

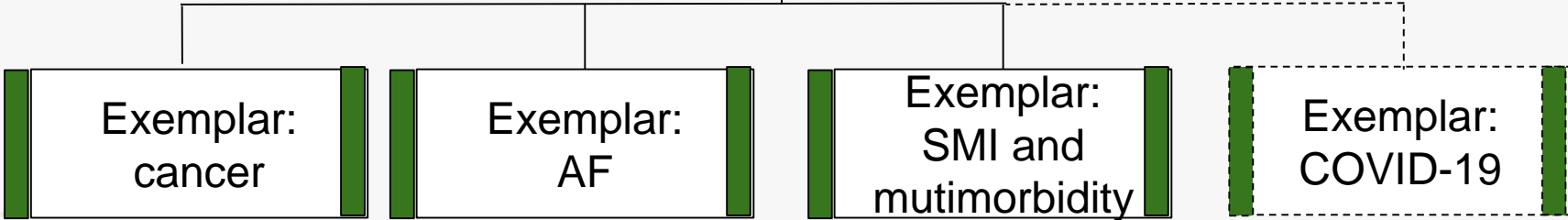
National Text Analytics Resource

Richard Dobson and Angus Roberts



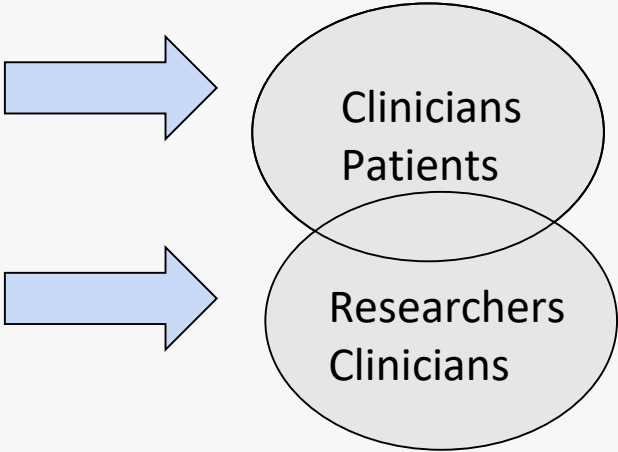
National Resource

NLP community

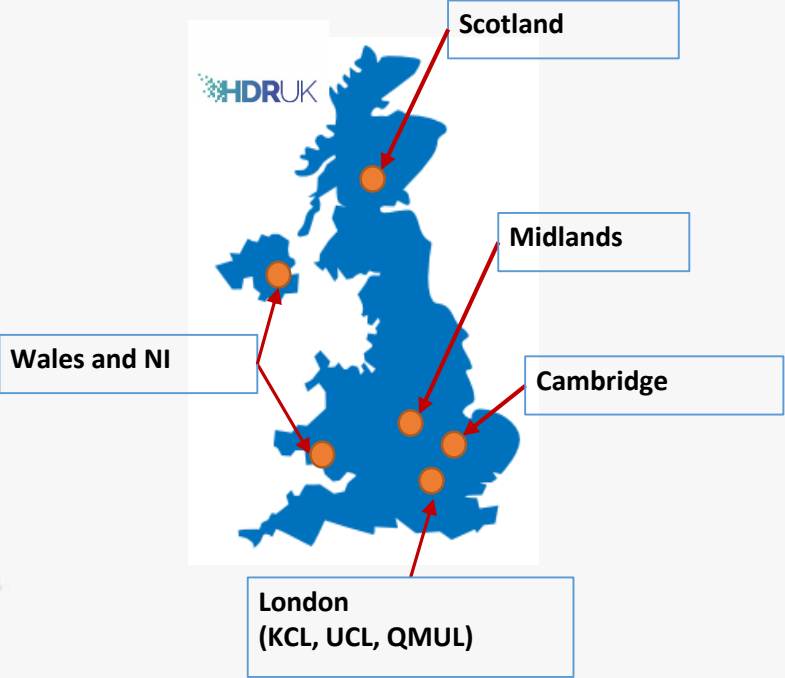


Infrastructure and data science integration

Training



Community



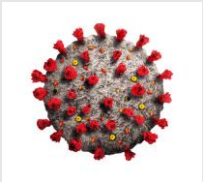
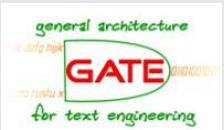
Linking within HDR UK



