Building Bucky

a generic, chatbot-based microwork platform.

Presentation outline

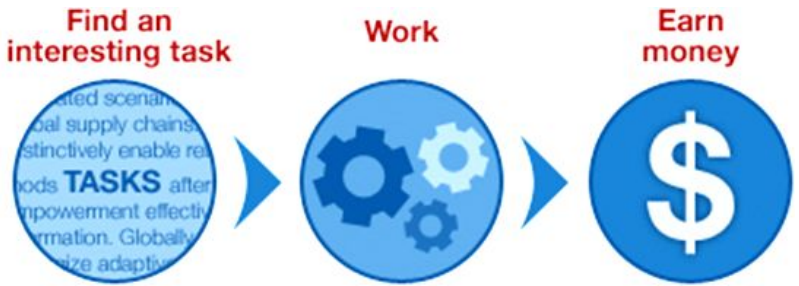
1. Introduction
2. Running example
3. Design
4. Demo
5. Evaluation
6. Future work
7. Conclusion

Introduction

What is microwork?

“**Microwork** is a series of small tasks which together comprise a large unified project, and are completed by many people over the Internet.”

- Wikipedia



The current state of microwork

Best practices

- Simple tasks
- Verifiable tasks
- Pre-testing users
- Show progress

Challenges

- Collaborative tasks
- Recruit/retain users
- Task complexity
- Evaluation

Sources:
[1] Crowdsourcing User Studies With Mechanical Turk
[2] Crowdsourcing systems on the world-wide web

A chatbot can add value

- Smartphone can be used at any place and time
- Does not require a special app
- Great natural interface for interaction

- Requires programming knowledge and effort to get running for a specific organization

A microwork chatbot platform is missing

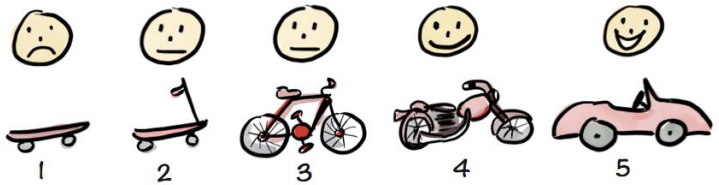
Requirements

Prototype

- Flexible and generic
- Efficient
- Quality control

Final product

- Natural language processing
- Task matching
- Contextual information
- Gamification

