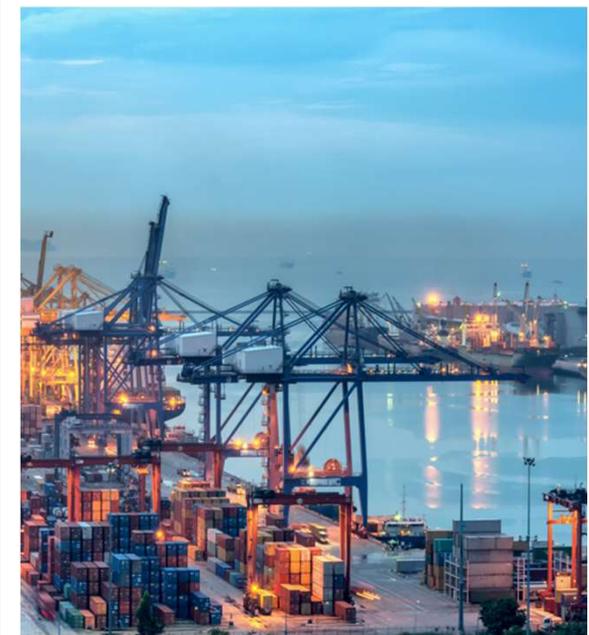


# Anders Alpsten Swedish Customs

18 November 2021

01100010110100101100101001  
**P**ROFILE

A project funded by the European Union's Horizon  
2020 research and innovation programme under the  
Grant Agreement No 786748.



# The Profile project

---

- Objective: Develop modern data analytics and leverage Big data and open sources for customs risk management
- Swedish customs is responsible for Work package (WP) 5 in the project
- In the WP Swedish customs, Norwegian customs, The Swedish defence research organization (FOI) and the Norwegian defence research organisation (FFI) work together
- The method for the work is called Living Lab, in this case it means that the customs has the “business” knowledge and the research organisations has the technical knowledge. The point is to work together to try to develop IT based risk analysis tools
- In this WP a Member state is working together with non Member state to try to improve risk analysis when goods crosses the border
- The results in the project are still in test



# The Profile project

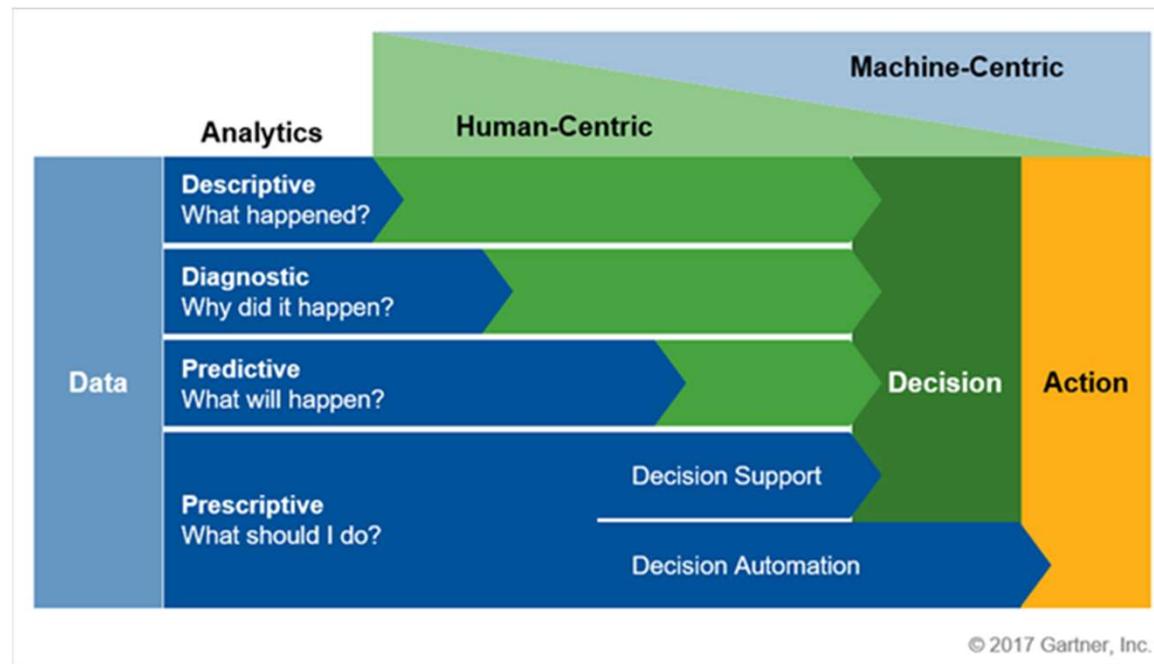
---

- With the methods we are trying to find:
  - Anomalies
  - Errors
  - Fraudulent declarations and shipments

