Joint project on detecting malicious registrations: a follow-up

Thijs van den Hout (.nl) & Maarten Bosteels (.be)

CENTR 23rd R&D / 49th Tech

Agenda

- Goal of collaboration
- Project recap
- Evaluation
- Lessons learned
- What's next?

Goals of joint project (defined last time)

Exchange ideas for more effective detection

Jointly develop code

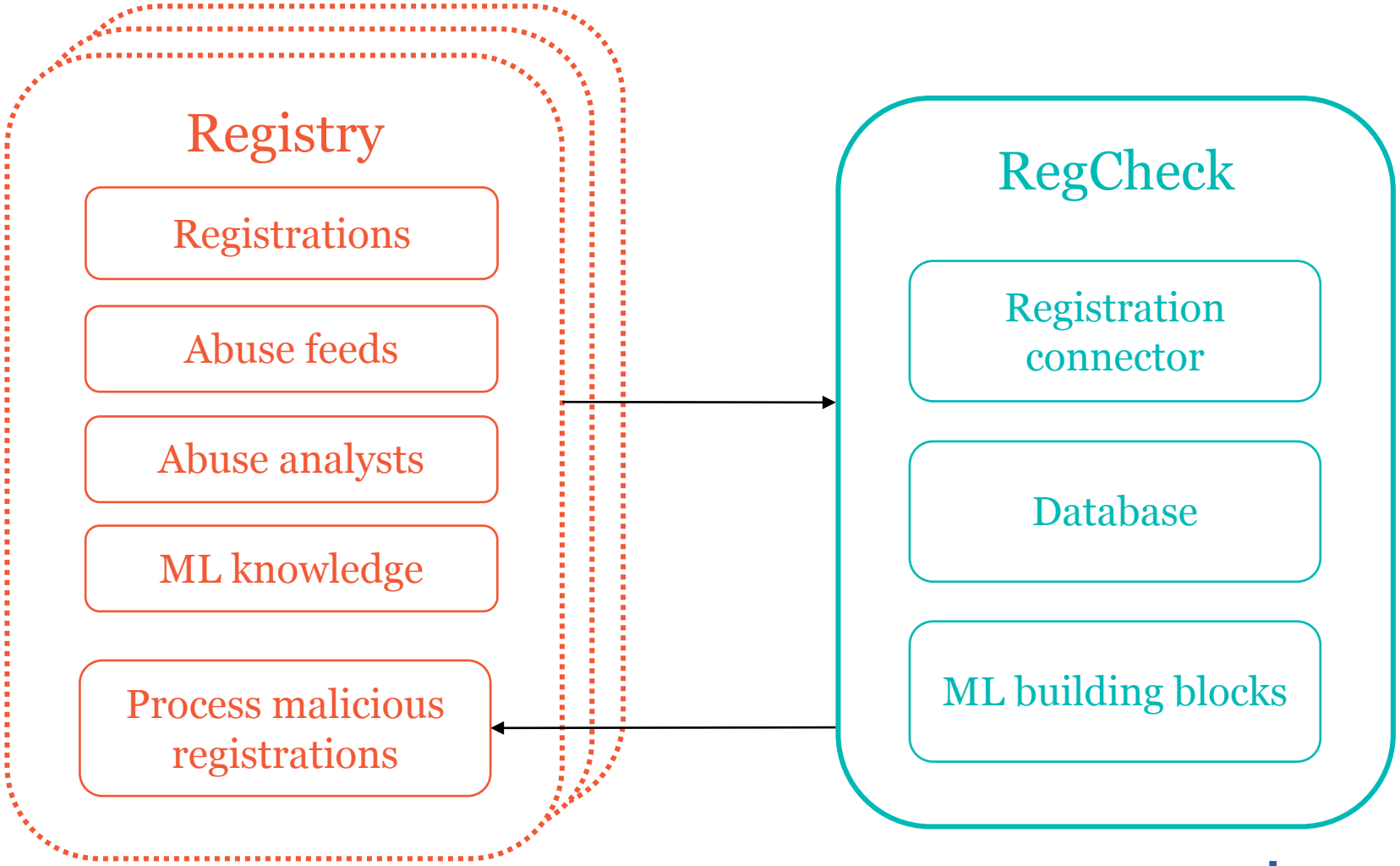
Blueprint for other registries

dns belgium SIDN LABS

Collaboration phases

March - December 2022	December 2022	January 2023 - now
Exploration	Agreement	Joint development

RegCheck design



RegCheck design: customizable

```yaml
version: 1
detector: logistic_regression
model_name: my_first_model
model_description: Logistic regression classifier trained using dummy data
save_path: /path/to/regcheck/models
datasets:
  - dummy_train_label1
  - dummy_train_label0

training_options:
  reputation_timeframe: 30
  features:
```