CogStack

- Better Care portfolio and NEWS2
- Gateway
- DATA-CAN Leeds
- The Human Phenome
- Shared leadership with Applied Analytics
- Synthetic data SIG
- In application: mental health hub

Reaching beyond HDR UK



International
NEWS2-CoV
collaboration



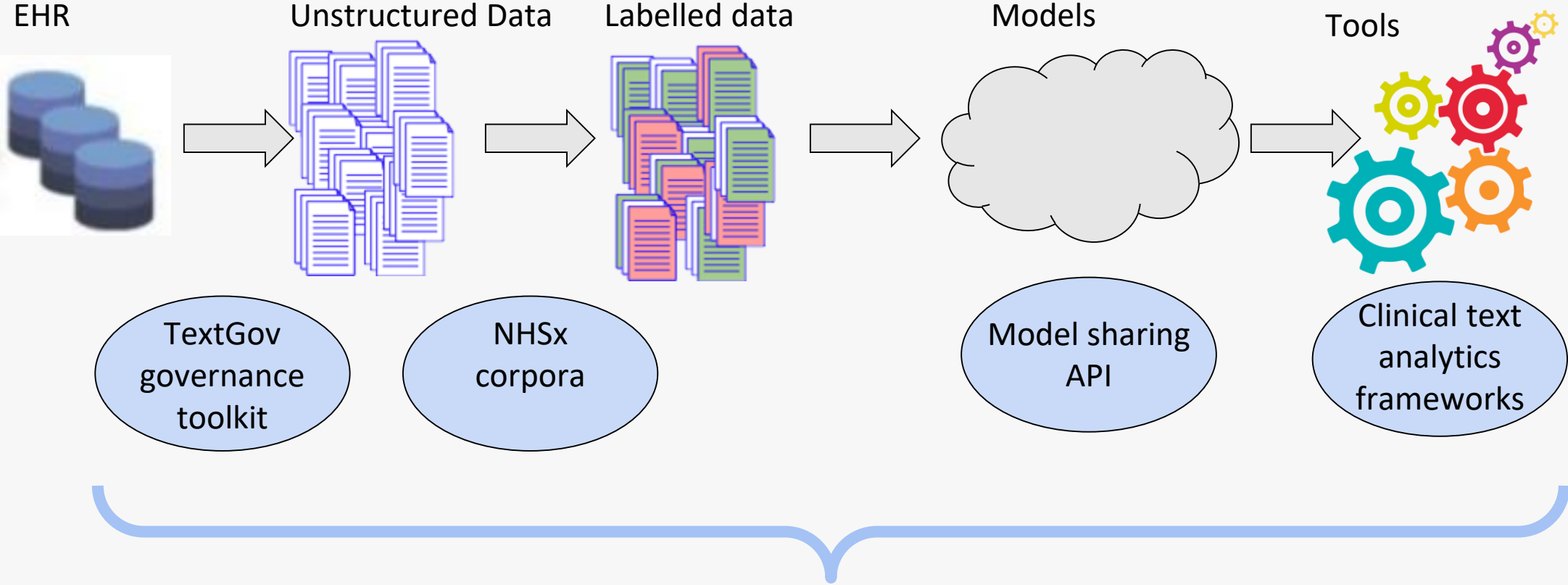
AI Statement
of Opportunities



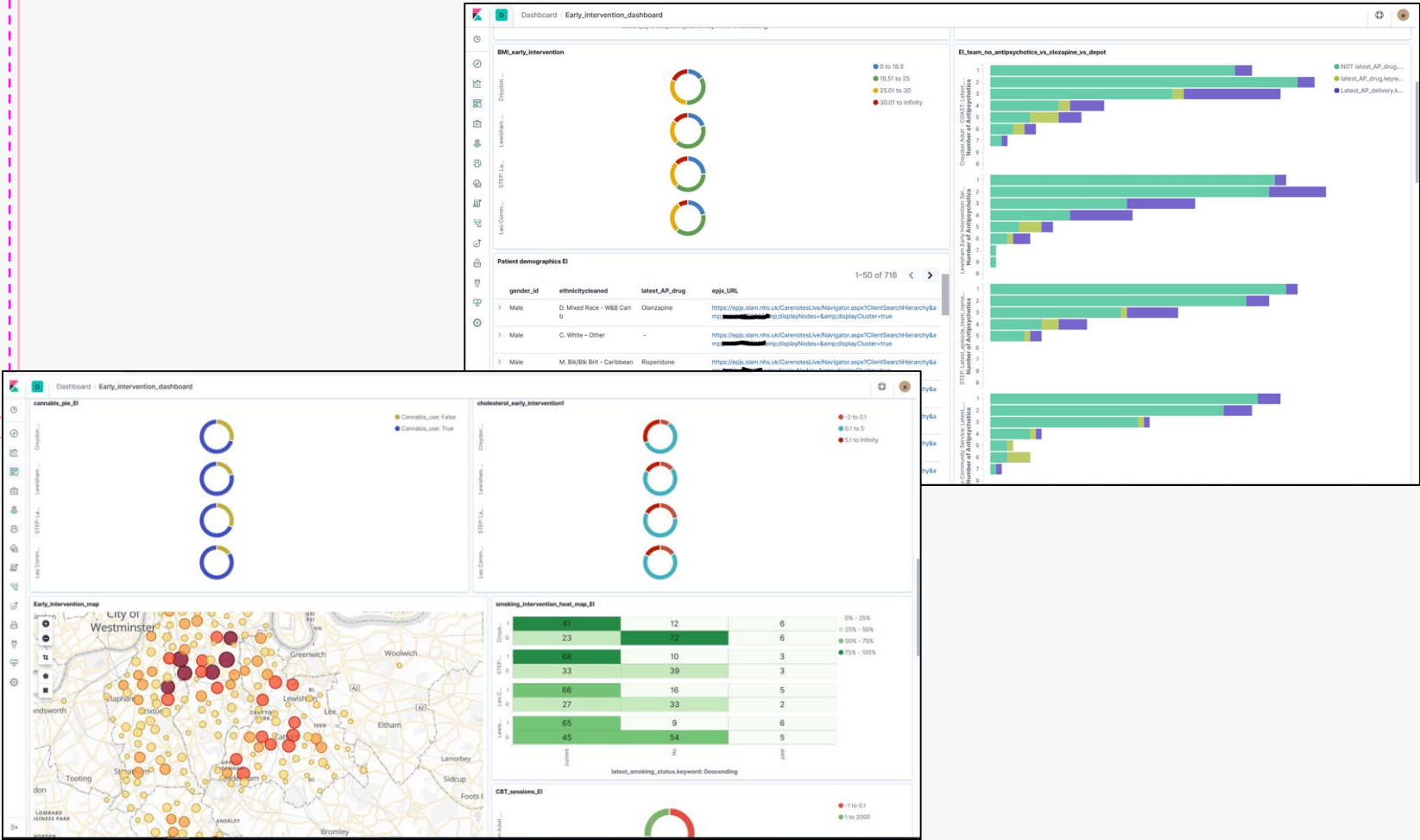
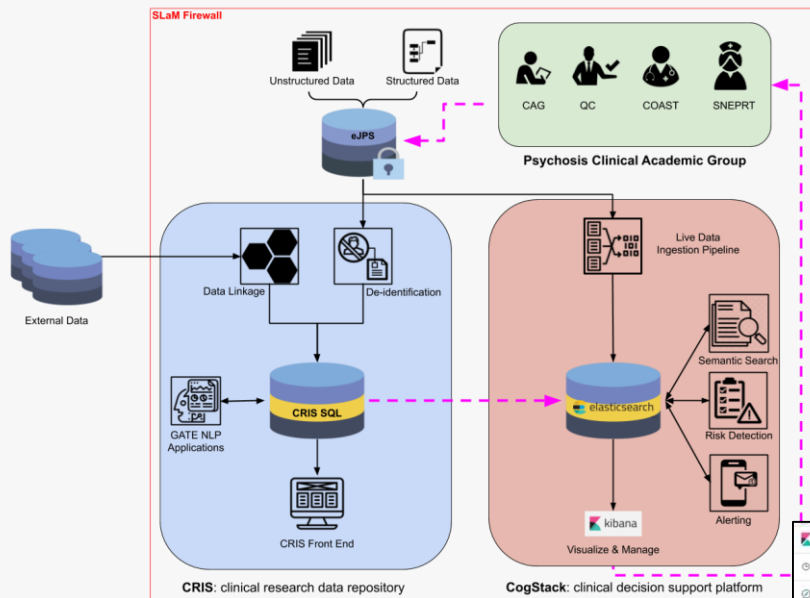
**CogStack
sites**

