and its User Feedback Sources



Frequently used sources[2]

- User forums
- Mobile application markets



Rarely used source[2]

- Social media platforms

Potential of social media platforms

1. **New motivation opportunities**: Entertaining & enjoyable activities such as **watching a video**[2]
2. **Wide reach**: Gather millions of views and solicit **thousands of comments**[3]

[1] E. Groen et al.: *The Crowd in Requirements Engineering – The Landscape and Challenges*, IEEE Software, 2017.
[2] J. Khan et al.: *Crowd Intelligence in Requirements Engineering: Current Status and Future Directions*, REFSQ, 2019.
[3] Vistisen and Poulsen: *Return of the Vision Video: Can Corporate Vision Videos Serve as Setting for Participation?*, Nordes 7(1), 2017.

Vision Video – One kind of Videos in RE

Apple Knowledge Navigator, 1987

- Fictitious tablet with voice assistant

A vision video is…

…a video that represents a vision or parts of it (problem, solution, improvement) for achieving shared understanding among all parties involved by disclosing, discussing, and aligning their mental models of the future system.[1]

Frequent use of vision videos

- Individual meetings
- Focus groups
- Workshops

} Stimulate discussion & Solicit feedback

Proposal[2]

Use of **vision videos for CrowdRE** by transferring their benefits to **social media platforms**

[1] O. Karras: *Supporting Requirements Communication for Shared Understanding by Applying Vision Videos in Requirements Engineering*. Logos Verlag, 2020.
[2] K. Schneider et al.: *Reframing Societal Discourse as Requirements Negotiation: Vision Statement*, 2nd CrowdRE Workshop, 2017.

