Cerberus and the future of risk at the UK border

Economic and Social Research Council

CenSoF Centre for Sociodigital Futures

Travis Van Isacker

Big Data and Society Colloquium: Data Decisions & Border Governance, 3 October, 2024

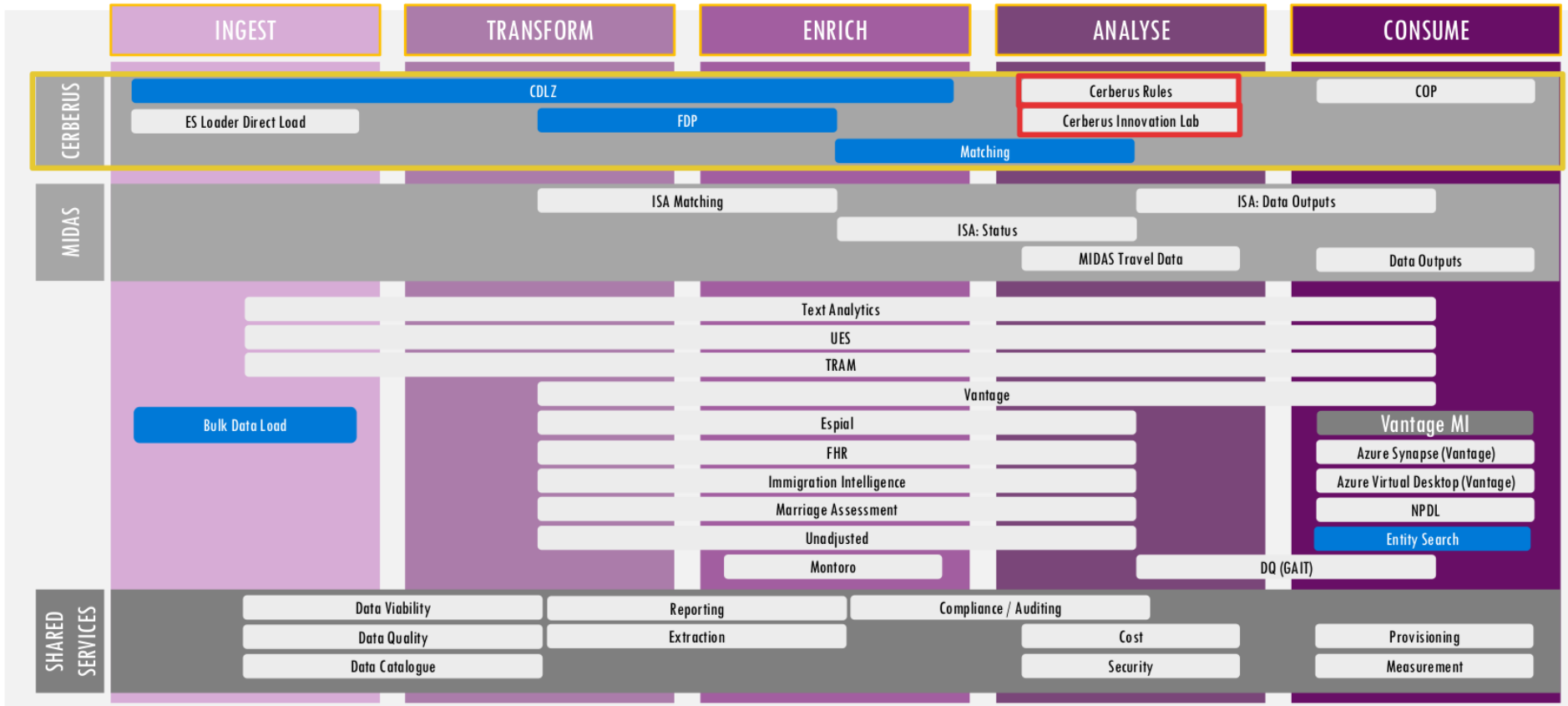
Data Futures Programme - Cerberus

- ‘Utilising data, analysis, and technology to create a holistic, cross-modal, threat agnostic targeting capability.’
- Built by BAE Systems Digital Intelligence (£38m)
- Managed by Home Office Data Services & Analytics

The ‘Cerberverse’

- Ingest
- Transform
- Enhance
- Match
- Risk
- Alert

To date DSA “Products” have been solution focussed …



Slide from Home Office Data Services & Analytics Market Engagement Event, 8 July, 2024.

