Influence of scientific research on policy: a funder's perspective

Chonnettia Jones, Ph.D.
Director of Insight & Analysis

Our origins

Founder Sir Henry Wellcome was an entrepreneur, collector and philanthropist.

On his death in 1936, his will established a charity for "the advancement of medical and scientific research to improve mankind's wellbeing".



Wellcome Success Framework



The framework sets out **9 shared ambitions** that express '*what success looks like*' to Wellcome.

We aspire to **improve health for everyone by helping great ideas to thrive**.

We will achieve this by:

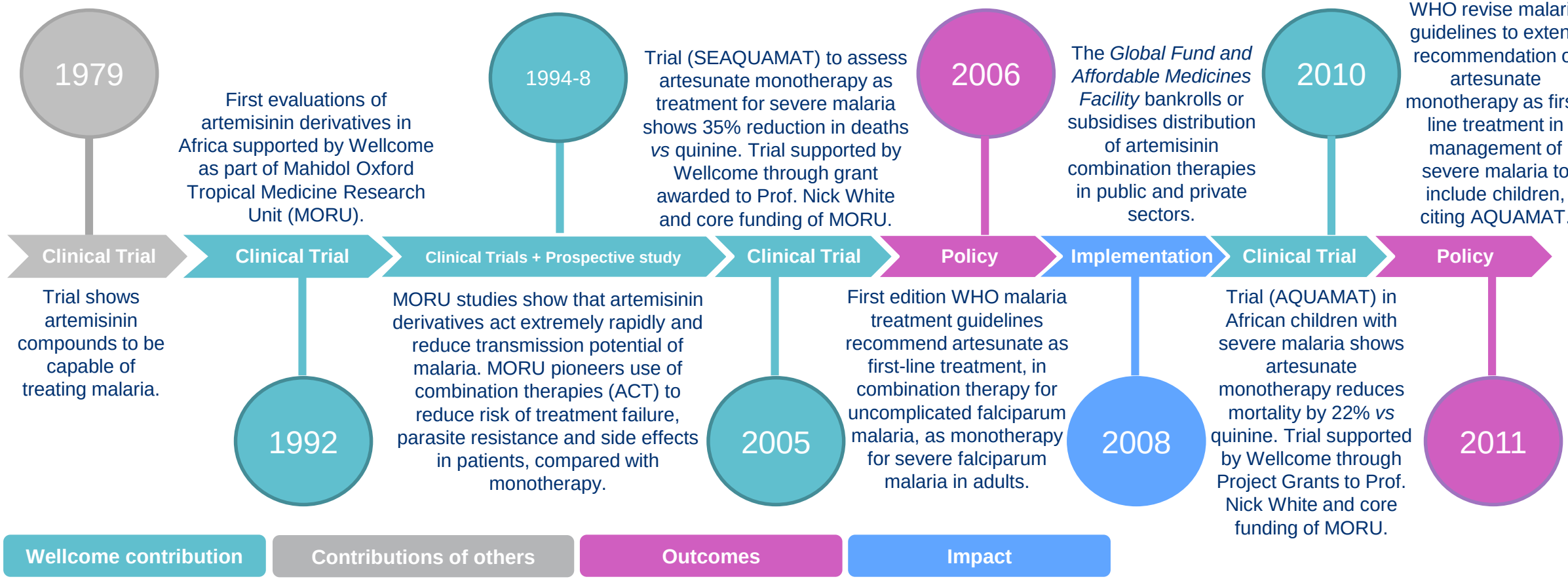
- Maximising the potential of **research** to improve health.
- Delivering **innovations** that prevent or treat health problems.
- Engaging **society** to shape choices that lead to better health.

We hold ourselves **accountable to society** for delivering Wellcome's mission, while using our **independence for public benefit**.