Research Question & Analysis Procedure

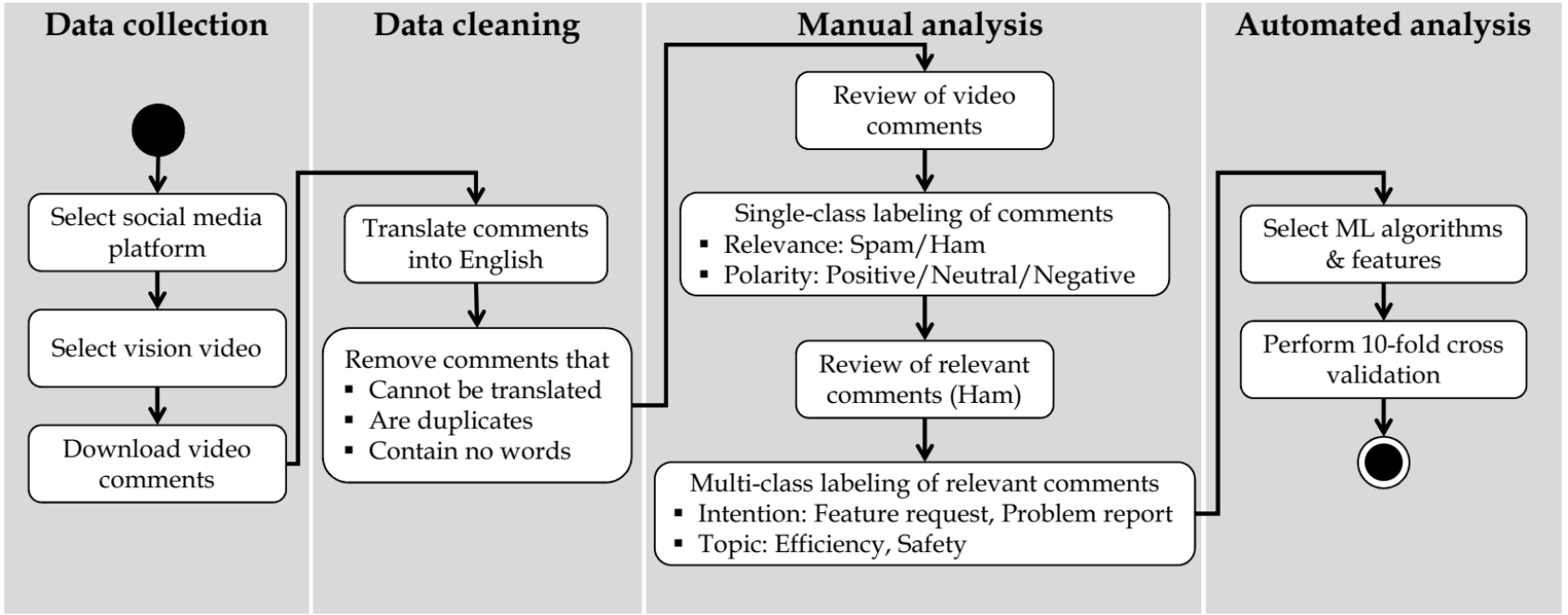
Research question

What is the **potential** of using **vision videos** on social media platforms to **solicit feedback** in the form of **video comments** for CrowdRE?