History of UK border risking systems

1994 – Warnings Index watchlist

2003 – e-Borders launched

2004 – Semaphore pilot for API

2010 – e-Borders cancelled

2014 – Digital Services at the Border replace legacy systems

2019 – Reset DSaB – failed to deliver targeting systems

2020 – Data Futures for risking systems

'The tech vision of e-Borders showed a remarkable obduracy in the face of repeated failure.'

- Boswell and Besse (2023) 'The strange resilience of the UK e-borders programme', p. 23.

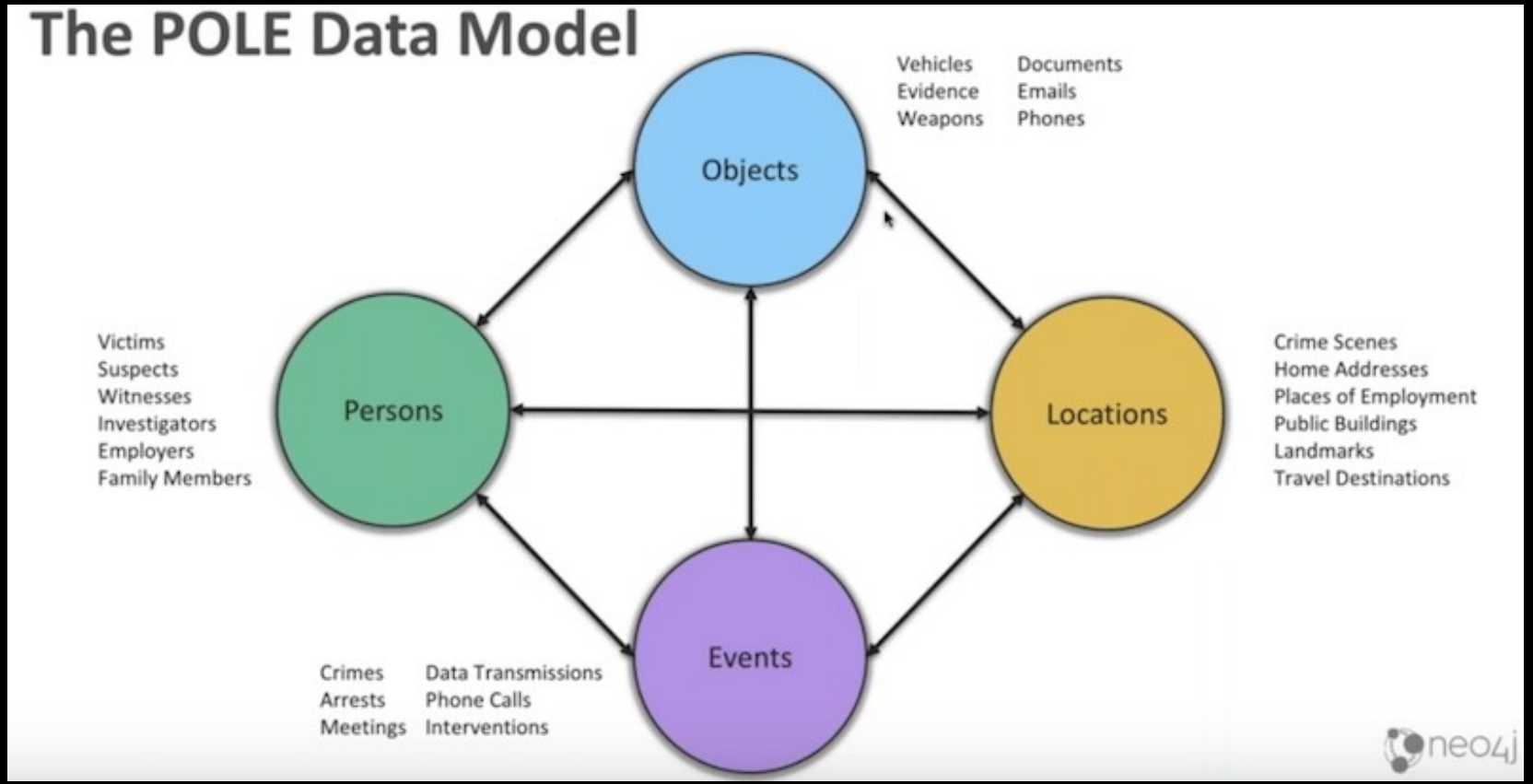
Multi-modal risking

- Single targeting system for ‘people, sea containers, fast parcels, post and freight.’
- Data from API/PNR, customs, safety & security declarations etc.

