DEX:
Domain name Ecosystem eXplorer

CENTR R&D – 26-11-2020
Maarten Wullink, Moritz Muller, Thymen Wabeke
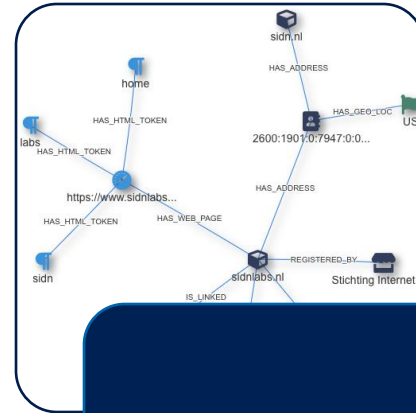
SIDN LABS

Agenda

Why, what, how

Neo4J

Demo & evaluation


Discussion

Motivation: safe .nl zone

- Tackling abuse requires various types of information

- Information stored in various data sources (silos)

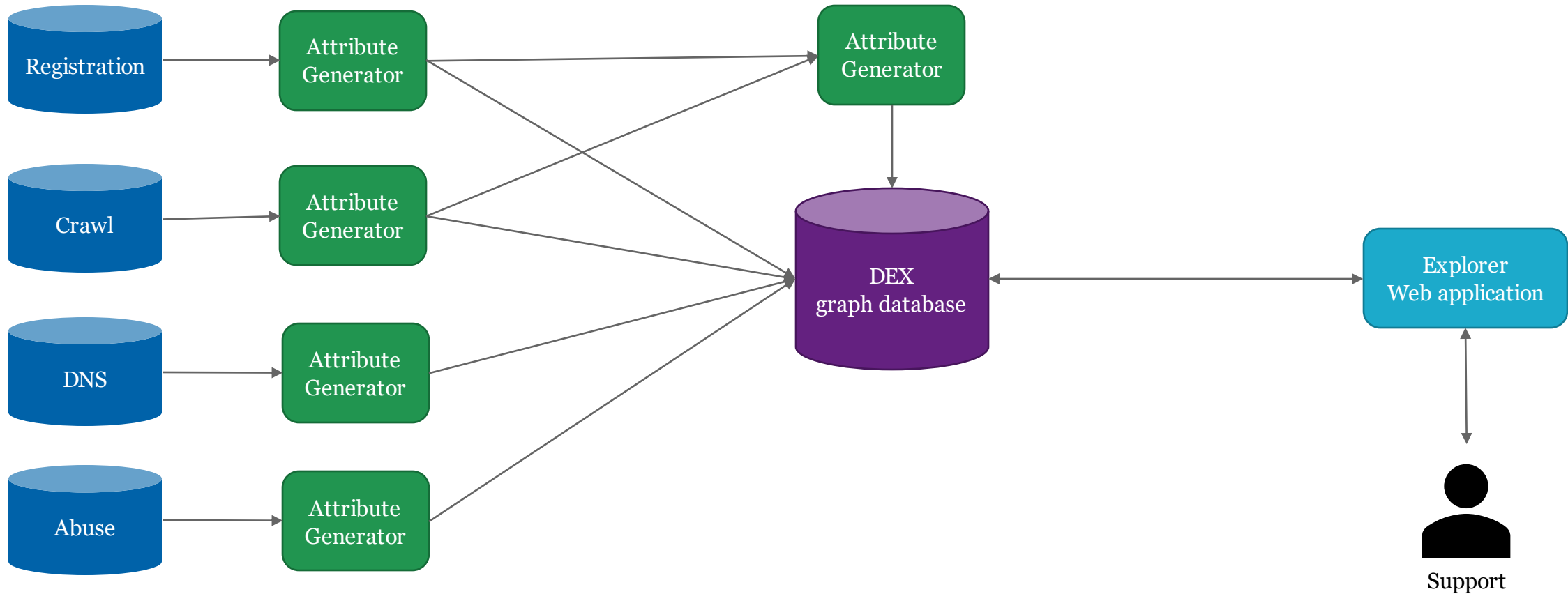- Information not easily accessible to Support