Case Study 1:

Treatment for Malaria

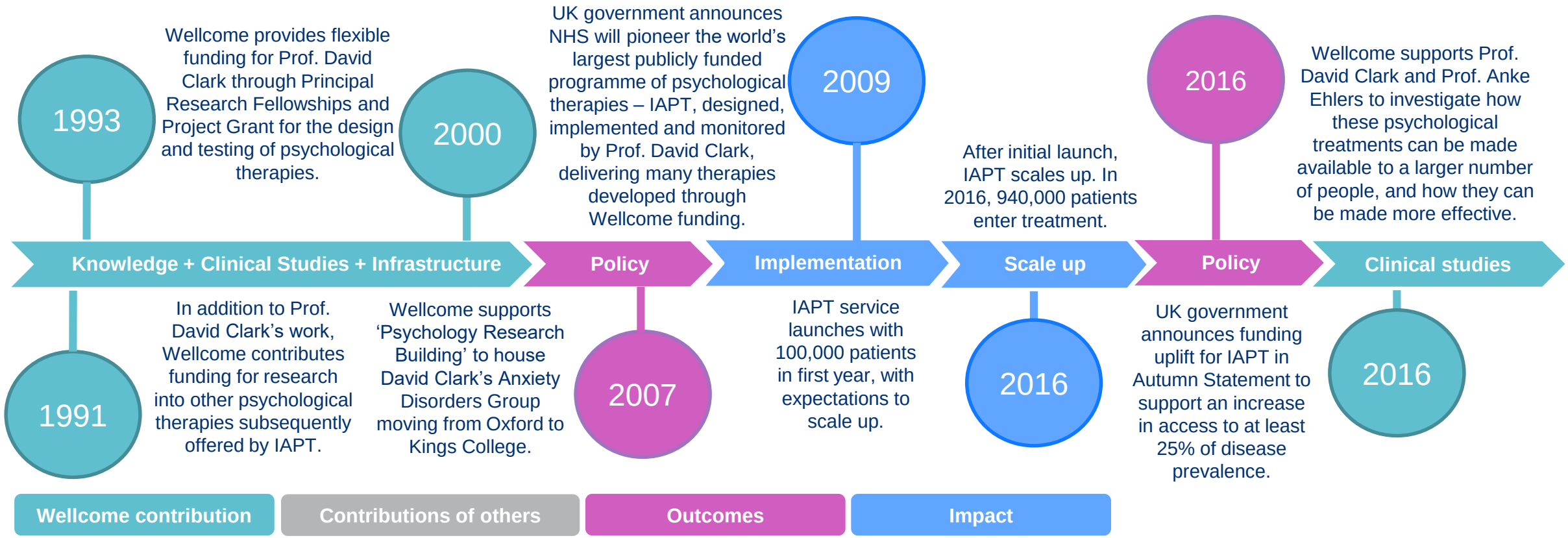
Artesunate-based Malaria Mono- and Combination Therapy



Case Study 2:

Psychological therapies

Improving Access to Psychological Therapies



What have we learned?

- Wellcome's contributions to health have primarily been the funding of research and early development of health interventions, including drugs, devices, diagnostics, and cognitive behavioural therapies.

- The average time that has elapsed between first research findings and improved health outcomes has been 10-15 years. Greater focus on the implementation of interventions reduced the time to health impact.

- Wellcome has played a 'passive' funder role in supporting research that could have policy relevance – that is, while Wellcome has provided funding, the responsibility for pushing for policy reform required to improve health outcomes has fallen to the investigators.

What could we do differently?

- Diversify Wellcome's contributions to health-related research beyond just funding – including convening and brokering multi-lateral conversations, advocating for reform, and actively seeking to influence policy.

- Develop an implementation strategy that involves early policy engagement alongside health-related research to facilitate the rapid uptake of research into policy and practice and ensure health interventions can reach the people who could benefit.