‘It’s bad people that bring bad stuff, after all. Dividing it on people and stuff, is that a good way of doing it? Maybe not.’

- Smith, T., former Director General of Border Force

POLE relational data structure



Add data extracted by the border – PNR, immigration history, freight consignments, etc. – and ‘contextual data’.

‘Use of the model helps us standardise the data, making it easier to share, access and match up with other data.’

- Gregory, A. DS&A Deputy Director

‘There’s a huge number of entities being matched… [with] Cerberus… we’re streaming data in and we’re matching millions of records per day.’

- Curphery, L. DS&A Programme Lead

Risking by association

‘cargo is not treated as a unitary object – a static “thing” whose very materiality speaks risk – but as a relational effect emerging out of associations between multiple data points and risk criteria.’

- Glouftsios, G. (2024) ‘Assess in advance, control where required’, p. 12.

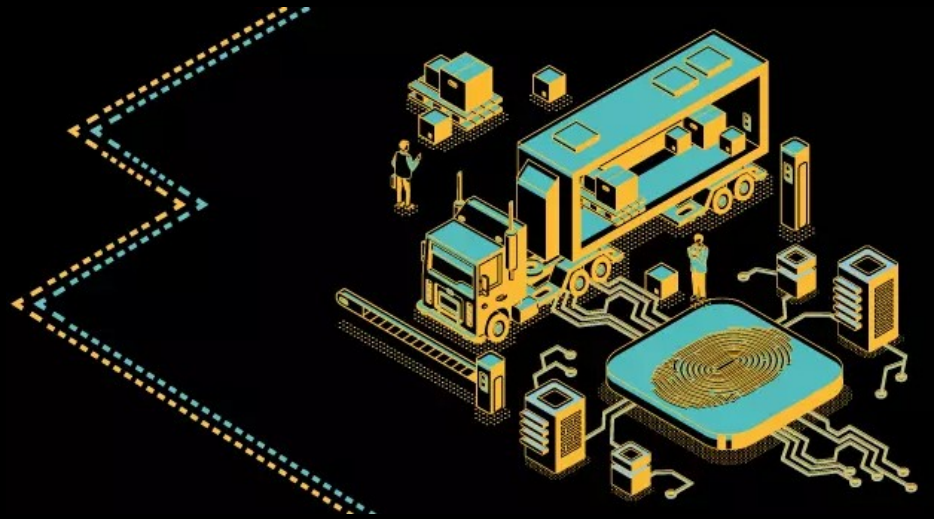
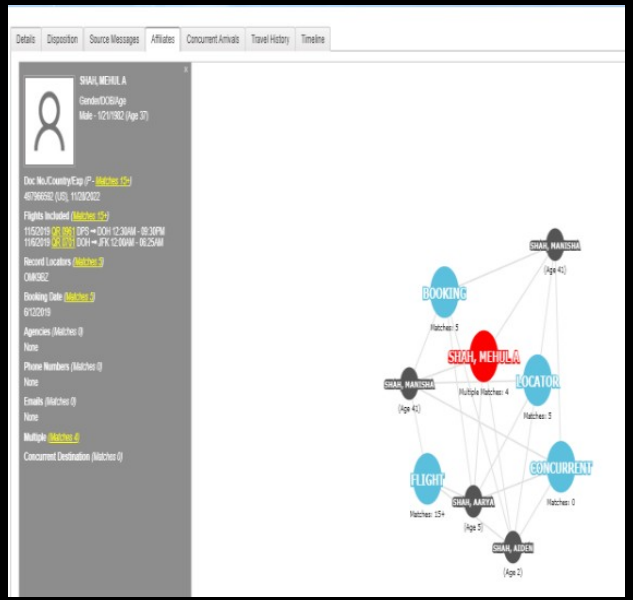
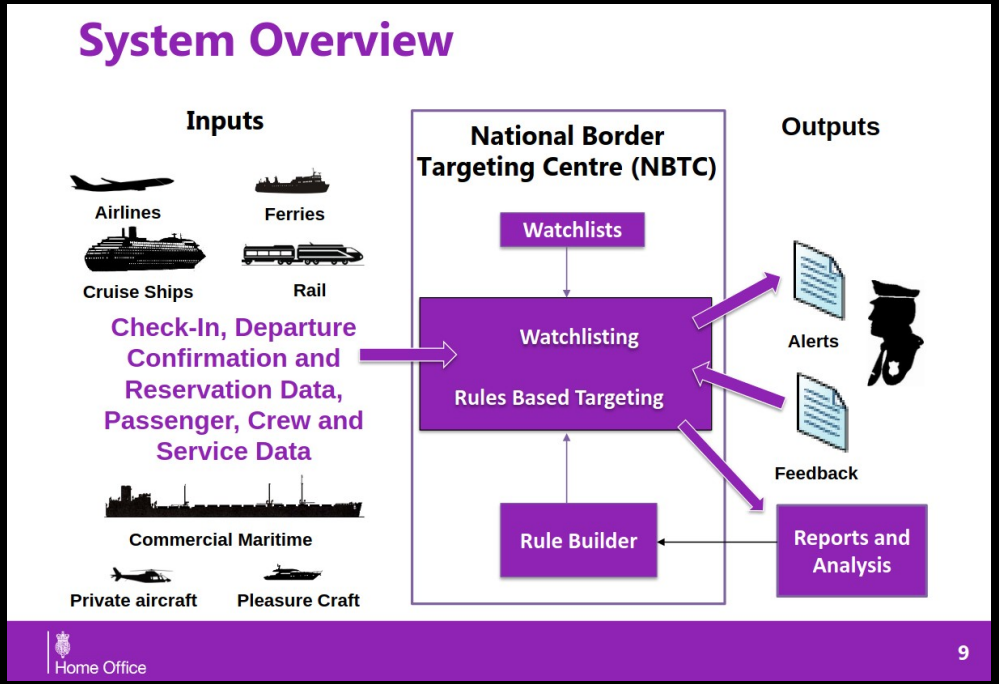