- SLaM
 - KCH
 - GSTT
 - MFT
 - UCLH
 - MerseyCare
- NHS** South London and Maudsley NHS Foundation Trust
NHS King's College Hospital NHS Foundation Trust

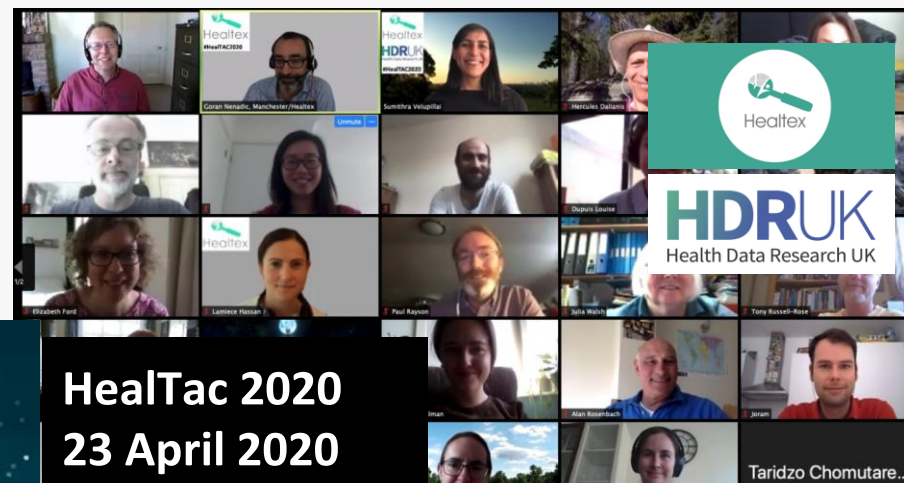
Sharing resources



Closing the loop: text analytics for direct clinical care



Training and dissemination




Portal, standards and APIs

Support HDR UK Gateway

Tools		
Manage your existing tools or add new ones		
Active (95) Pending approval (0) Rejected (0) Archived		
Last activity	Name	Author
8 March 2021 21:38	Freertext Matching Algorithm	HDR UK Text Analytics