Goal: tackle abuse thoroughly and efficiently

- Create an easy to navigate domain name ecosystem for non-technical employees

- Identify domain names associated with malicious domain names

- Recognize trends and new types of abuse earlier
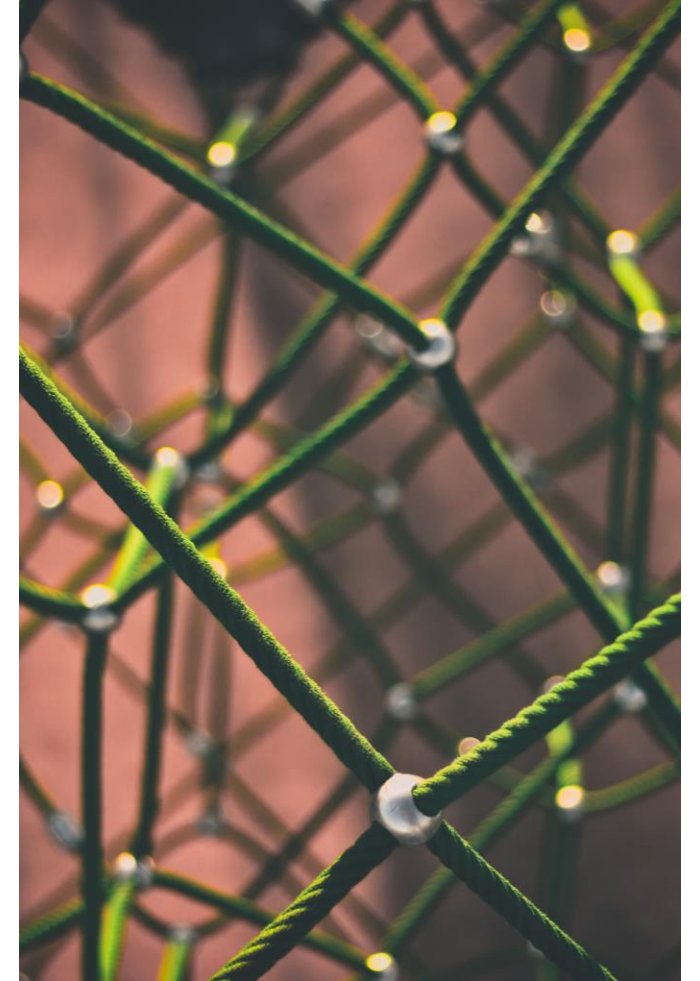
Architecture

NEO4J experience (pros)

- Natural fit for highly connected data

- No need to join (large) tables during query execution

- Many client libraries available (e.g., JavaScript, Go, Python)

- Powerful query and visualization tool ("Neo4j Browser")

  - Not a perfect tool for non-developers

- Graph algorithms, e.g. for centrality and community ("Graph Data Science")

NEO4J experience (cons)

- Cypher queries differ from SQL queries

  - Declare relations instead of joining tables

- Requires indexes, otherwise performance is low

- For optimal performance, the database should fit in memory

  - 18 GB for ~6M domain names, ~41M nodes and ~113M relations

- Not great for historical data

- Community (free) edition lacks some functionality

  - E.g., logging, run graph algorithms using multiple cores

Demo time

https://www.youtube.com/watch?v=wHDkC6KOTpI

User experience

- Used by SIDN Support team for ~2 months

- DEX mainly used for exploration,
  less for case building

- New attributes very valuable (such as subdomains,
  social media accounts, Chamber of Commerce)