Data science methods to assess the influence of scientific research on policy

Dr. Danil Mikhailov
Founder and Head of Wellcome Data Labs

Contents

Context

Team Structure & Focus

Part I: Using ML to understand Reach and Impact

- Problem Definition
- Methodology
- Results
- Issues
- Future Development

Part II: Practical Ethics in Product Development

- Our Ambition
- What have we achieved?
- The challenges
- Acknowledgements

Context

Contextualising this research

- This report is intended as a meta analysis of the approach Wellcome Trust has taken to apply machine learning (ML) to the problem of understanding the reach and influence of Wellcome funded research in Policy.
- Emphasis is on analysing the process, including challenges and issues, rather than publishing the results of that process.
- As such the results are presented unvarnished, following a pilot run of a set of ML tools that are collectively known as the Wellcome Reach tool.
- The results purposefully include errors and duplicates to enable analysis of the cause of the issues and discussion of possible ethical challenges.
- The intention is to iterate the development of the Reach tool until we get the desired accuracy levels and then publish the results.

Team

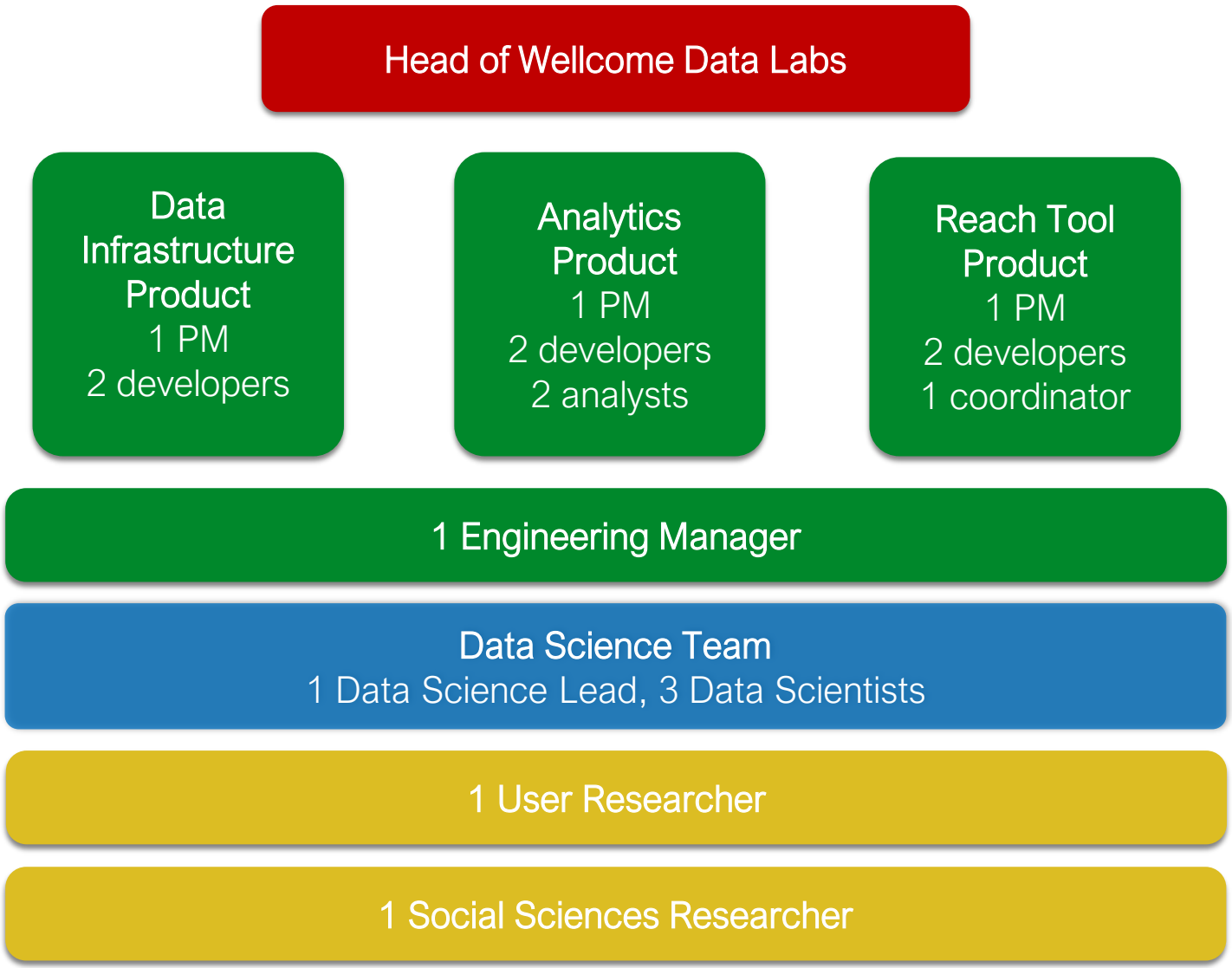
Who are Wellcome Data Labs?