Curate text analytics Github



Health Data Research UK Text

Resources and tools from Health Data Research UK Text

United Kingdom <https://hdruk-text.org/>

Repositories 2

Packages

People 5

Teams

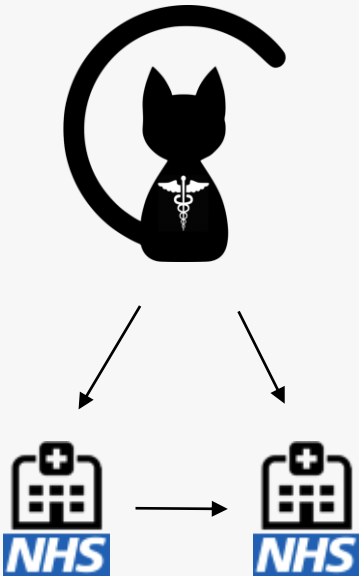
Type

Language

resources

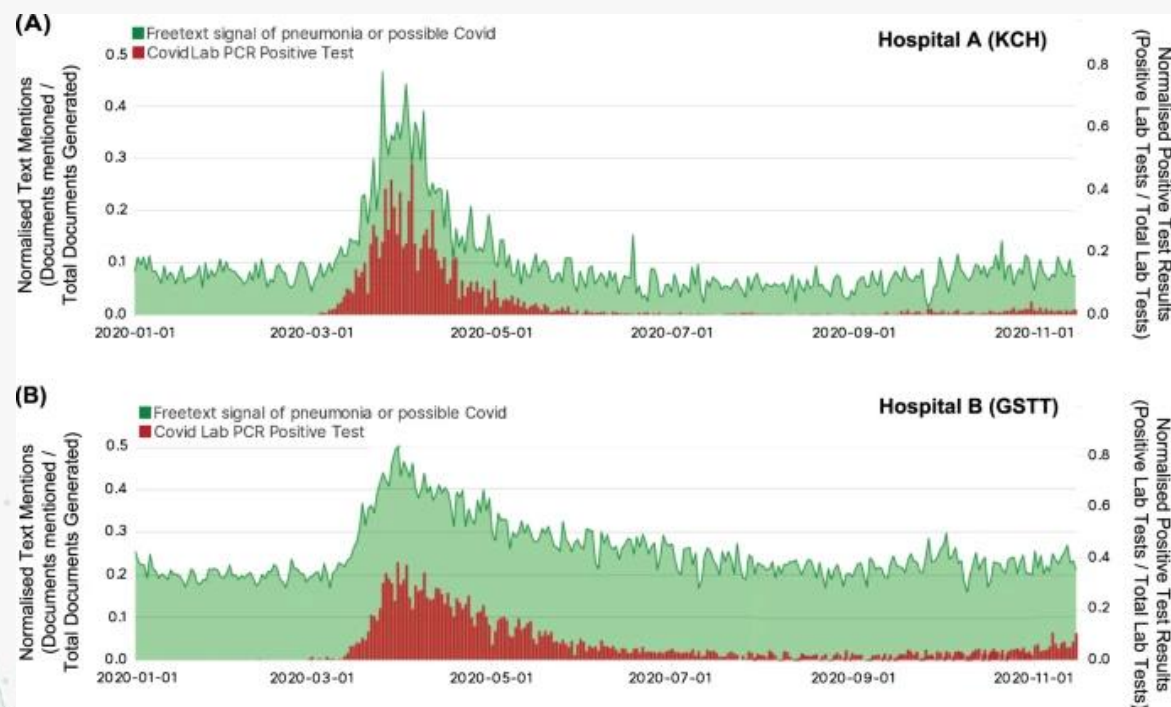
Applications and datasets for healthcare text analytics developed and shared by the HDR UK Text community.

