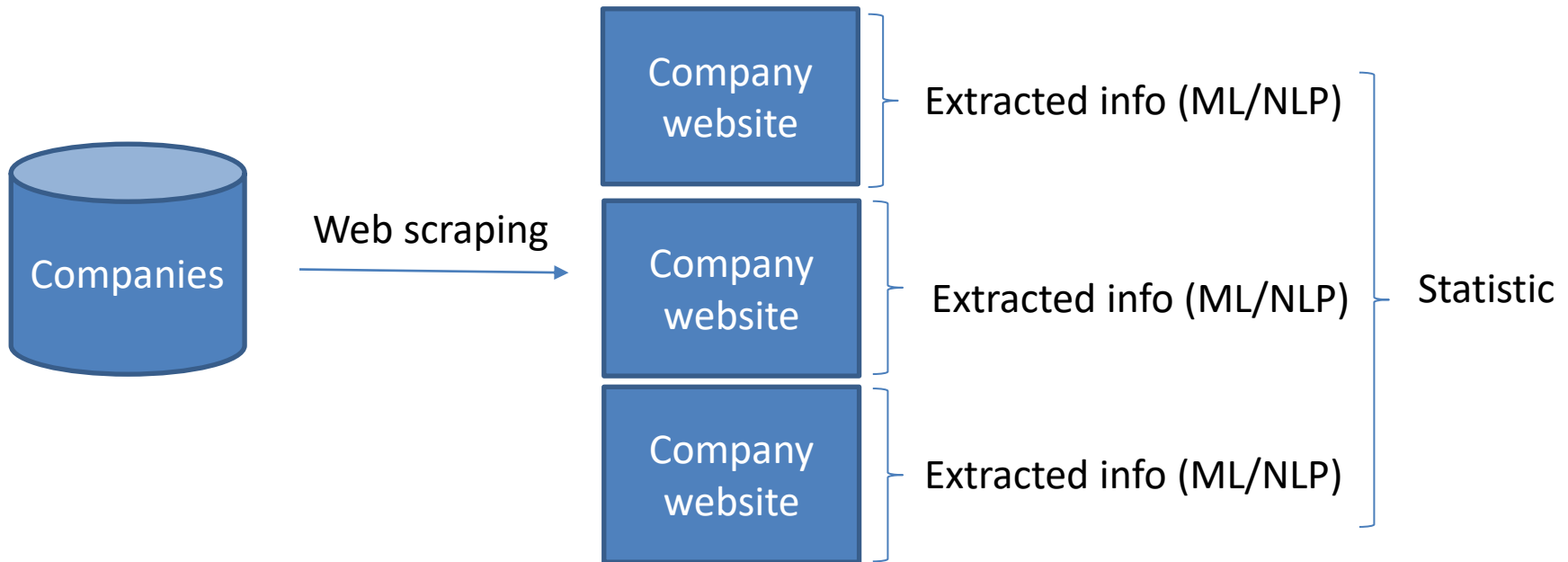


Co-creating new indicators using text scraped from company websites

Michael Reusens
Data Science Coordinator - Statistics Flanders
michael.reusens@vlaanderen.be

Goal

1. Make new indicators on companies using text scraped from company web pages



Goal

1. Make new indicators on companies using text scraped from company web pages
2. Benefits w.r.t. traditional survey approach
 1. Response burden
 2. Coverage
 3. Frequency

Approach

- With interested countries, define a company-related indicator of interest, for example:
 - Number of innovative companies (cfr. Work by CBS)
 - Number of companies working in
 - Data science
 - Circular economy
 - ...
- Design a methodology together, implement per region
 - Usefull work has been done (e.g. CBS Netherlands)

Expected lessons learned

- Web scraping – NLP – machine learning
- Doing it in together and in parallel
 - Efficient collaboration towards the same goal
 - How well does the method generalise over regions?
 - How comparable are the statistics over regions?

Looking for

- Organizations who want to develop such an indicator
 - Note on data: url sample is enough to participate
- Feedback group
 - Will not develop anything themselves

Thank you

- michael.reusens@vlaanderen.be