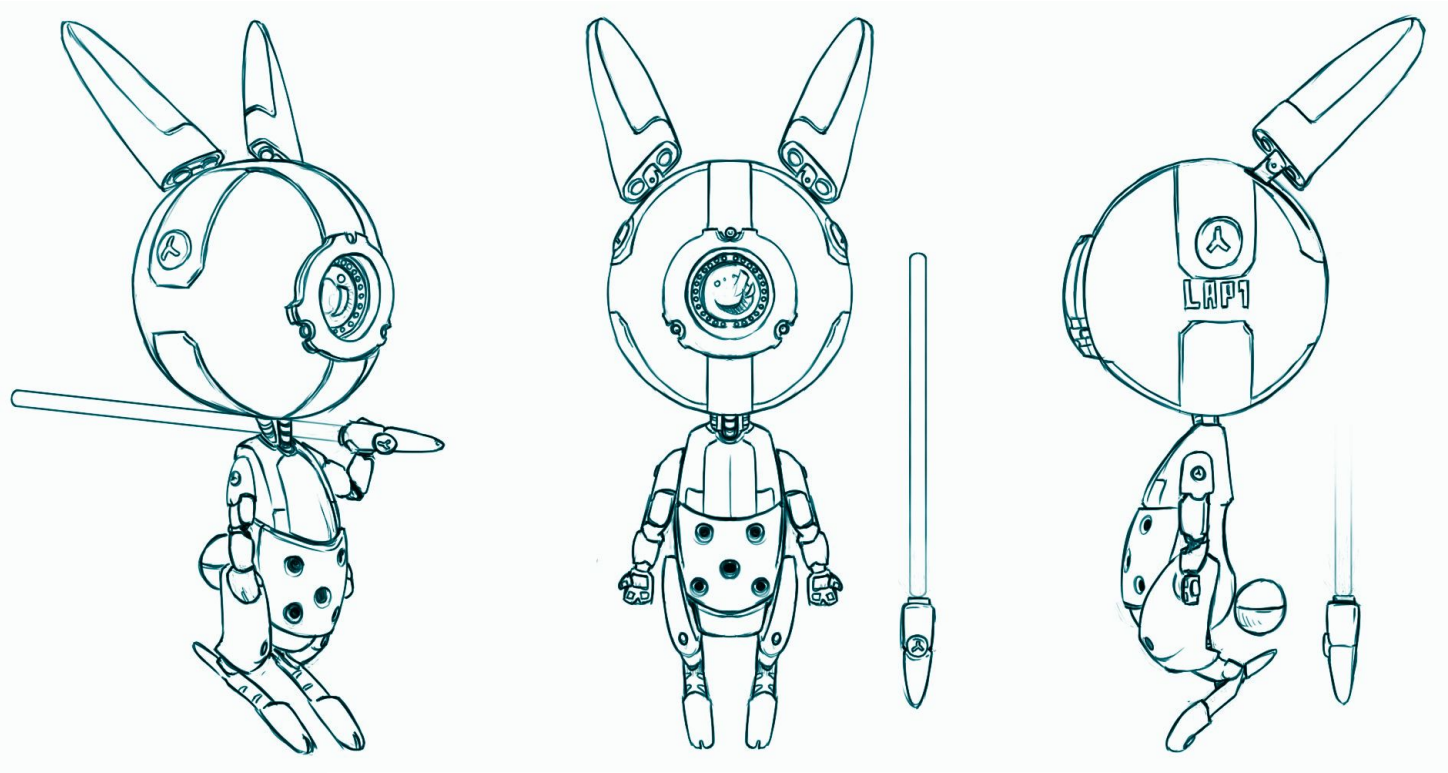


Running example

Running example: Delft municipality

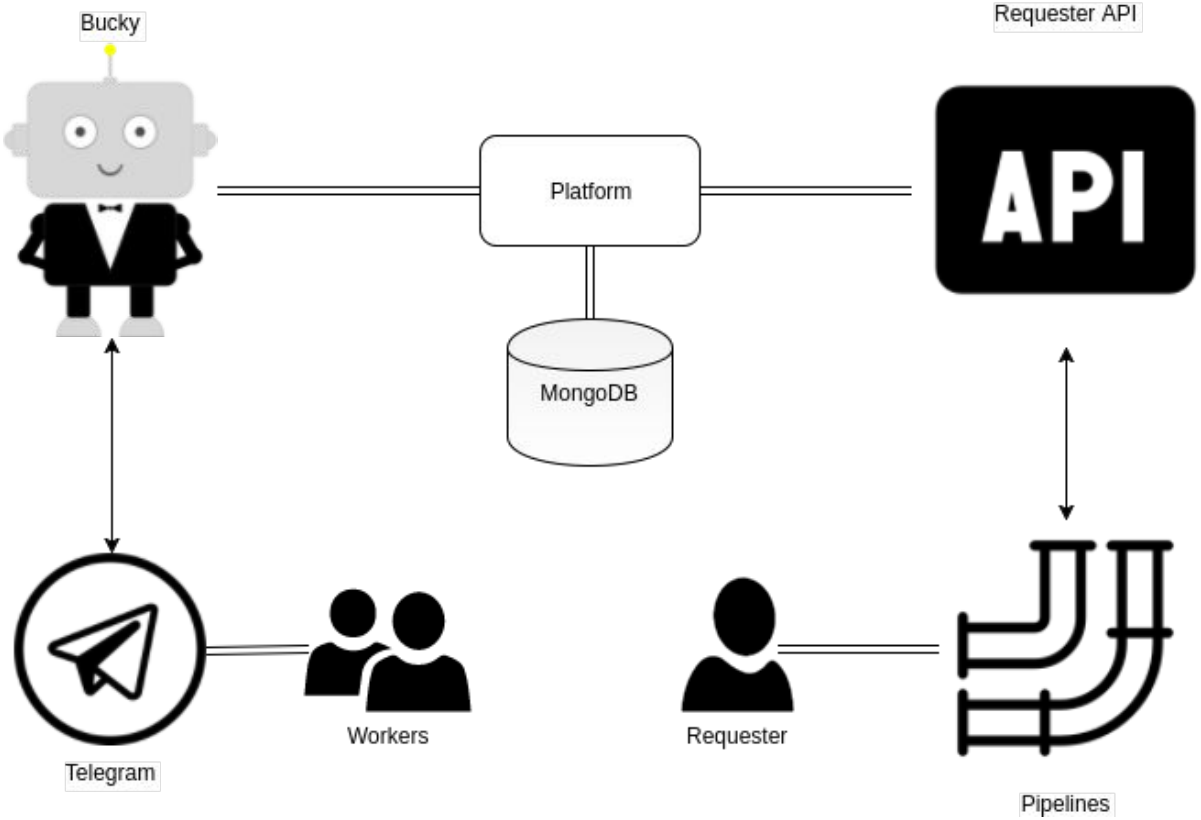
Gemeente Delft

Municipalities and citizens both benefit from better public services.