Quantity
Views & Comments

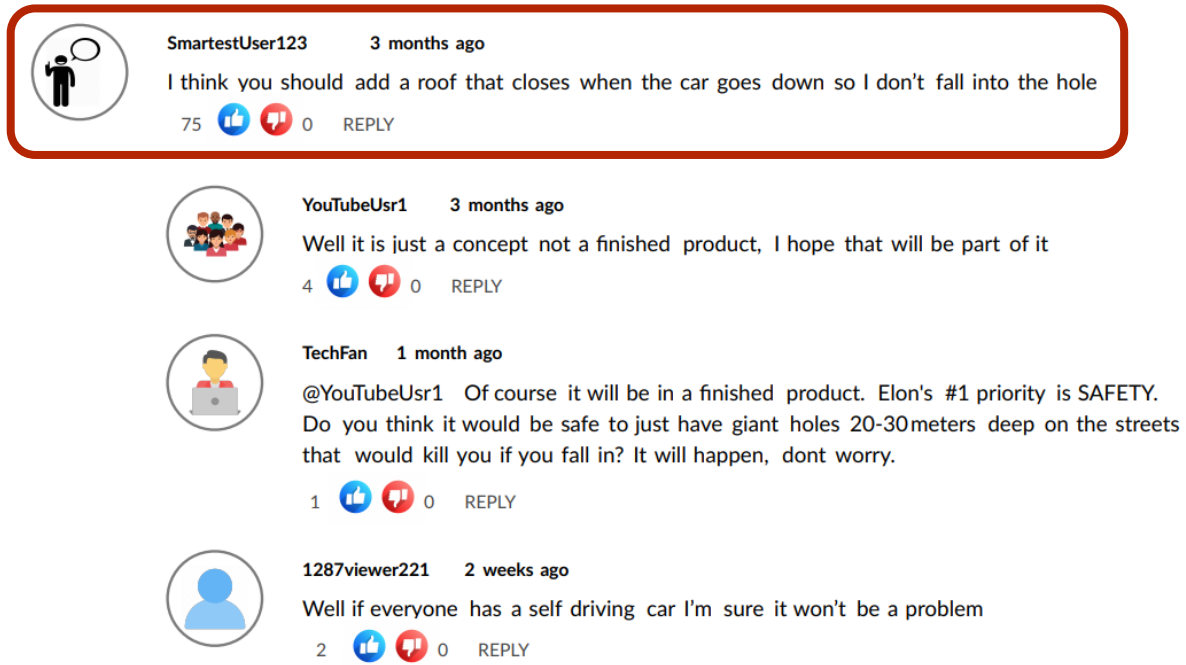
Content
of comments

Suitability
for analyses



Platform, Vision Video, and Data Set

Tunnels by The Boring Company, 28th Apr 2017

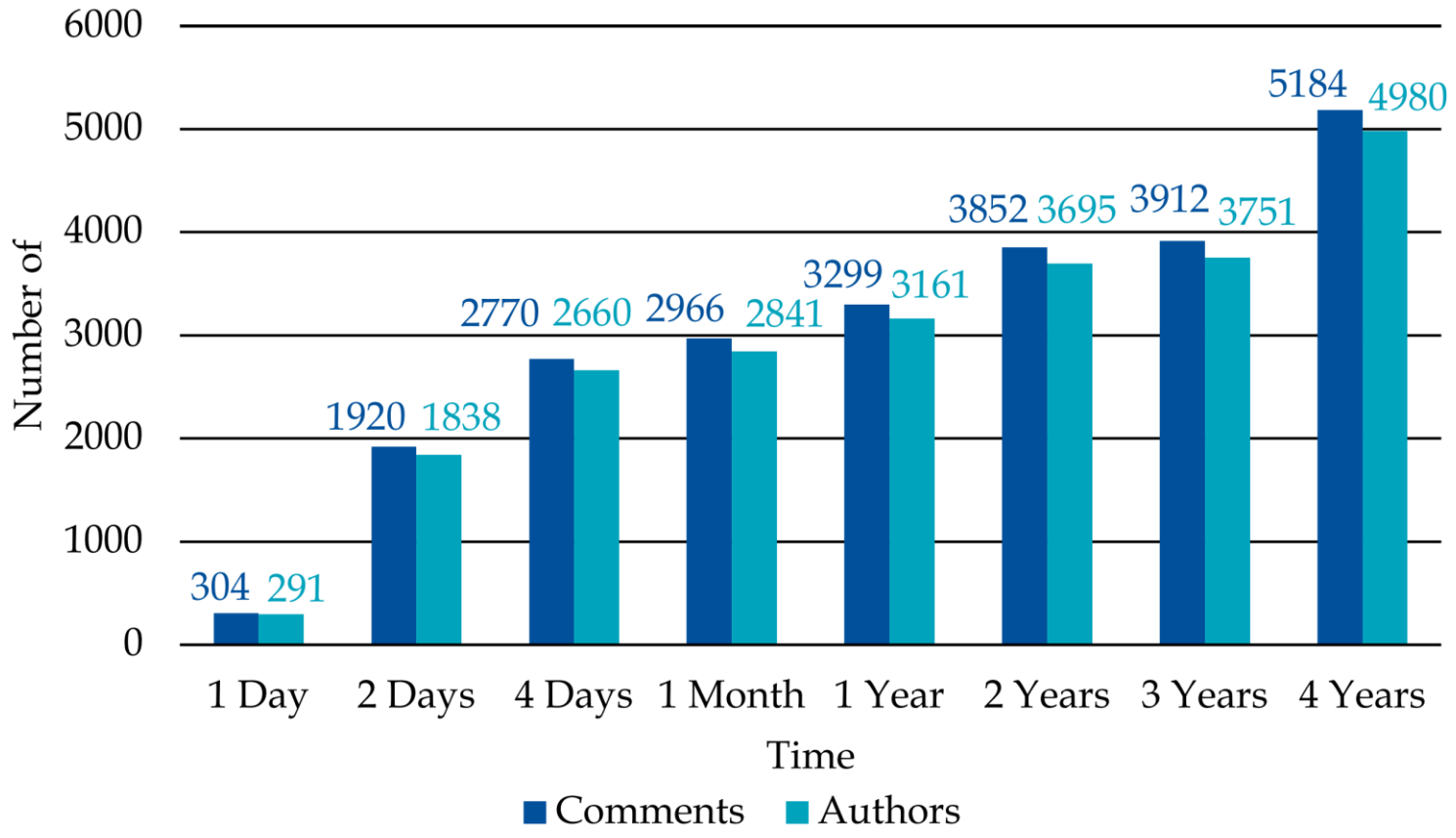


Date	Views	Likes	Dislikes	Comments	Replies	Total
Download (Oct 13th 2020)	6.9m	59621	4287	4505	2107	6612
Publication (Apr 28th 2021)	8.3m	88457	4913	5184	2652	7836