We are an interdisciplinary team of social scientists, technologists and data scientists founded as an internal start-up in Wellcome in 2017.

We focus on:

- Working on funding data to improve decision making in Wellcome and other funders
- Supporting Wellcome's Priority Areas on their data and technology needs
- Embedding ethics, social and behavioural analysis into technology and data science work

Data Labs Product Structure



Problem

Wellcome Success Framework



Ambition 7.

Health is improved through changes in policy and practices.

Ambition 7 - Indicators

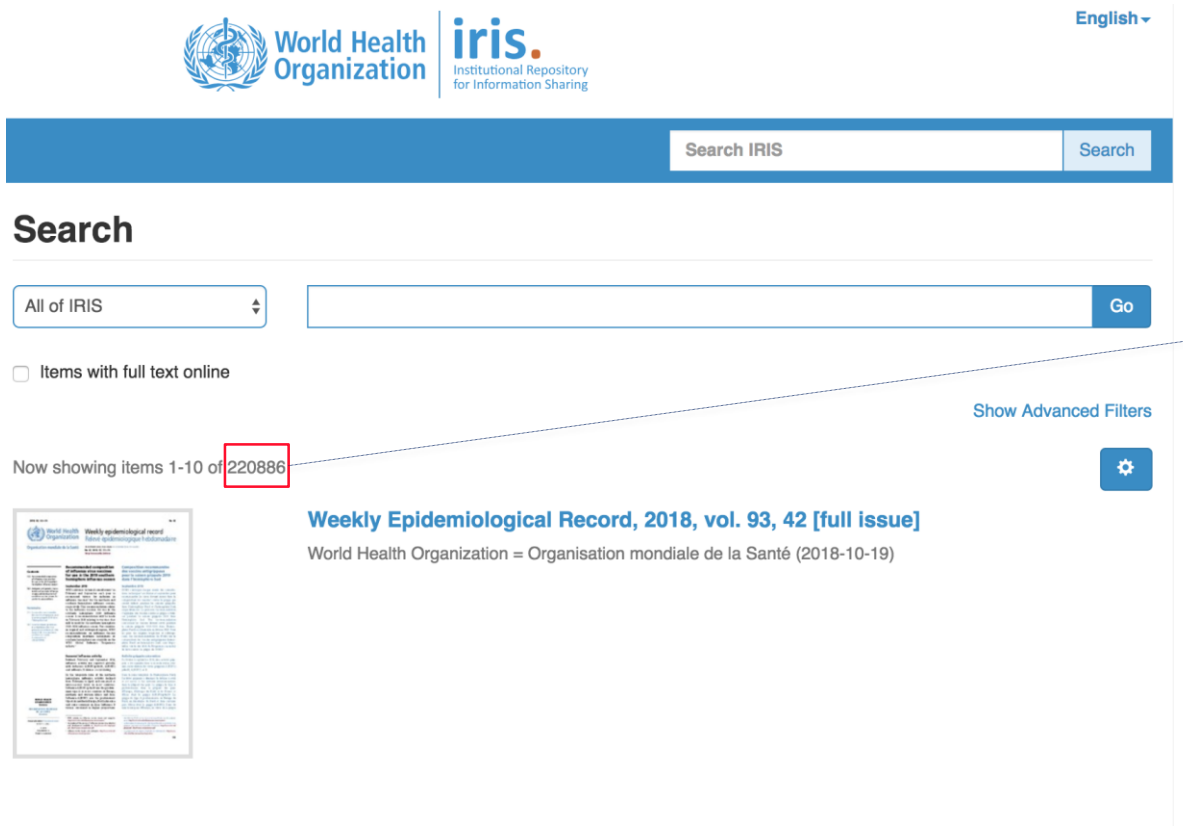
a. Policy and practice are informed by research and researchers