Benefits of collaboration

- [some items redacted]

- Discussion about design and policy choices

Evaluation at .nl

	Art. 16	Art. 18
Domain names	1100 (45% of total)	390 (40% of total)
ID verified	56	28
Registrants	258	208
ID verified	10	12

Table 1: Verification of registration data procedures initiated due to a RegCheck notification.
(January through September 2023)

Evaluation at .nl

	✓ RegCheck	✗ RegCheck
✓ Netcraft	46	442
✗ Netcraft	2,247	369,279
	2,293	**369,721**

Table 1: Comparison between RegCheck and Netcraft notification

	✓ RegCheck
✓ High-risk	425
✗ Low-risk	1,658
	2,083

Table 2: Analyst labels for RegCheck notifications

(February through June 2023)

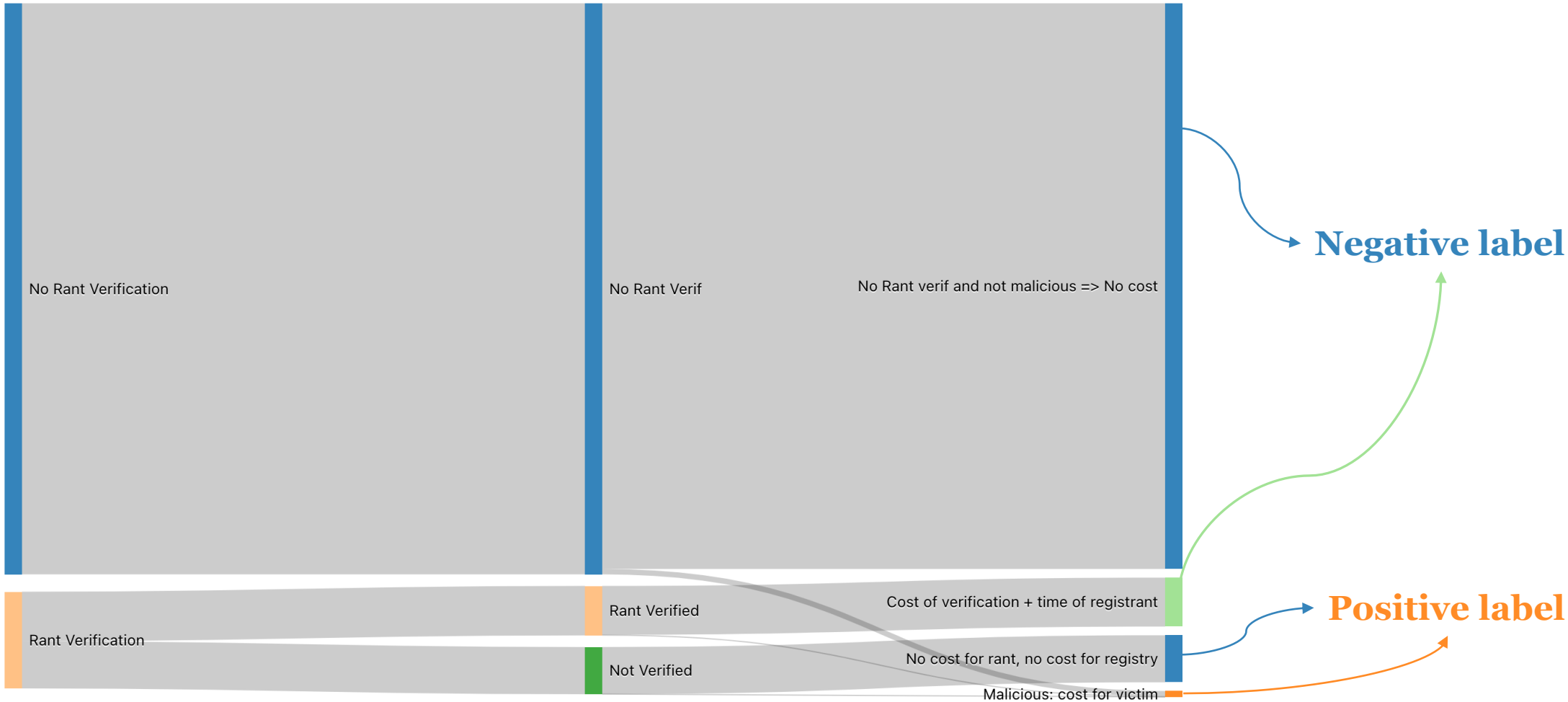
dnsbelgium SIDN LABS

Evaluation is hard!