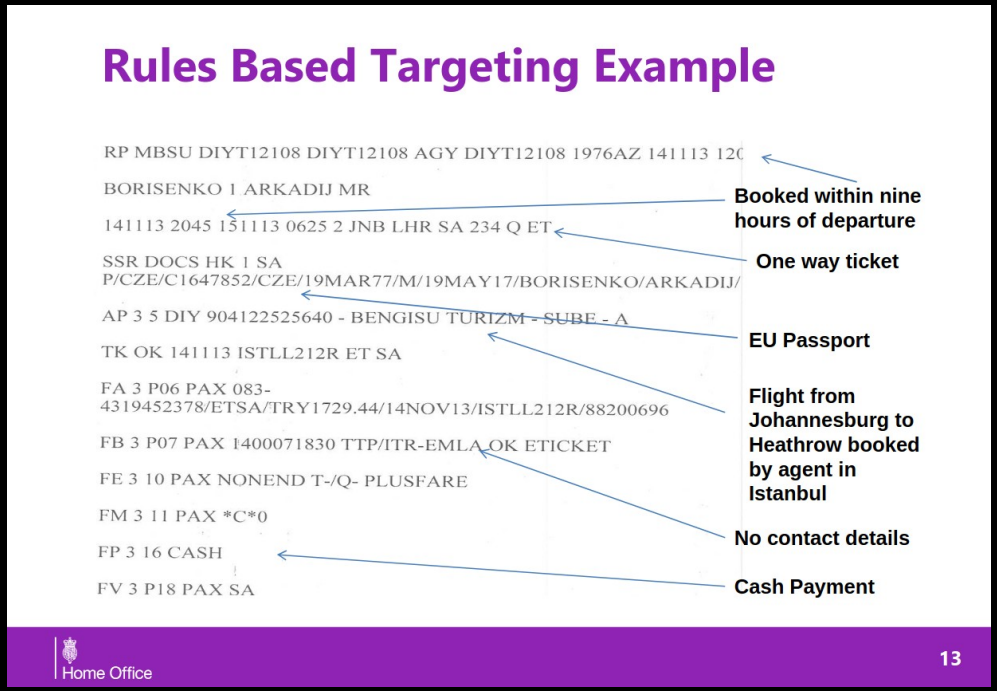
Image from BAE Systems Digital Intelligence, ‘Using Data to Secure the UK Border’