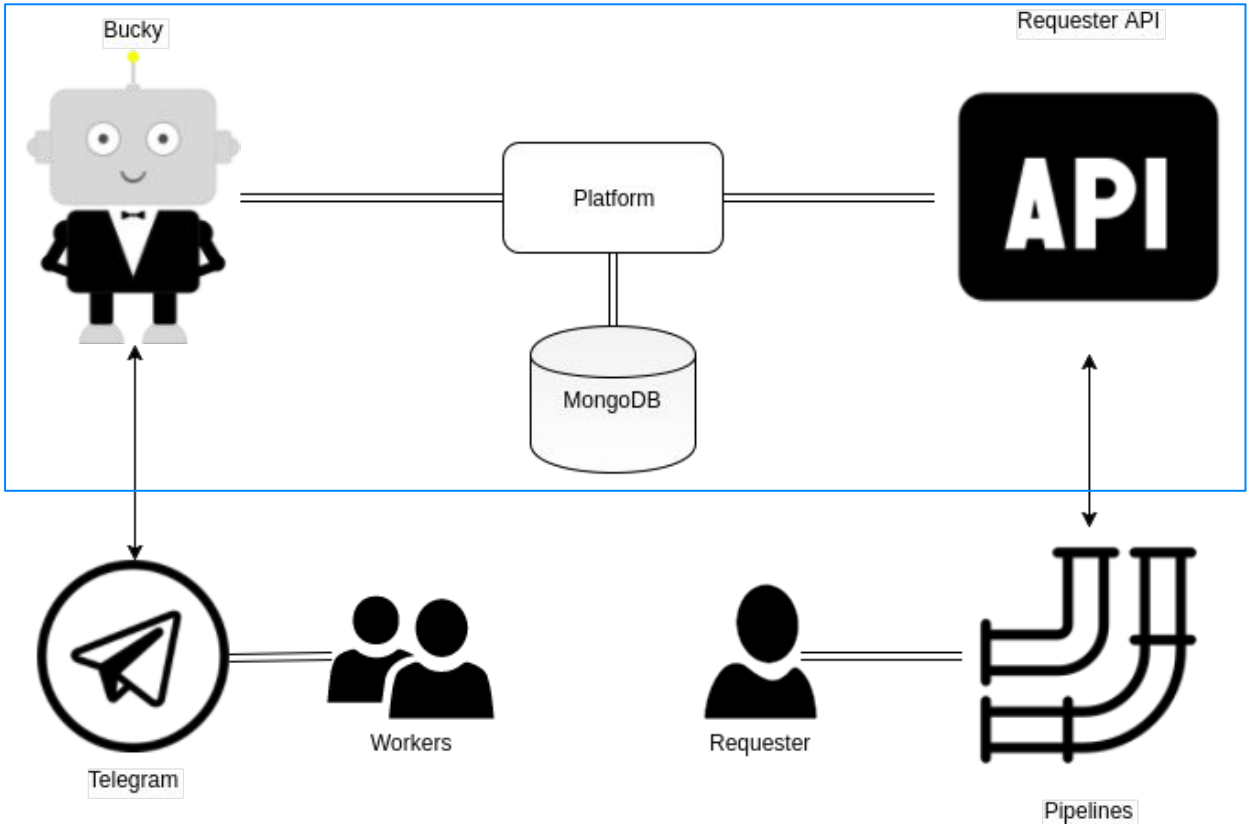


Design

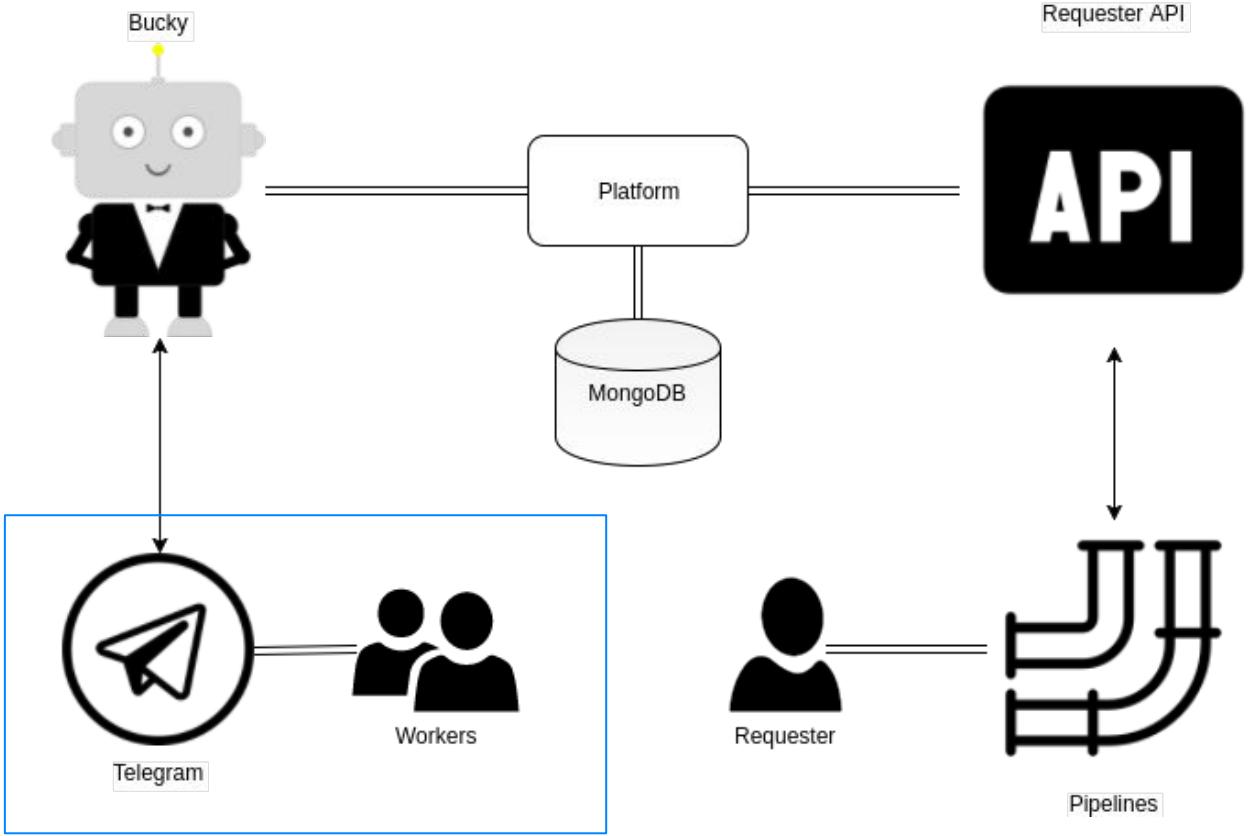
System overview



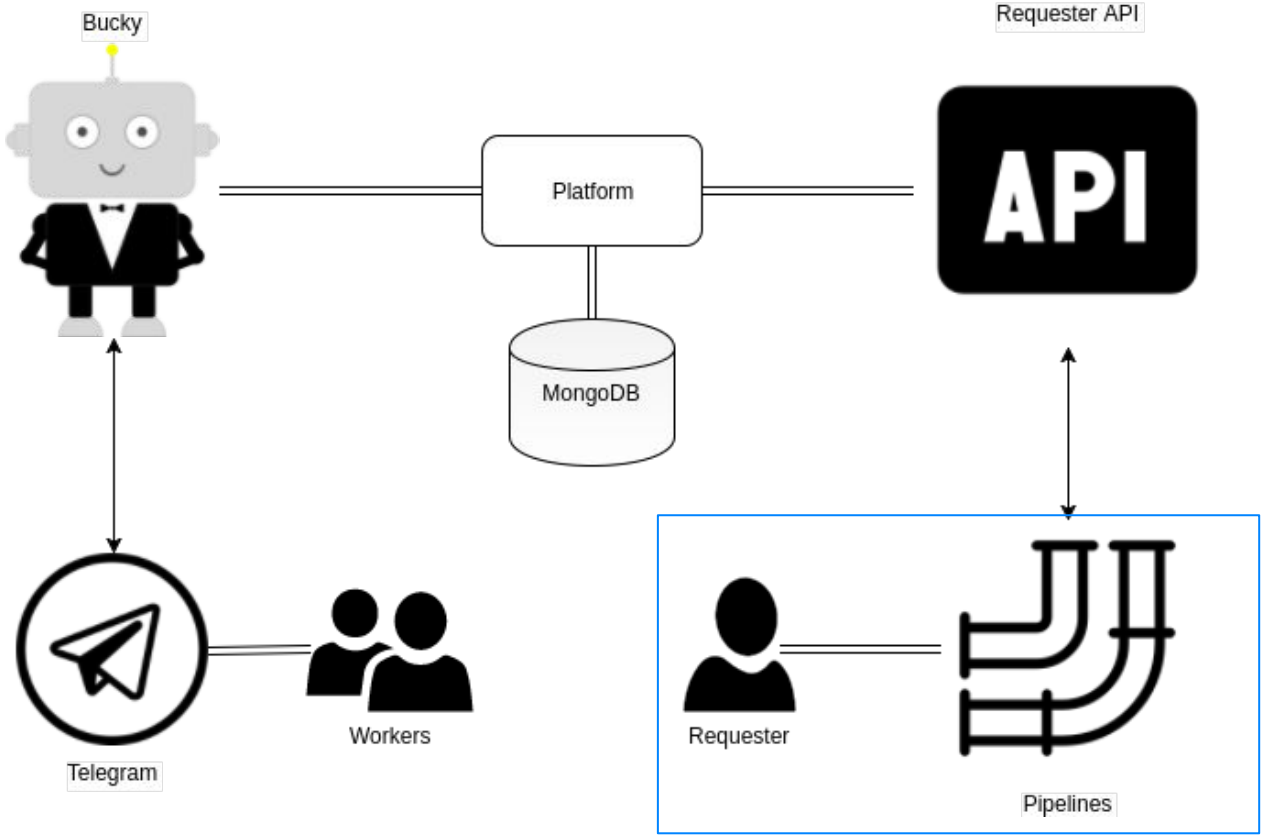
System overview



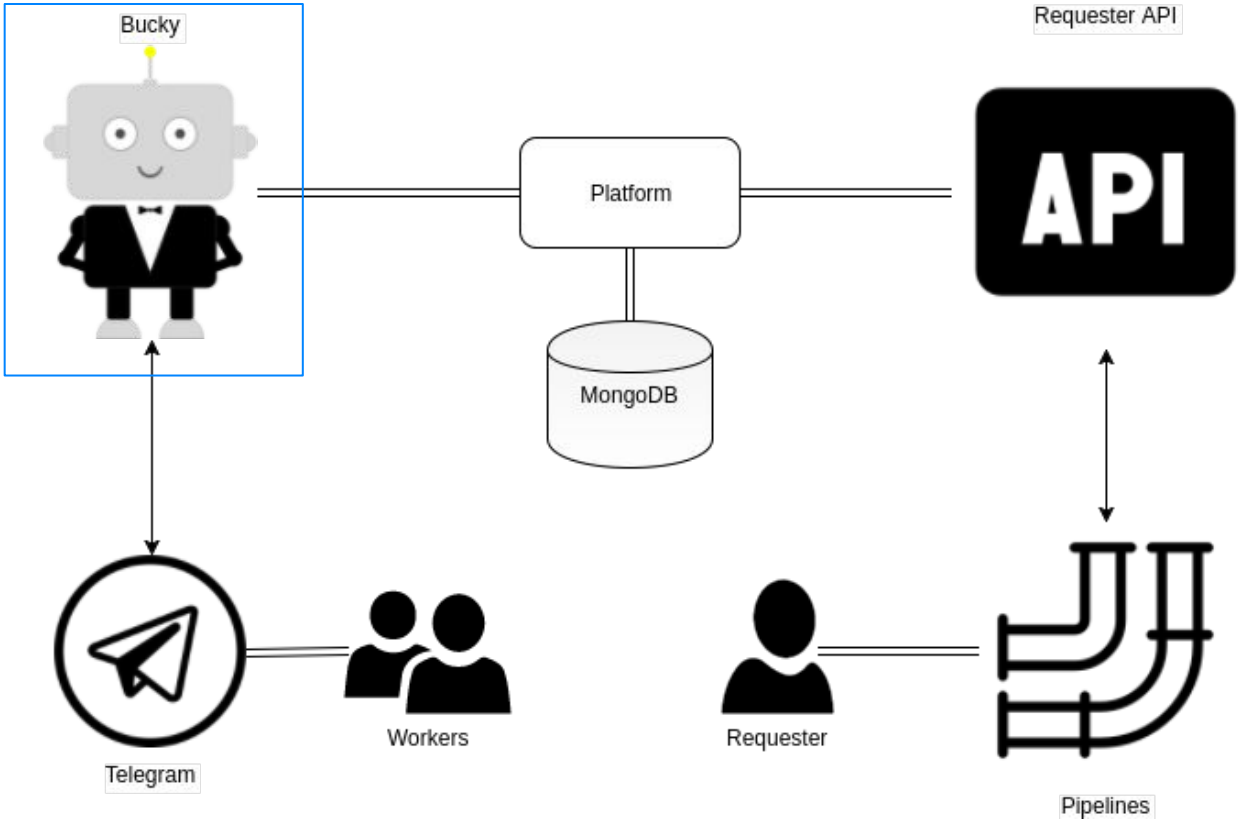
System overview

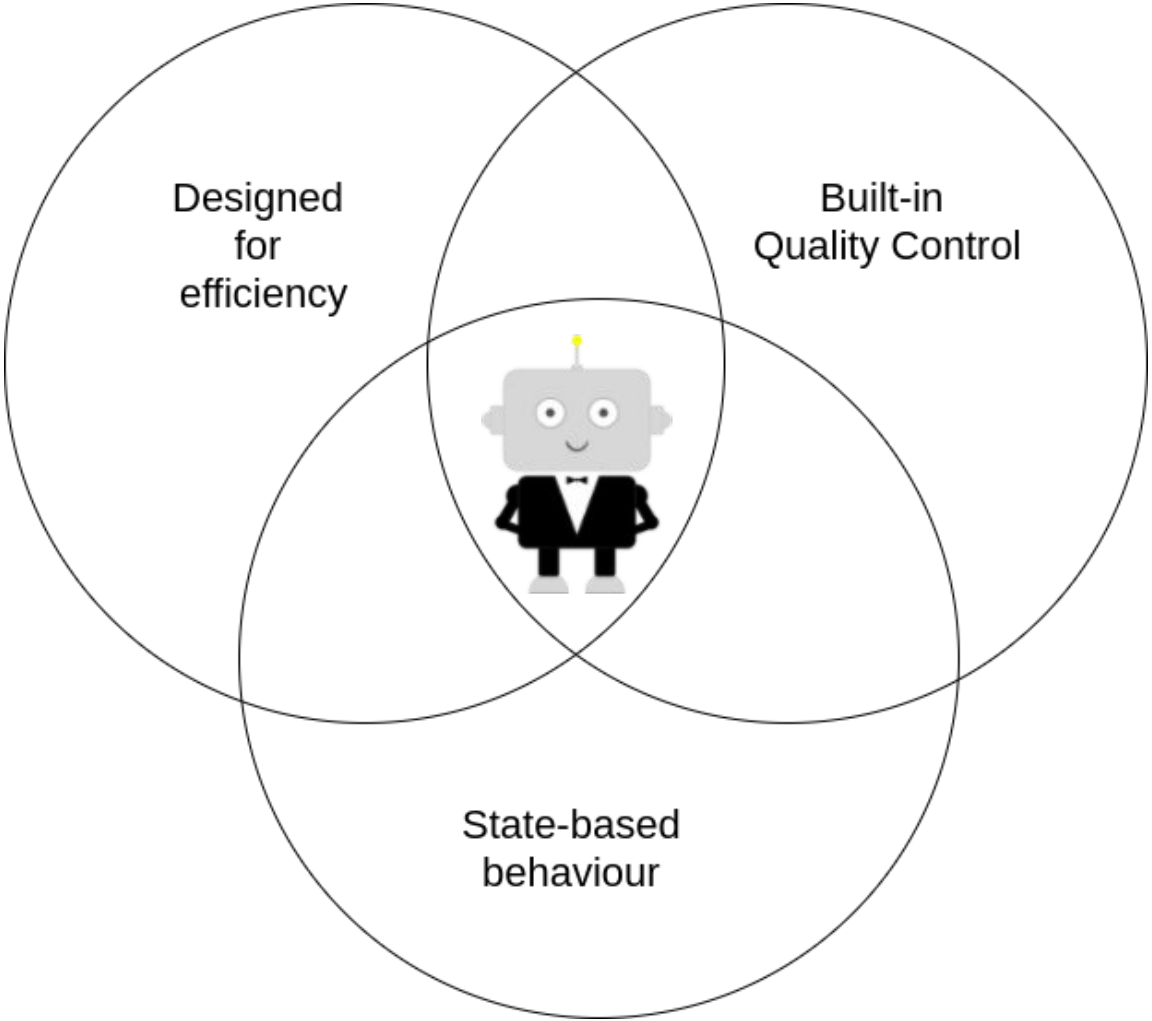


System overview

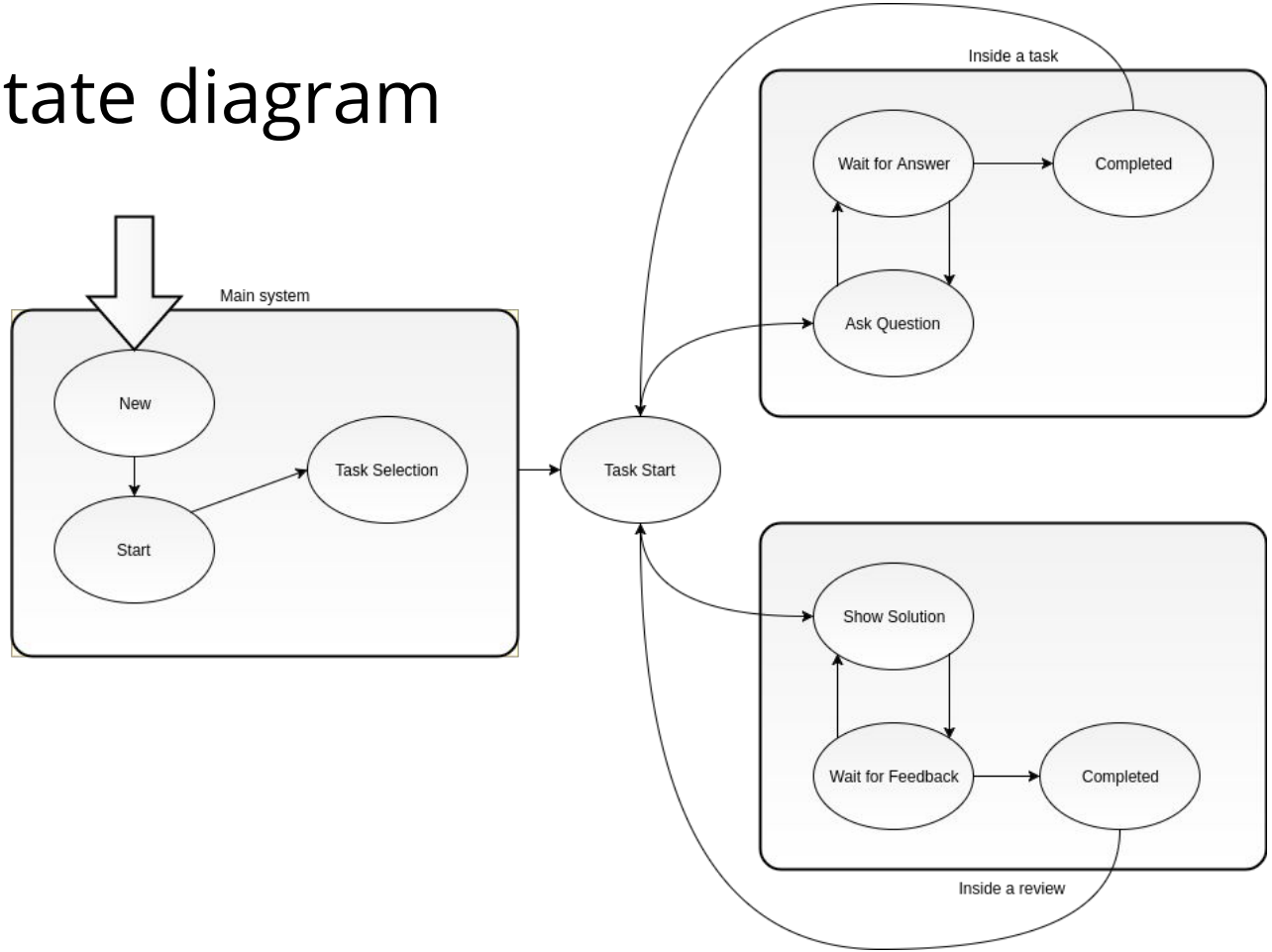


System overview

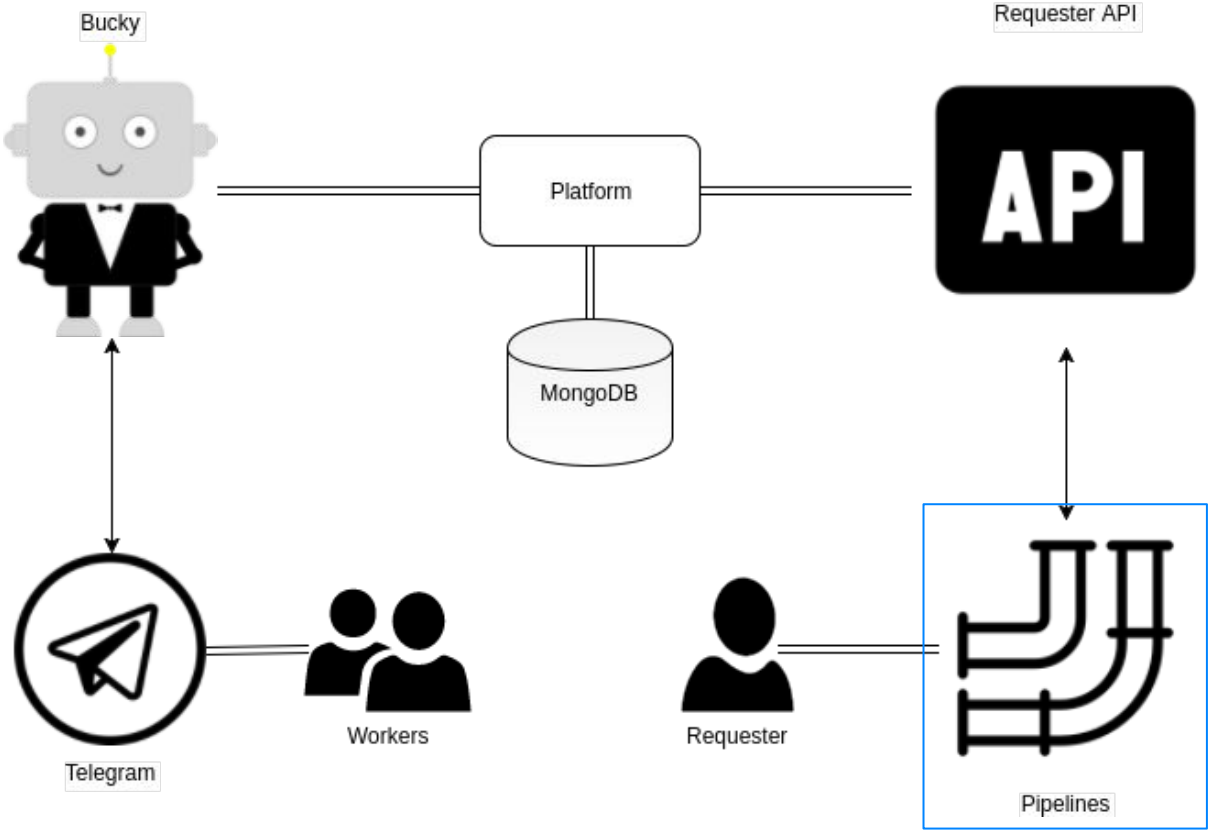



Designed
for
efficiency
Built-in
Quality Control
State-based
behaviour

Bucky state diagram

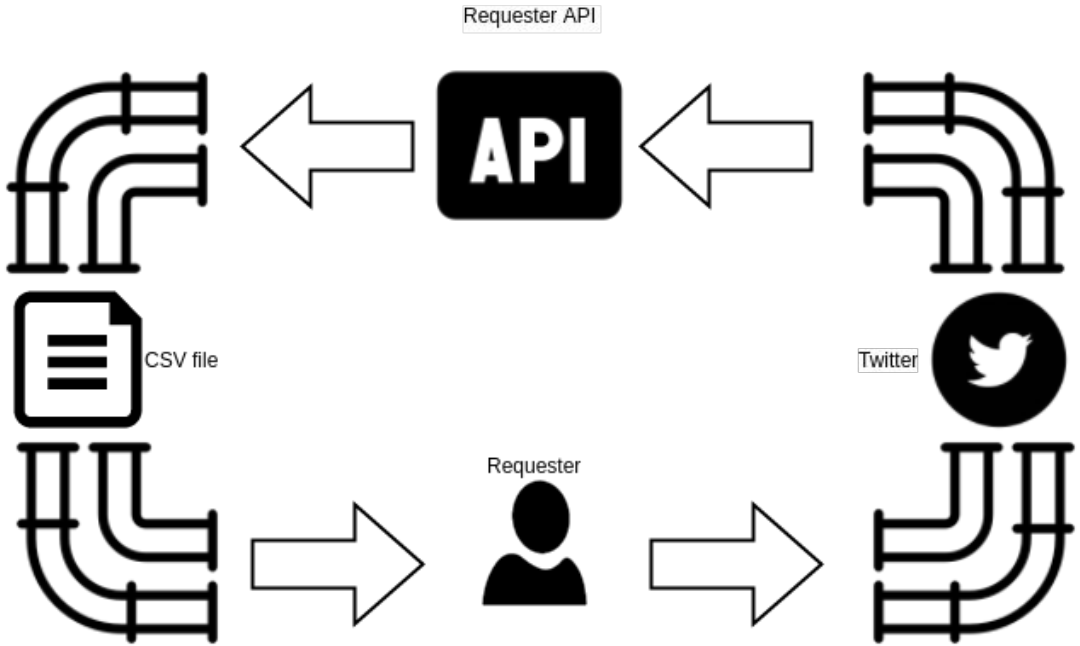


System overview



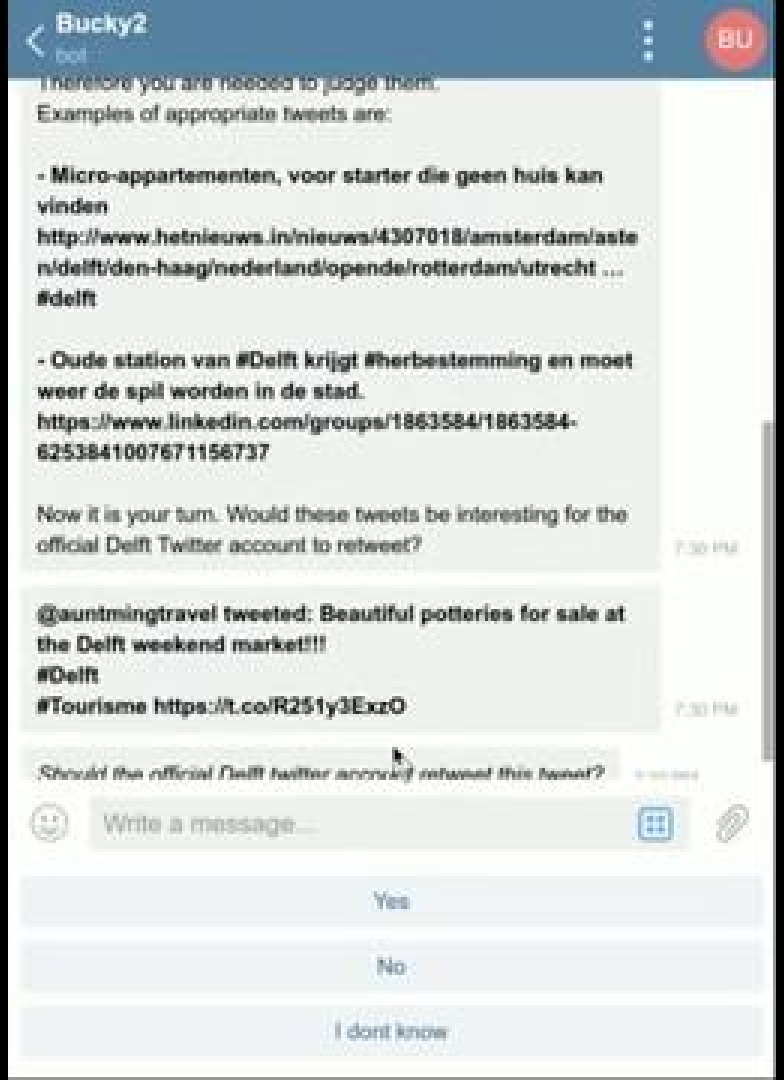