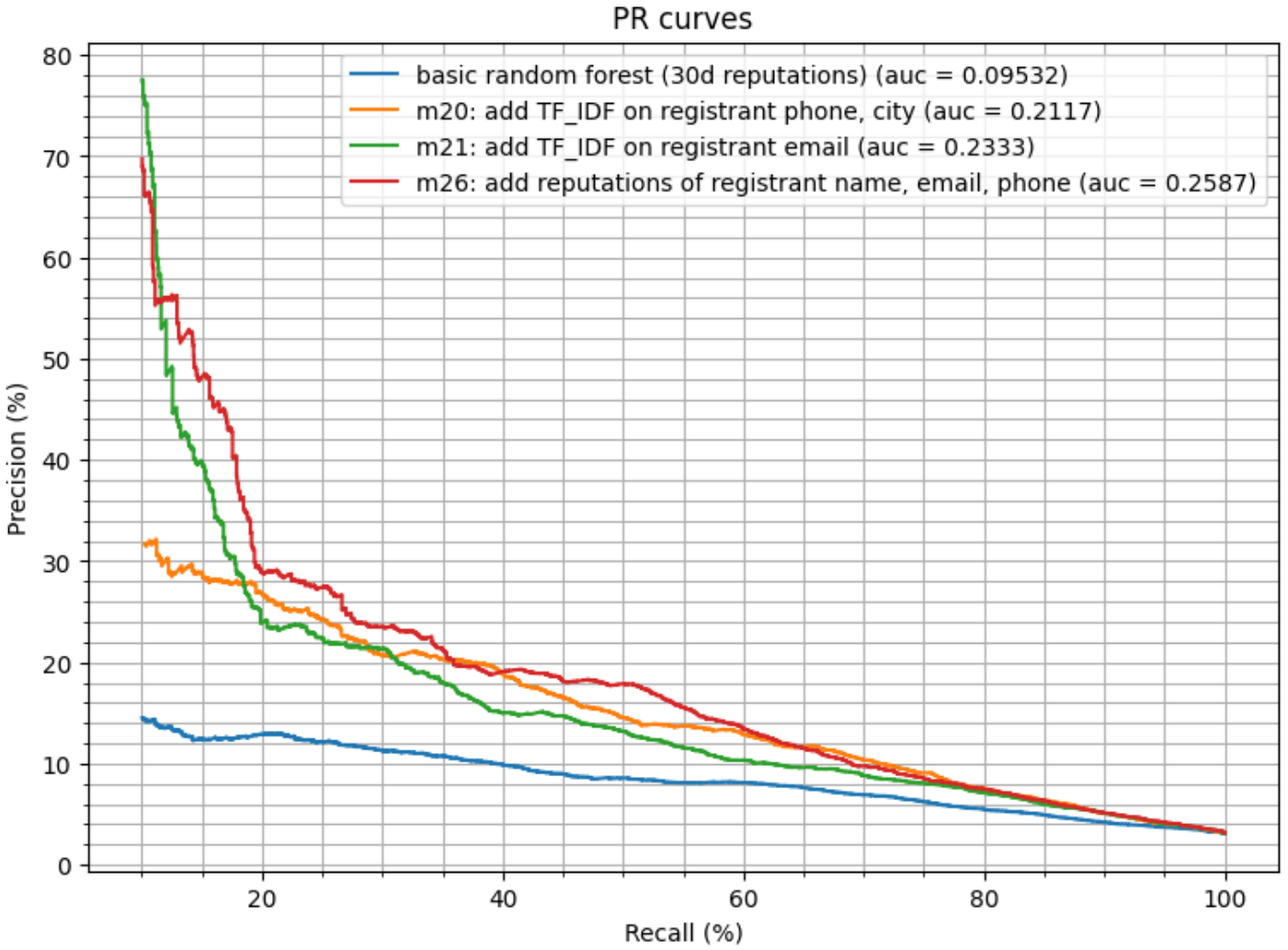
- Recall is a heuristic metric
- Deployment in the wild = interaction with evaluation set
- How much abuse is prevented?
- Qualitative results are positive

dns belgium SIDN LABS

Labeling the data at .be



Evaluation at .be



Still to do:

- Hyper parameter tuning
- Higher sample weights for malicious regs

dnsbelgium SIDN LABS

Lessons learned

- Sharing experiences and different views on abuse is valuable
- Collaboration works, even with diverging policies
- Collaboration inspires innovation
- More people = more opinions

Project goals update

Exchange ideas for more effective detection

Jointly develop code

Blueprint for other registries

What's next?

.nl

- Automatic ID verification
- Share scores with registrars

.be

- Hyper param tuning
- Registrant verification
- Close feedback loop
- Sample weights

collaboration

- More registries
- Add features
- Improve prediction latency

dnsbelgium SIDN LABS

Q&A

thijs.vandenhout@sidn.nl
thymen.wabeke@sidn.nl
maarten.bosteels@dnsbelgium.be