1. Number of policy cases that reference WT funded research
2. …

b. Decision makers take up Wellcome's position on policy and practice

1. Number ..

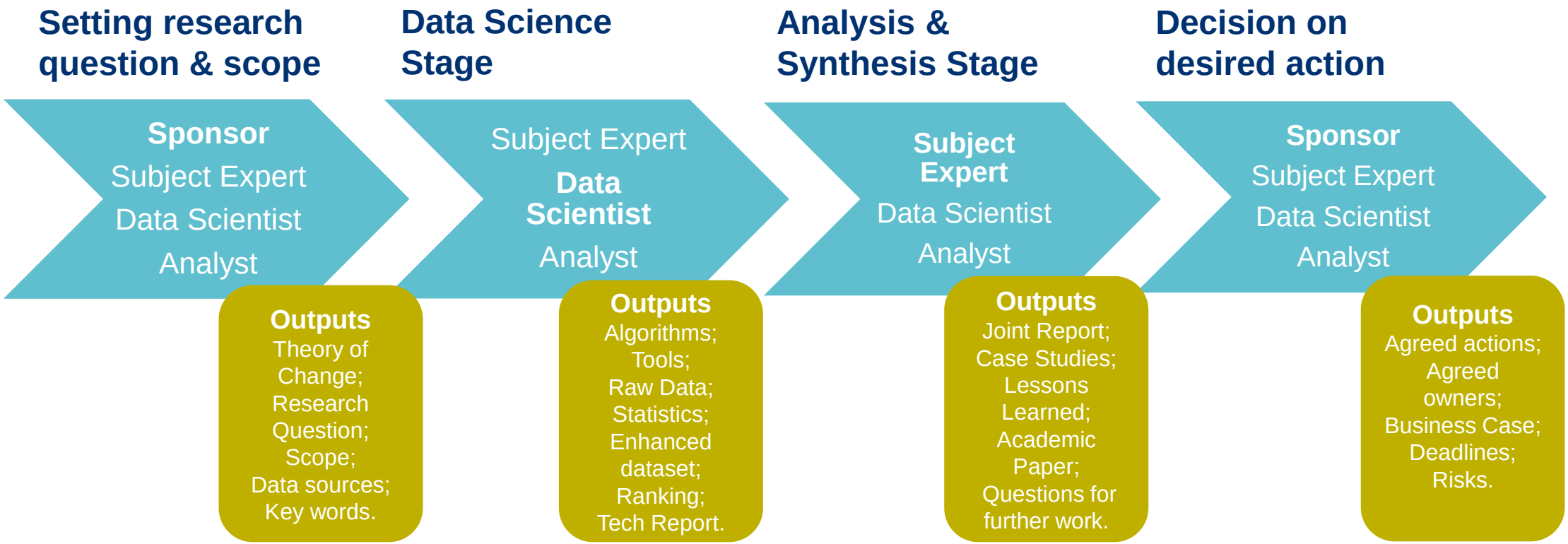
Finding policy cases: scale justifies use of ML & automation