- Investigating what new features to add next year,
  develop an ethical/legal framework for broader usage
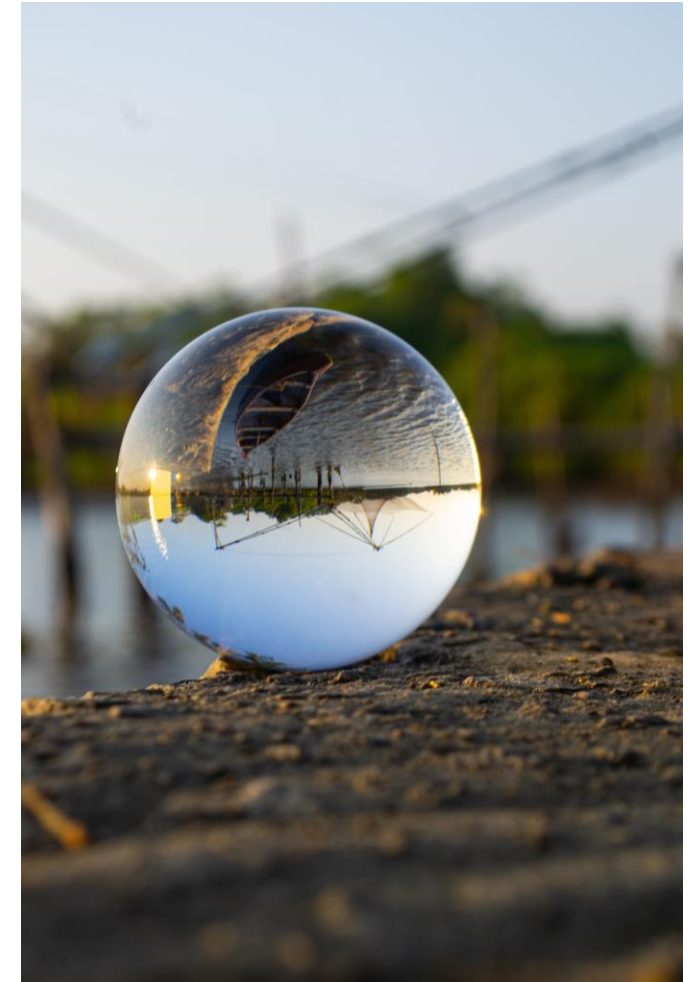
Questions and discussion

- Feedback?

- Additional attributes?

- Other use cases?

- Interest in using the tool? Join us in further development.
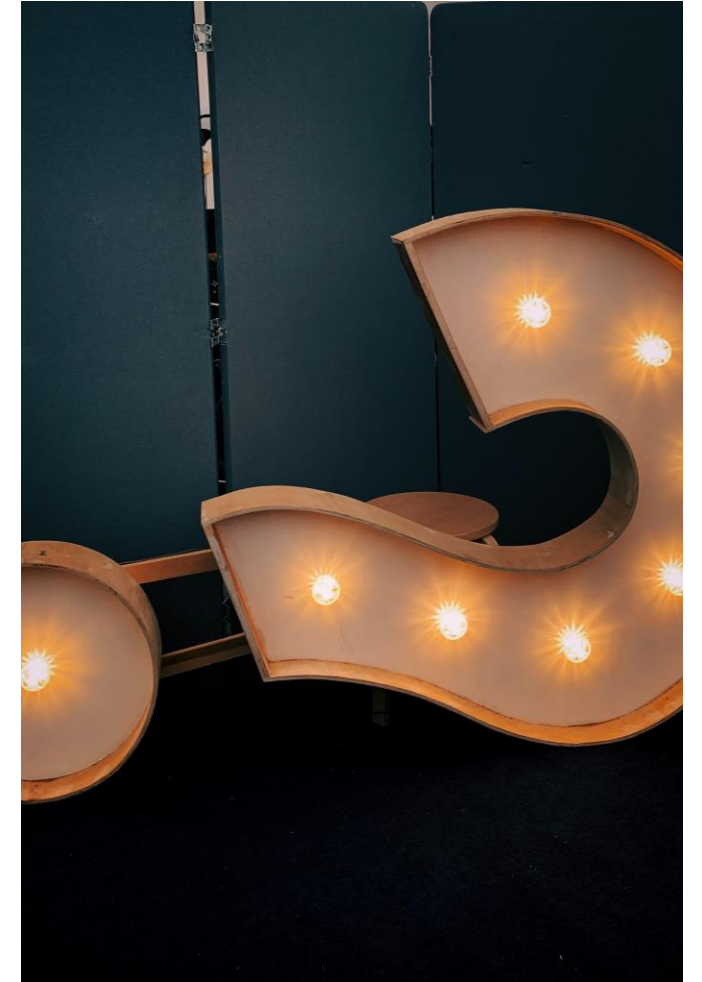
Follow us

SIDN.nl

@SIDN

SIDN

Thanks for your attention

SIDN LABS