After data cleaning
4400 comments

Replication package: https://zenodo.org/record/4698969

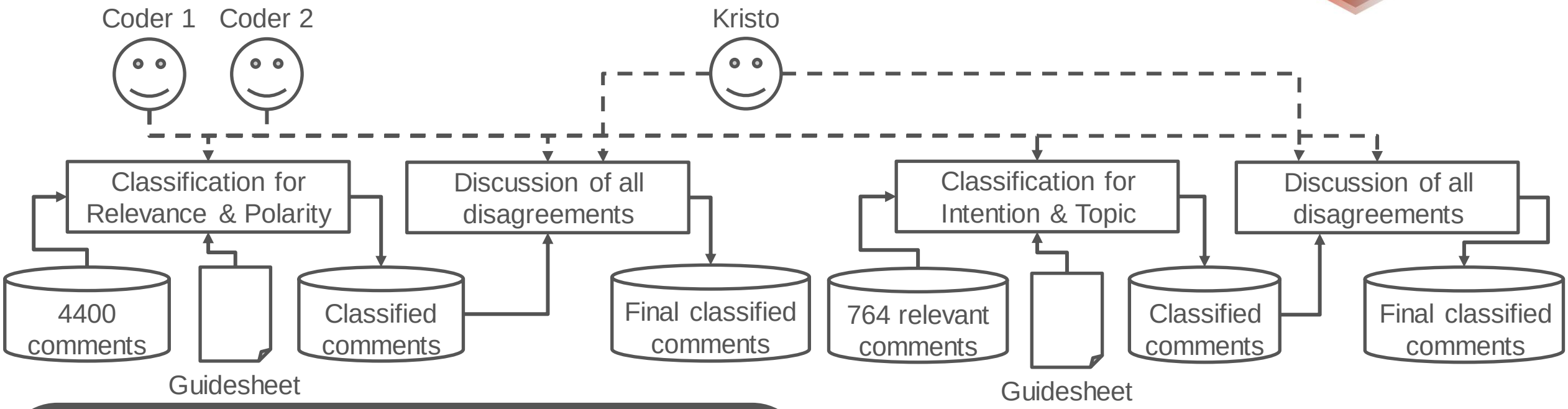
Quantity



Finding 1:

- After **4 days** there were **2770 comments** (approx. **53% of all comments in 4 years**)

Content



Finding 2:

- **Only 764** comments are **relevant** (approx. **17%**)
- **Feature request** and **problem report** most frequent **intentions**
- **Efficiency** and **Safety** are frequent **topics** in the comments
- **Content** of comments is **related to taxonomy** of user feedback classification categories[1]

Category	Positive	Neutral	Negative	Overall
Overall	400	3033	967	4400
Spam	368	2486	782	3636
Ham	32	547	185	764
Feature request	7	115	20	142
Problem report	16	277	116	409
Efficiency	5	101	54	160
Safety	16	205	91	312

[1] Santos et al.: *A Taxonomy for User Feedback Classification*. 2nd NLP4RE Workshop, 2019.

Suitability

Category	Metric	Naive Bayes		SVM		Random forest	
		BOW	TF-IDF	BOW	TF-IDF	BOW	TF-IDF
Relevance: Ham	Accuracy	0.663	0.695	0.799	**0.818**	0.808	0.807
	Precision	0.675	0.697	0.799	**0.819**	0.809	0.808
	Recall	0.663	0.695	0.798	**0.818**	0.808	0.806
	F_1	0.657	0.694	0.798	**0.817**	0.807	0.806
Intention: Feature request	Accuracy	0.672	0.644	0.712	**0.750**	0.745	0.734
	Precision	0.680	0.656	0.726	**0.757**	0.753	0.742
	Recall	0.676	0.647	0.719	**0.752**	0.750	0.738
	F_1	0.666	0.634	0.708	**0.745**	0.742	0.731
Intention: Problem report	Accuracy	0.566	0.600	0.644	**0.675**	0.641	0.650
	Precision	0.570	0.600	0.651	**0.681**	0.644	0.656
	Recall	0.566	0.599	0.650	**0.680**	0.644	0.654
	F_1	0.556	0.594	0.604	**0.673**	0.640	0.648
Topic: Efficiency	Accuracy	0.642	0.632	0.723	0.737	**0.747**	0.724
	Precision	0.648	0.650	0.728	0.740	**0.751**	0.726
	Recall	0.643	0.637	0.726	0.740	**0.746**	0.725
	F_1	0.636	0.623	0.722	0.734	**0.742**	0.720
Topic: Safety	Accuracy	0.652	0.663	0.814	**0.823**	0.816	0.803
	Precision	0.670	0.667	0.822	**0.827**	0.822	0.806
	Recall	0.651	0.664	0.815	**0.824**	0.816	0.805
	F_1	0.638	0.658	0.811	**0.821**	0.813	0.801

Algorithms & features[1]

- SVM and naïve bayes most frequently used
- Random forest rarely used, but good performance
- BOW and TF-IDF most frequently used