Slide from U.S. Customs and Border Protection, Automated Targeting System - Global

Rules-based targeting

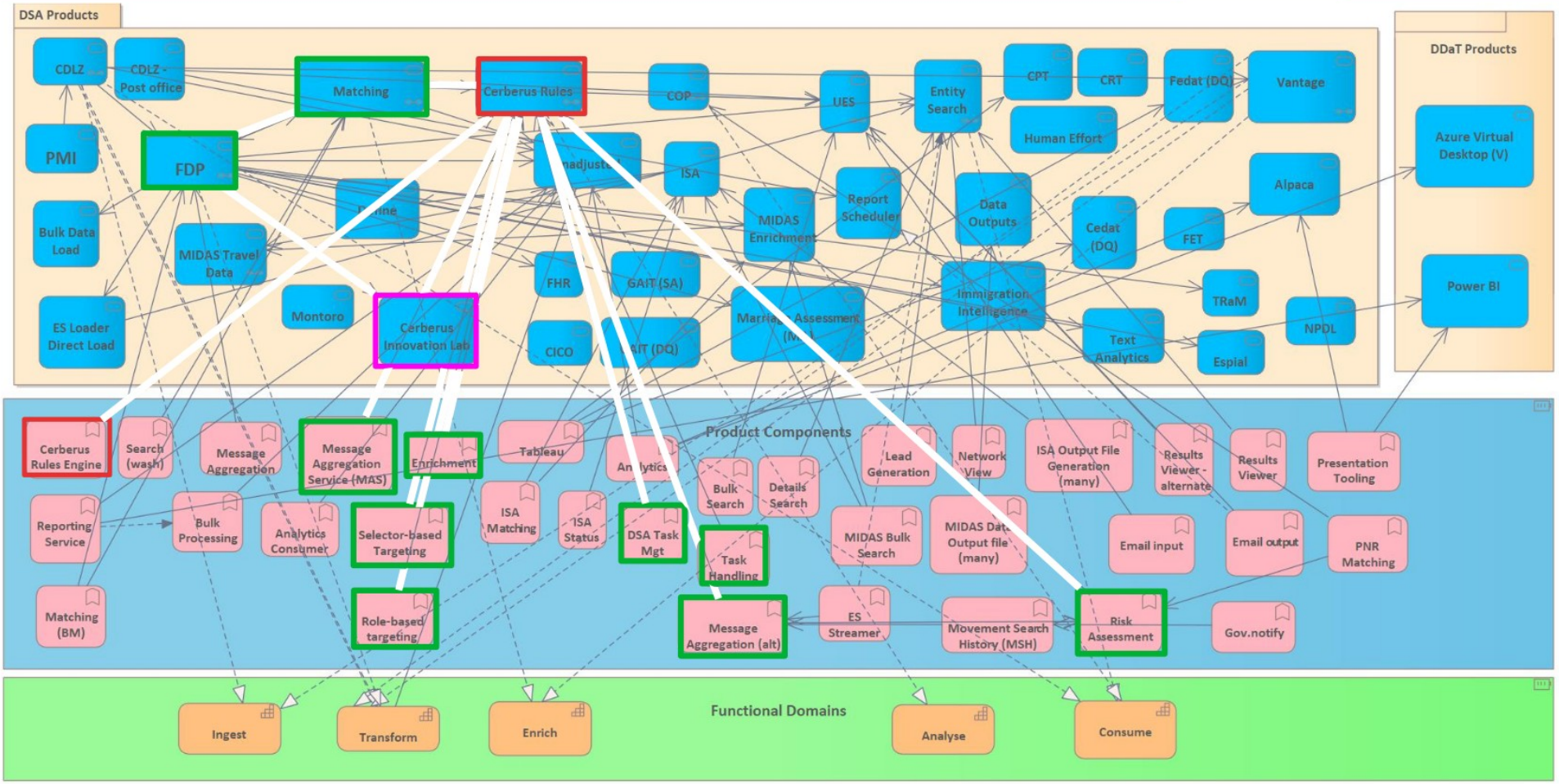




'Analysts set the rules based on patterns of observed suspicious behaviour. The [Border Risking and Targeting Capability] then uses these rules to analyse a wide range of data sets. Intelligence analysts can amend rules based on the outcomes of the intelligence.'