Twitter pipeline

- Fetches tweets with #Delft from Twitter
- Pre-processing to remove tweets with swear words
- Post-processing to create CSV file with aggregated results



BUT WAIT
THERE'S MORE!

Demo


Bucky2
bot
BU
Therefore you are needed to judge them.
Examples of appropriate tweets are:
- Micro-appartementen, voor starter die geen huis kan vinden http://www.hetnieuws.in/nieuws/4307018/amsterdam/aster/delft/den-haag/nederland/opende/rotterdam/utrecht ... #delft
- Oude station van #Delft krijgt #herbestemming en moet weer de spil worden in de stad. https://www.linkedin.com/groups/1863584/1863584-6253841007671156737
Now it is your turn. Would these tweets be interesting for the official Delft Twitter account to retweet?
7:30 PM
@auntmingtravel tweeted: Beautiful potteries for sale at the Delft weekend market!!! #Delft #Tourisme https://t.co/R251y3ExzO
7:30 PM
Should the official Delft twitter account retweet this tweet?
Write a message...
Yes
No
I dont know

Bucky2
bot

BU

How well do you think this image belongs to category PORTRAITS? 8:34 PM

8:34 PM **Not at all**

Good job! You finished the task. Lets do another one! 8:34 PM

8:34 PM

How well do you think this image belongs to category HISTORY? 8:34 PM