220,886 documents on WHO website

Methodology

Data Labs approach to collaboration



Data Science Method

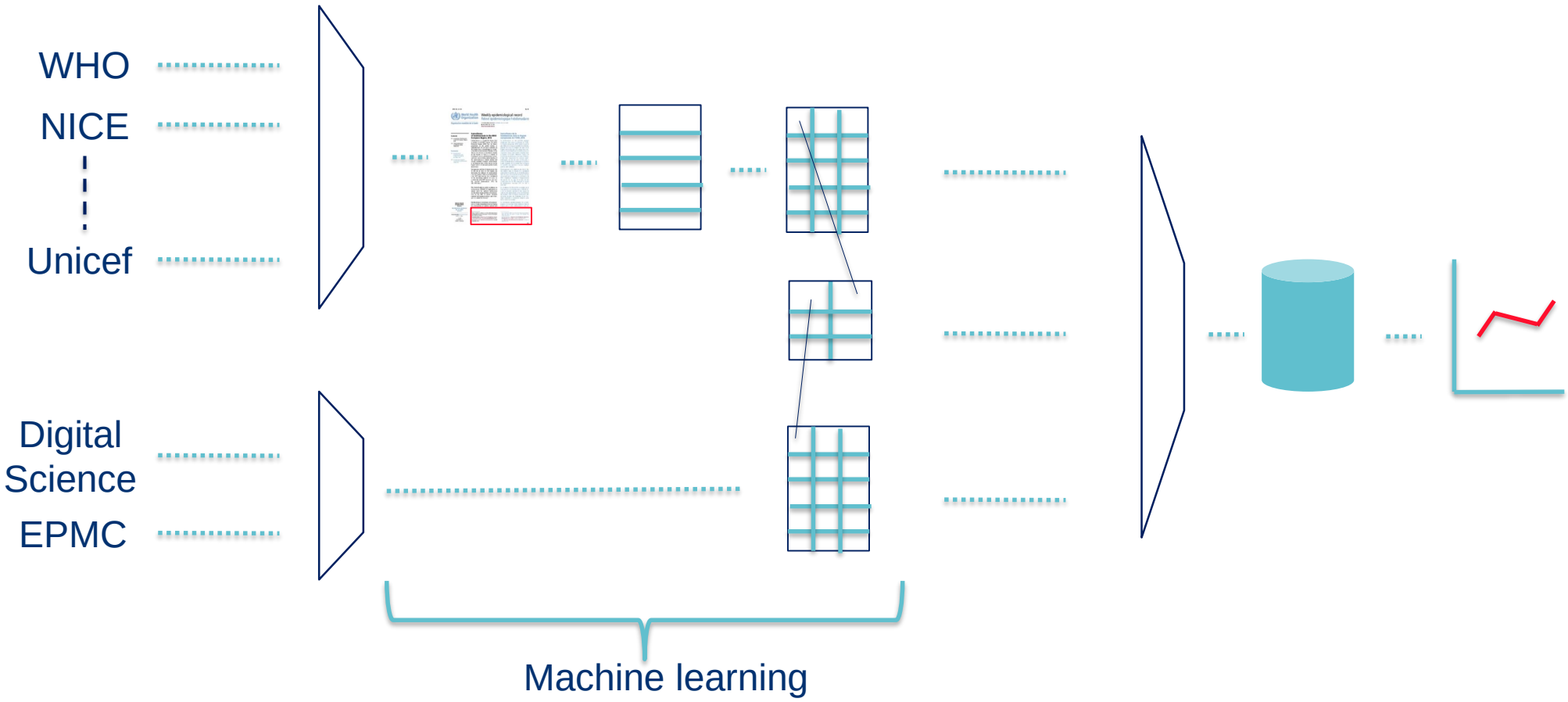
At this stage in the process we take the Theory of Change and other elements agreed with the analysts to begin to create Machine Learning tools and the data pipelines needed to productionise the quantitative analysis.

To understand the reach of Wellcome funded research into policy we created software tools to automatically scrape policy documents from WHO, Unicef, NICE and MSF's publicly accessible online portals.

Then we developed ML tools using Naïve Bayes to **Find** citation sections in the doc, **Split** them into individual citations, **Parse** them into constituent buckets of "title", "author", "date" etc, and **Match** them with a structured dataset of Wellcome funded publications derived from Digital Science's Dimensions tool.