So we look for intentional and unintentional errors, but in the long run we are looking for frauds. But we don't know what is what in the beginning.

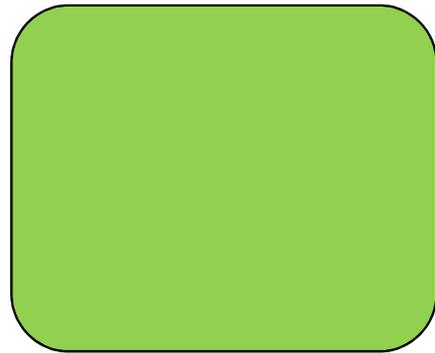


# Profile project – analytics

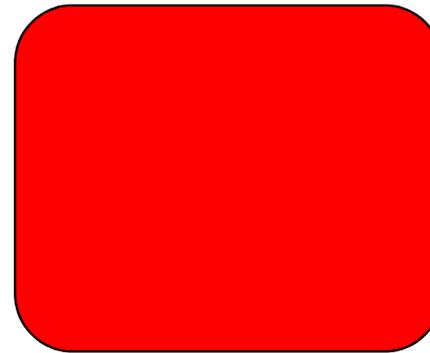


# Analysis results

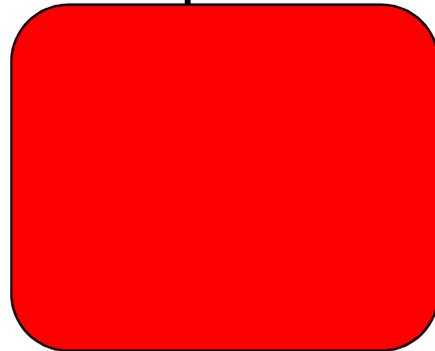
---



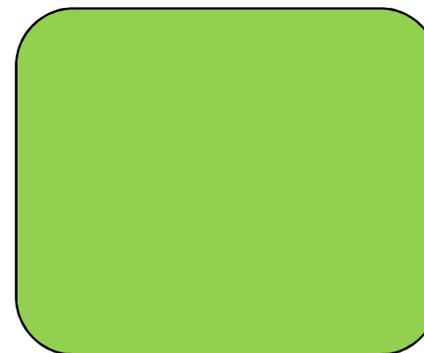
True positives



False negative



False positive



True negative



# The Profile project – techniques used

---

- **Random forest classifier**
- **Exchange of models**
- **NLP on free text data fields**
- **Clustering**
- **Data visualization**
- **Mirror statistics**



# The Profile project

---

- **Random forest classifier**
- Allows to make a more dynamic analysis
- Predicts the correct Harmonized system (HS) code
- Uses a number of fields in the declaration
- **Exchange of models**
- Models trained on Norwegian data have been used on Swedish data and vice versa. Is a good way have a second opinion on models



# The Profile project

---

- **NLP on free text data fields**
- Compares free text fields to HS code
- A bit hard to use because of that the texts are very short, normally one word
- **Clustering**
- A way to find "outliers" compared to the normal picture of the use of for example a HS code



# The Profile project

---

- **Data visualization**
- Can give a quick overview of possible fiscal gains over periods
- Can aid analysts in search of fraudulent behaviors
- **Mirror statistics**
- Compares export and import statistics between countries
- In theory import – export between countries should match



# Thank you for listening Swedish Customs

18 November 2021

**P**<sub>01100010110100101100101001</sub>  
PROFILE

A project funded by the European Union's Horizon  
2020 research and innovation programme under the  
Grant Agreement No 786748.