Write a message...

very much

Quite well

Medium

Not really

Not at all

Bucky2
bot

BU

during the weekends

Regarding the above message, is the following answer correct?
Q: Please take a picture of the people at this Point of Interest
A:

8:41 PM

8:41 PM

8:41 PM yes

The review is complete! 8:41 PM

Markt (Plaza between City Hall and Nieuwe Kerk) during the weekends

8:41 PM

Please take a picture of the people at this Point of Interest

8:41 PM

Write a message...

Recap

- Use pipelines to create tasks and fill them with content.
- Easy and quick to do tasks, even when you have a short amount of time.
- Specific answers in buttons to reduce effort even more.
- Review tasks from other people to assure quality of answers.

Evaluation

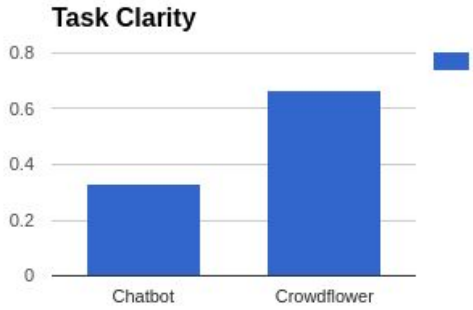
Evaluation

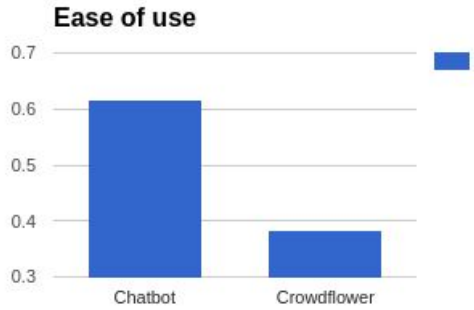
Goal

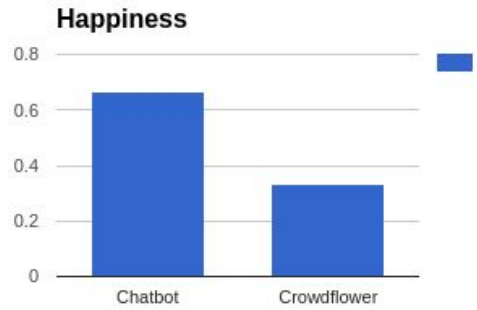
Setup

Experiment execution

Results







Performance

	Chatbot (sec)	Crowdflower (sec)
From start until task	20.43	33.36
Per task	15.11	11.36

Evaluation Conclusion

Without experience

	Chatbot (sec)	Crowdflower (sec)
From start until task	20.43	33.36
Per task	15.11	11.36

With practise

	Chatbot (sec)	Crowdflower (sec)
From start until task	13.51	30.89
Per task	7.70	9.03

Future work

Prototype

- ✓ Flexible and generic
- ✓ Efficient
- ✓ Quality control

Final product

- ✓ Flexible and generic
- ✓ Efficient
- ✓ Quality control
- ✗ Natural Language Processing
- ✗ Task matching
- ✗ Contextual information
- ✗ Gamification

Conclusion

Conclusion

Microtasks & why a chatbot

Running example: Municipality of Delft

Design decisions

Demo & cross check requirements

Experiments & Evaluation

Future work

@Buck_a_bot

We have the answers. He has the tasks.