CBS RobotTool

An *interactive* tool for price analysts to detect price changes on websites
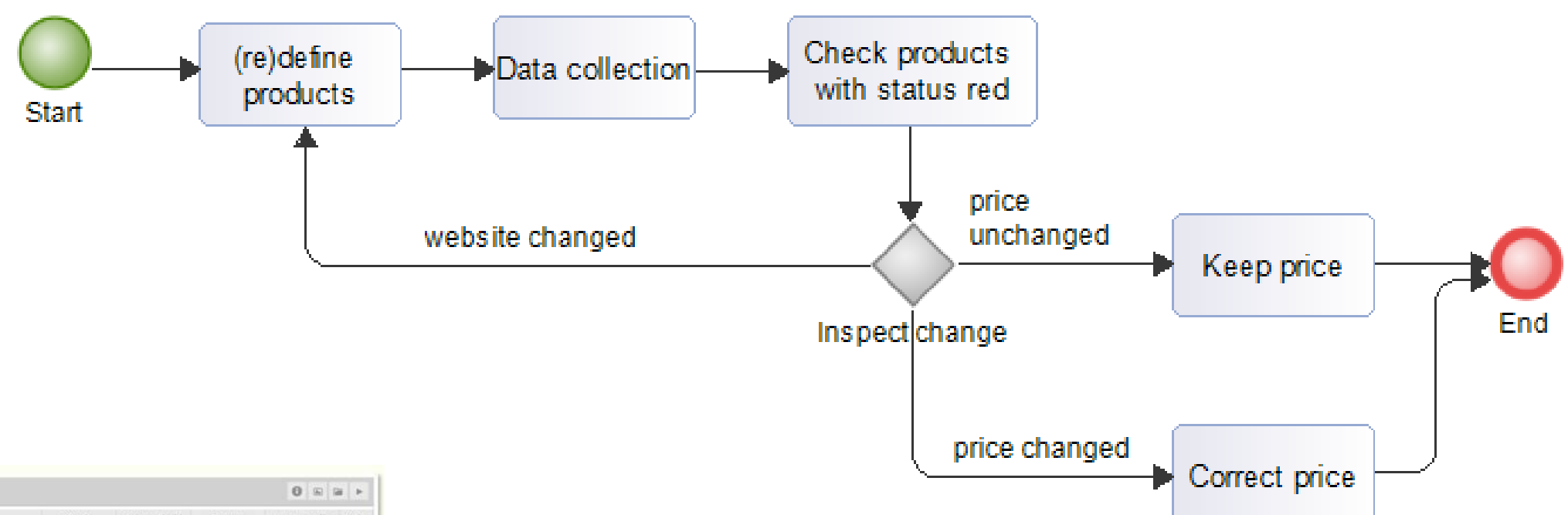
The concept

- **Configure** products
- **Run** data collection
- **No change** detected:
    => fully automated processing
- **Change** detected:
    => manual control

Typical use

- When bulk-scraping is not applicable
- Efficient semi-automated price collection
- No change in methodology needed: use basket

Workflow

In more detail

1. Run data collection:

2. Change detected: inspect change

3. Configure by XPath:

Production use, 2020 Q1:

- Number of products: 123
- Number of prices / month: 2218
- Number of concurrent users: 5
- Increase next years:
    266 products
    1072 price observations

4. Add XPath for product:

5. Export prices (csv) into database

Statistics Netherlands

hjm.windmeijer@cbs.nl & o.tenbosch@cbs.nl
In corporation with price analysts of Statistics Netherlands

Download

https://github.com/SNStatComp/RobotTool