Procedure

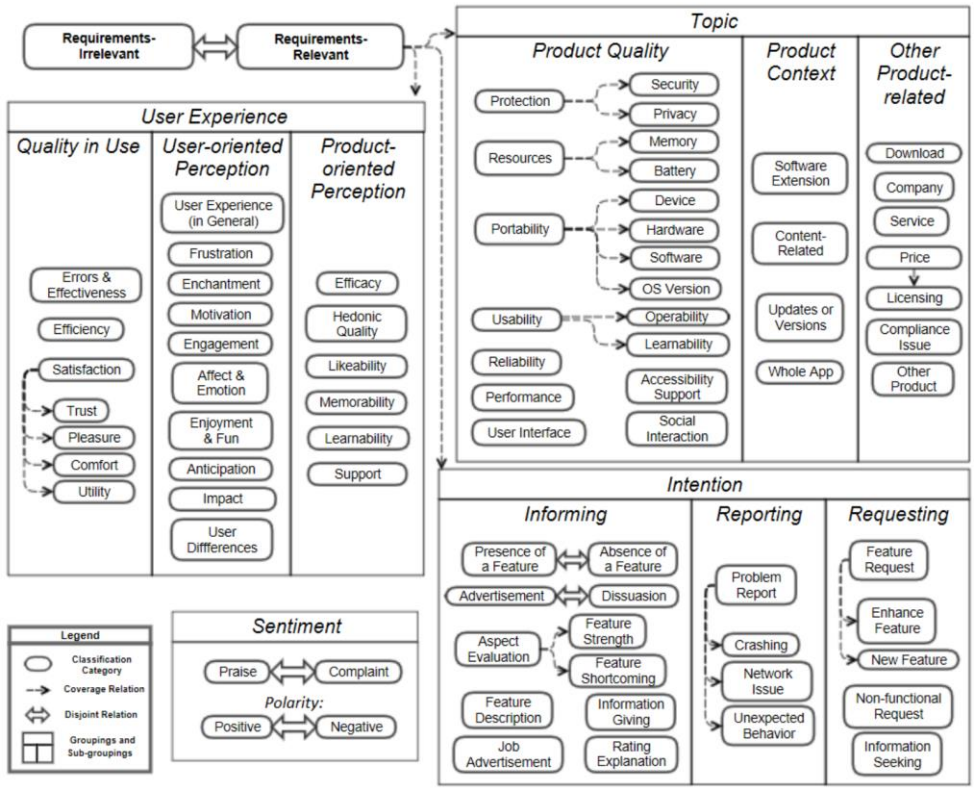
- Balanced data sets for each category
- 10-fold cross validation for binary classification
- 10 times repeated to reduce data splitting bias

Finding 3:

- **SVM with TF-IDF** achieved **best** results overall
- **Ham** and **safety** can be classified **well**
- **Problem report** is most **difficult** to classify

[1] Santos et al.: *An Overview of User Feedback Classification Approaches*. 2nd NLP4RE Workshop, 2019.

Discussion – Feedback on Safety



From 43 papers:
78 categories divided in 4 groups

Taxonomy for user feedback classification[1]

- Developed with SLR on “Classifying user feedback”
- Safety ∉ Taxonomy

Vision videos are concrete by showing system and its use

Empower stakeholders to **experience** a **system** and **report** their suspected ***quality in use* problems**, such as economic, health, and safety risks, even if they have **never actually used** the system.

[1] Santos et al.: *A Taxonomy for User Feedback Classification*. 2nd NLP4RE Workshop, 2019.

Discussion – Selected Vision Video

Vision Video: Tunnels by The Boring Company

- Company of **Elon Musk**
- Already has a **strong social media presence**

Does not hold for average companies!

➢ Need strategy for building social media presence

However, we need even more!