Please see our tech documentation on GitHub for more details and to review code

Data Science Pipeline



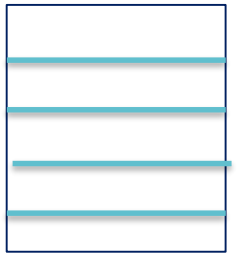
ML accuracy by stage

Find



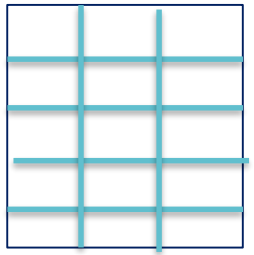
85%

Split



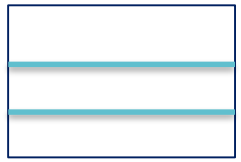
57%

Parse



87%

Match



96%

Going from Reach to Influence

The data science work produced a signal which indicates a potential connection between a Wellcome funded publication and a policy document that we define as “reach”.

To get from the weaker claim of “reach” to a stronger claim of “influence” we do a number of follow up steps of more qualitative analysis to validate the claim.

Validation of case studies makes use of multiple methods to clarify the validity of the claimed influence, such as: timeline review, documentation review, consideration of alternative explanations for the observed influence and other independent sources of information.

Preliminary results of Wellcome Reach Tool 1st trial run

Citations of Wellcome funded publications by policy source

Organisation	Documents	WT References*
WHO	7,484	867
NICE	289	5
Unicef	189	33
MSF	69	1

*** Results include errors and duplicates, see Issues Section**

Issues

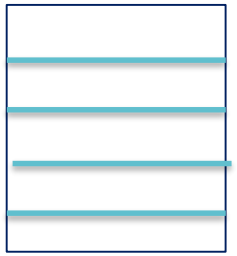
Split function causes accuracy issues

Find



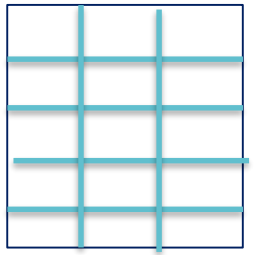
85%

Split



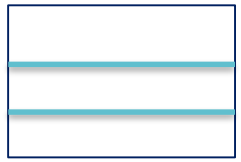
57%

Parse



87%

Match



96%

Example: 10 ‘most cited’ papers

The following table shows the ten Wellcome funded papers with the highest number of citations in the Policy document corpus analysed with the Reach Tool. It was initially assumed that the highest cited papers would be best candidates for follow on evaluation.

But:

- **8 out of 10 had duplicate issues** (the most striking example is the publication ID 1054751447 – it is supposed to have 56 policy citations and 55 were duplicates).
- **4 out of 10 are false matches** (the policy docs do not seem to cite the publication)

case study ID	#Policyci	DOI	Publication ID1	Right doi	Title	policy tool - issues
1	10	10.1016/s0140-6736(15)6	1021970111	10.1016/s0140-6736(15)6	Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990â€“2013: a systematic analysis for the Global Burden of Disease Study 2013	6 duplicates (1 policy doc repeated 6 times)
2	14	10.1016/s0140-6736(15)0	1011656620	10.1016/s0140-6736(15)0	Attention deficit hyperactivity disorder	for this publication the tool picks up exactly the same policy docs as the different publication below with the same tile (Pub ID 1021807036) and none of the policy docs cites this specific publication. (1 policy doc is repeated twice)
3	9	10.1016/s0140-6736(16)3	1004704916	10.1016/s0140-6736(16)3	Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19Â·2 million participants	1 policy doc repeated 4 times and 1 other policy docs repeated twice
4	8	10.1016/s1473-3099(10)7	1010822696	10.1016/s1473-3099(10)7	Priorities for tuberculosis research: a systematic review	1 policy doc repeated 8 times and it does not cite this paper (title of the policy doc very similar to the title of the publication)
5	8	10.1371/journal.pmed.100	1003600333	10.1371/journal.pmed.100	Multidrug Resistant Pulmonary Tuberculosis Treatment Regimens and Patient Outcomes: An Individual Patient Data Meta-analysis of 9,153 Patients	ok
6	8	10.7448/ias.15.2.17383	1031727257	10.7448/ias.15.2.17383	Quantifying and addressing losses along the continuum of care for people living with HIV infection in subâ€Saharan Africa: a systematic review	1 policy docs repeated twice
7	19	10.1016/s0140-6736(12)6	1037992837	10.1016/s0140-6736(12)6	Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990â€“2010: a systematic analysis for the Global Burden of Disease Study 2010	2 policy docs repeated twice
8	8	10.1371/journal.pmed.100	1011501074	10.1371/journal.pmed.100	Global Estimates of Syphilis in Pregnancy and Associated Adverse Outcomes: Analysis of Multinational Antenatal Surveillance Data	OK
9	14	10.1385/nmm:8:4:461	1021807036	10.1385/nmm:8:4:461	Attention Deficit Hyperactivity Disorder	The publication was not actually cited in any of the policy documents the tool picked up (1 policy doc is repeated twice)
10	56	10.1016/j.trstmh.2007.03.	1054751447	10.1016/j.trstmh.2007.03.(	The treatment of severe malaria	55 same policy doc repeated +1 and they both do not cite the paper