Build model sharing API



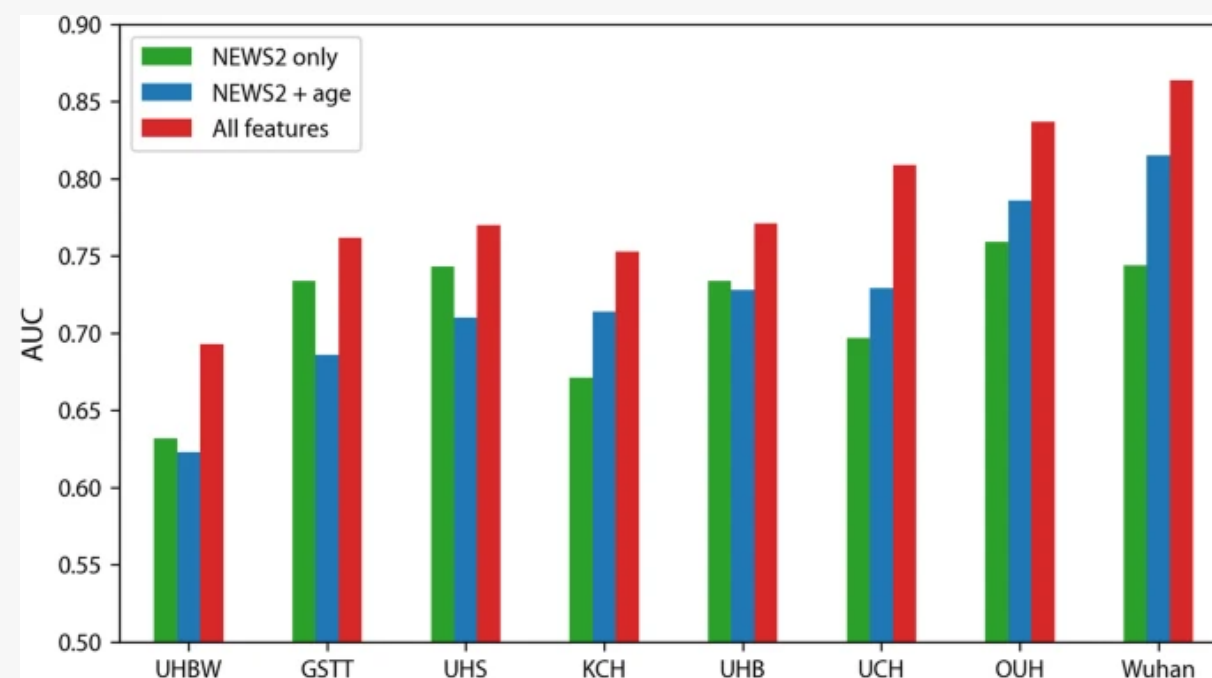
Exemplars: COVID-19

Real-time text feeds from EHRs



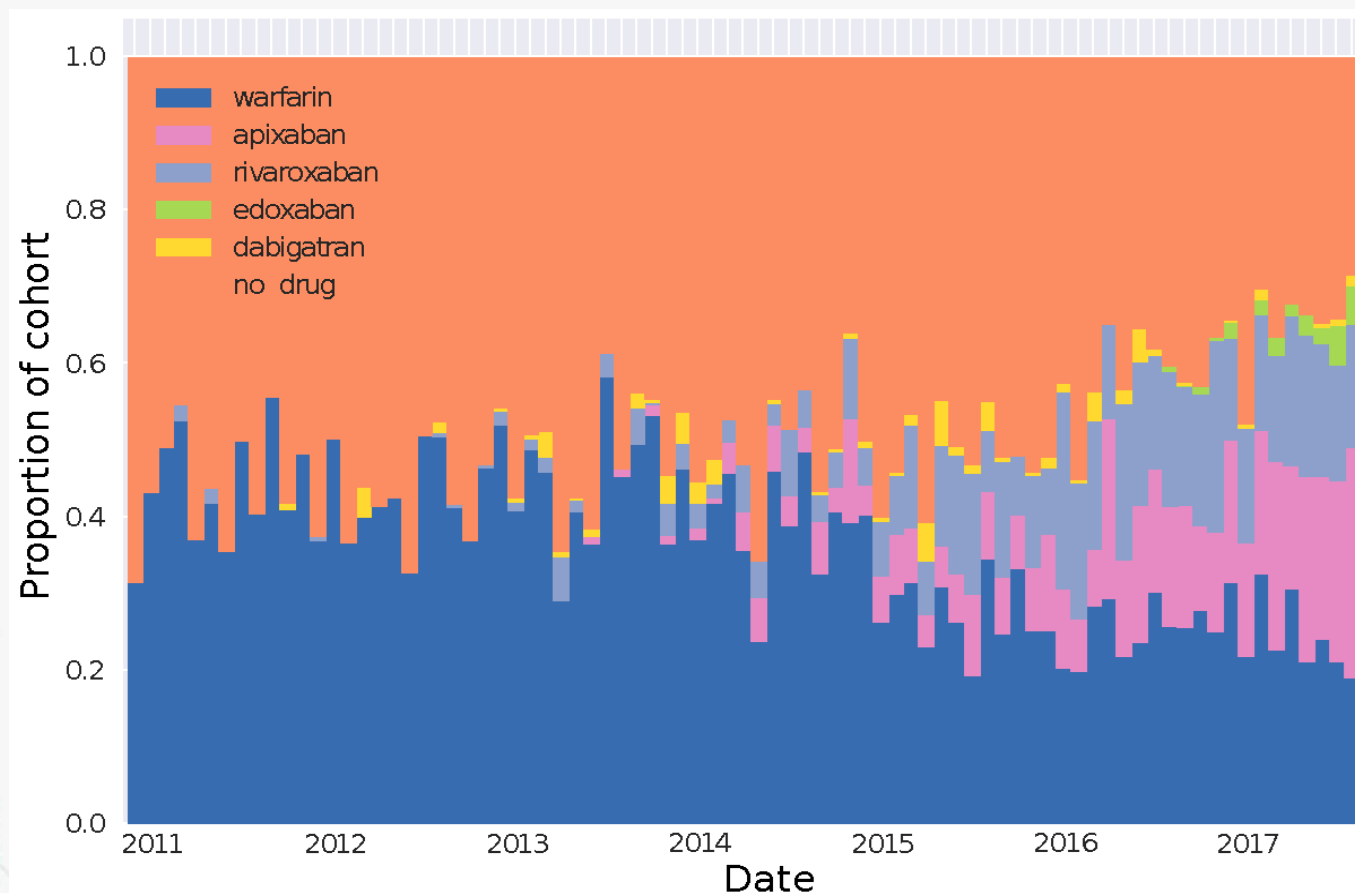
James T. H. Teo, Vlad Dinu, William Bernal, Phil Davidson, Vitaliy Oliynyk, Cormac Breen, Richard D. Barker & Richard J. B. Dobson
- <https://www.nature.com/articles/s41746-021-00406-7>