- Gregory, A. (2022) 'Analysing Data to Identify Risks at the UK Border.'

… resulting in a complex environment of functionality



Slide from Home Office Data Services & Analytics Market Engagement Event, 8 July, 2024.

Cerberus Innovation Lab

- Nascent machine-learning application?
- ‘unknowns unknowns’
- ‘Not predictive’ but pre-emptive and adaptive

Still from ‘Data to Secure the UK Border’ video, 00:37, BAE System Digital Intelligence.

- **Predictive Modeling**: Provide data-driven modeling to further modeling efforts incorporating state-of-the-art analytical tools, harnessing the power of machine learning, and focusing on analytical work. Work to discern anomalies that would escape the human eye as well as subtle common features that are highly predictive of future behavior.

From tender for U.S. CBP’s Automated Targeting System. ‘Prediction’ later omitted from the DPIA.

THE TIMES
THE SUNDAY TIMES

SPONSORED CONTENT

BAE SYSTEMS

Could AI-driven analytics transform Britain's border?

Huge passenger numbers require digital intelligence on a colossal scale to help minimise queues while also keeping us safe by monitoring threats in real time

October 2 2024, 12:01am

BAE Systems sponsored piece in the *The Times*, 2 October 2024

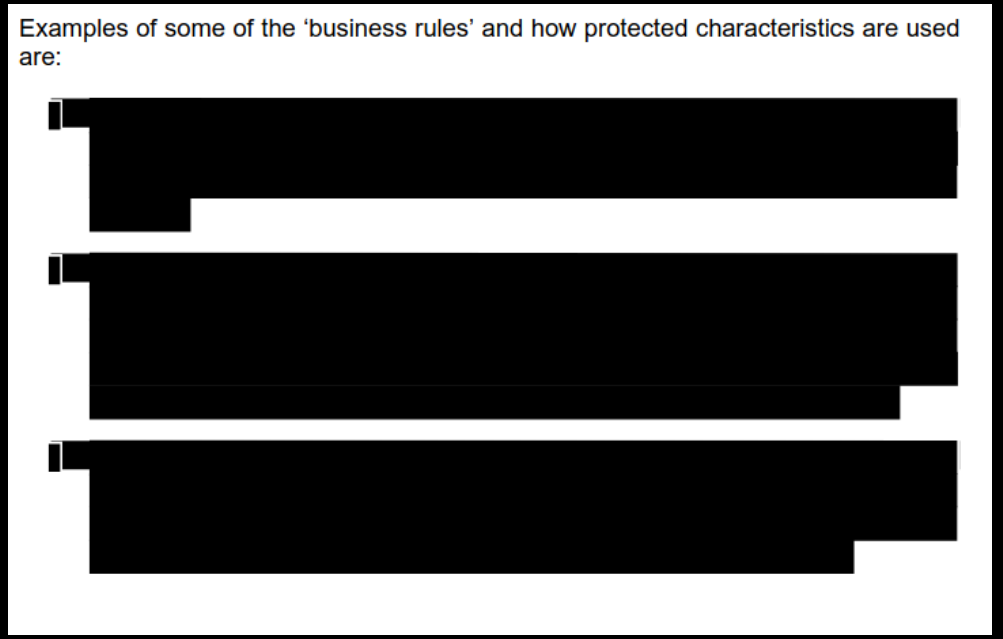
‘where rules-based algorithms sought to capture the relationship of variables within a dataset – often with human domain specialists crafting the rules – deep learning seeks to generate the rules from features that are not pre-programmed in advance.’

‘the border is able to experiment and iterate, to generate features from the examples to which it is exposed, and to be invoked against a person at any future moment’

- Amoore, L. (2023) ‘The Deep Border’, p. 4

Cerberus transforms risk

- Generates risky groups
- Opaque, ever mutating, models
- Analysis without use of ‘protected characteristics’, bias still hard-coded
- Can be easily ‘flexed’ to new political priorities


Examples of some of the ‘business rules’ and how protected characteristics are used are:

How to assess Cerberus' social risk?

- What more information can be found?
  - So far cobbled together from accounting officer reports, market engagement, and 'industry events'
  - No social transparency
- What 'contextual data' does Cerberus ingest?
- How is it being applied?
- What is planned for its future?