➢ Bridging the gap to the development

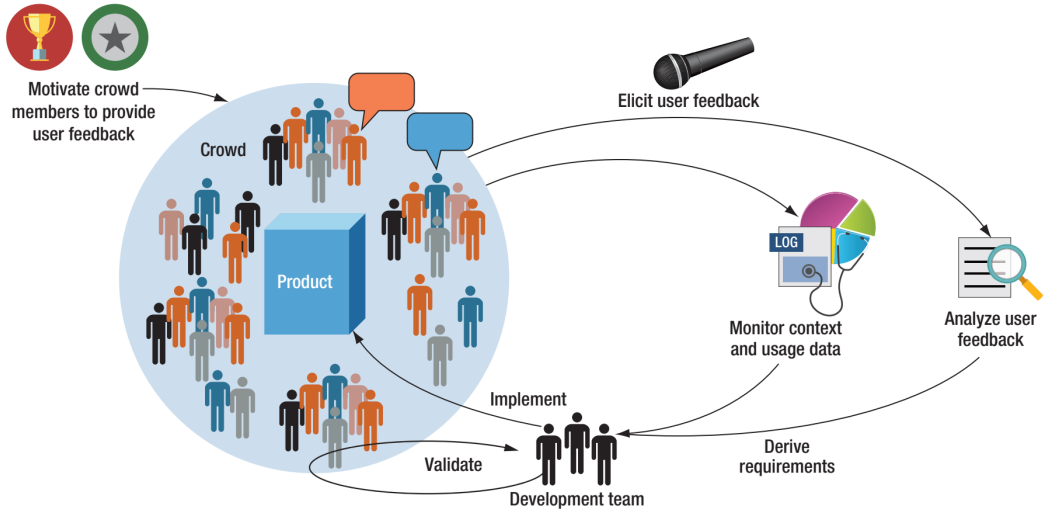
Need holistic approaches for CrowdRE

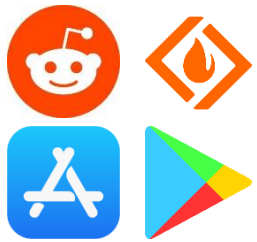
➢ Covering the entire software development process to operationally involve a crowd

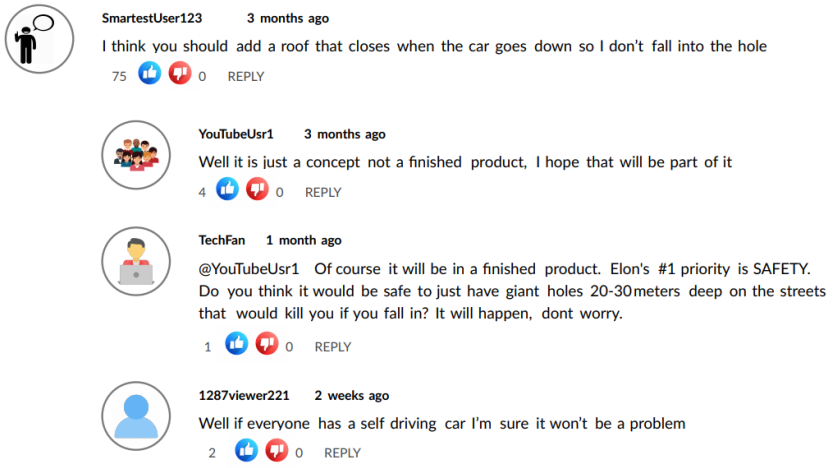
Answer to research question:

1. There is a **potential** of using **vision videos on social media platforms** for soliciting feedback.
2. The number of **comments** and their content are **promising for CrowdRE**.
3. Vision **videos** can be a suitable **option to motivate crowd members** to actively participate by **writing comments** that are a **valuable source of feedback**.

Conclusion







Future work

1. New trend in text classification using deep learning, e.g., BERT[1]
2. Replications & extension of this study, including further data sets
3. Investigate replies
   - Do not address the video itself but discuss the associated comment
   - For example: Use *CrowdRE-Arg* framework[2] to identify arguments for or against a given statement

[1] Devlin et al.: *Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding*. Arxiv, 2018.
[2] Khan et al.: *Conceptualising, Extracting and Analysing Requirements Arguments in Users' Forums: The CrowdREArg Framework*, JSEP 32(12), 2020

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY

TIB

https://www.linkedin.com/in/oliver-karras/

@KarrasOliver

Google Scholar – Oliver Karras

https://www.researchgate.net/profile/Oliver-Karras

Contact
Dr. rer. nat. Oliver Karras
TIB – Leibniz Information Centre for Science and Technology
Data Science and Digital Libraries Research Group
oliver.karras@tib.eu

Leibniz Association