Evaluation and improvement of the National Early Warning Score (NEWS2)



Ewan Carr, Rebecca Bendayan, [many others...],
Richard J. B. Dobson -
<https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-020-01893-3>

Exemplars: Atrial Fibrillation

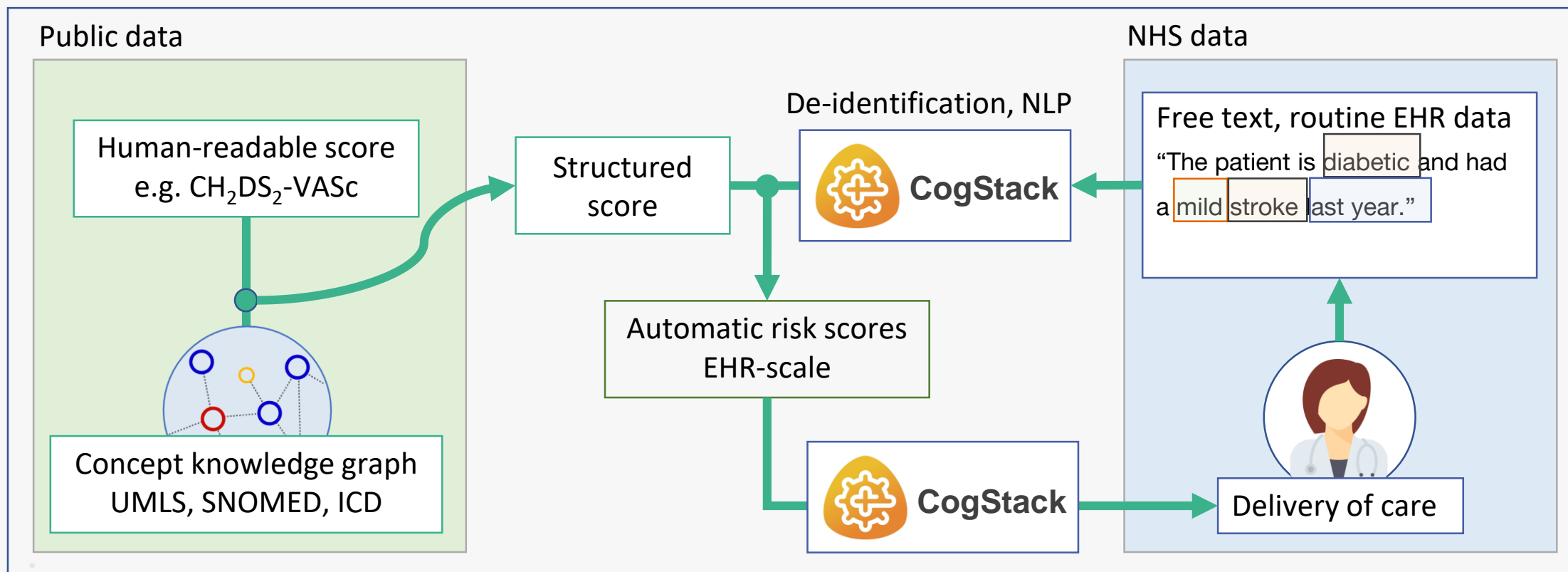


Why are anticoagulant medications not given to these at-risk patients?

Can routinely collected, unstructured EHR data help explain why and inform clinical decisions?

Daniel M. Bean, James Teo, Honghan Wu, Ricardo Oliveira, Raj Patel, Rebecca Bendayan, Ajay M. Shah, Richard J. B. Dobson, Paul A. Scott - <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0225625>

Exemplars: Atrial Fibrillation - developing a pipeline for automated risk scores across NHS Trusts



Prototype at KCH, validation and further development at UCLH, SLaM and UHB

Summary

- Huge opportunity to use text analytics to improve research, hospital service planning and clinical decision making
- Challenges remain in navigating regulatory access to free text EHR data and deploying consistent, replicable methods
- Progress can be made by
 - starting small
 - reaching out to the wider data science / text analytics / clinical community
 - explicitly focusing on not reinventing the wheel
 - persistence!