Example: 10 most cited papers (continued)

- It is clear from the results that there is a problem with initial hypothesis that highest-cited papers in the Reach results would make better targets for further evaluation to build evidence of influence. Namely, that many of the highest-cited papers may actually be those with most duplication and accuracy issues. This will require further investigation.
- Nevertheless the Reach tool is finding a significant number of real targets for further evaluation, which the analysts able to confirm 6 of the top 10 highest-cited papers as verified cases of “reach” and suitable for qualitative evaluation.
- The evaluation analysis is instrumental in catching these errors and feeding back into the design of the machine learning method to improve it.

Future

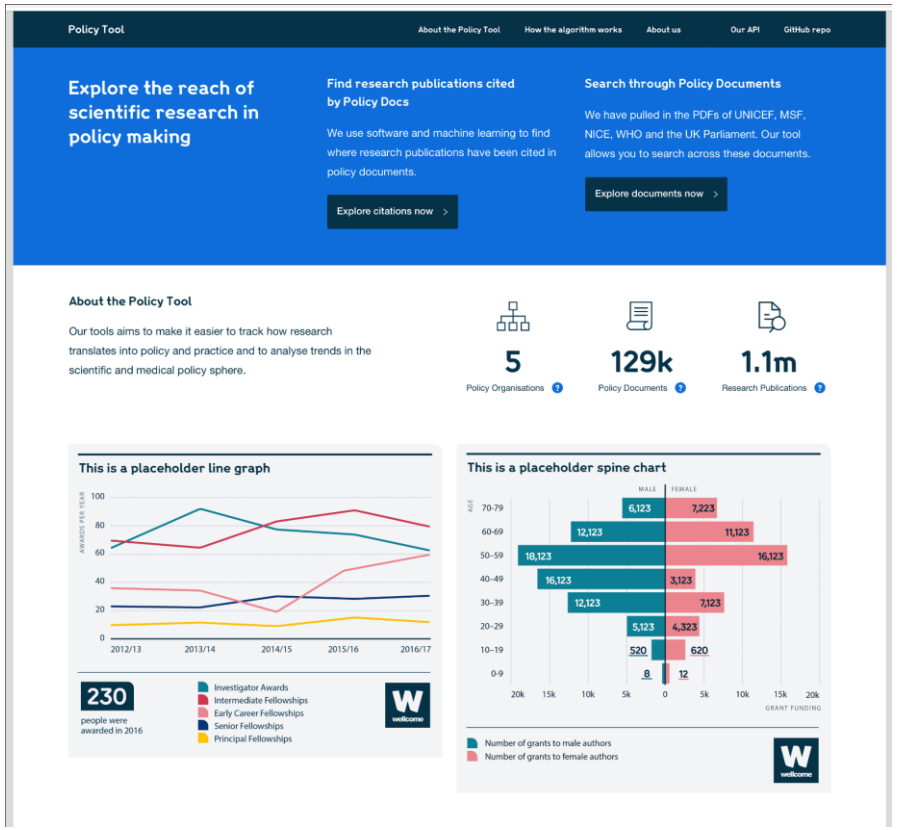
Improvements to ML

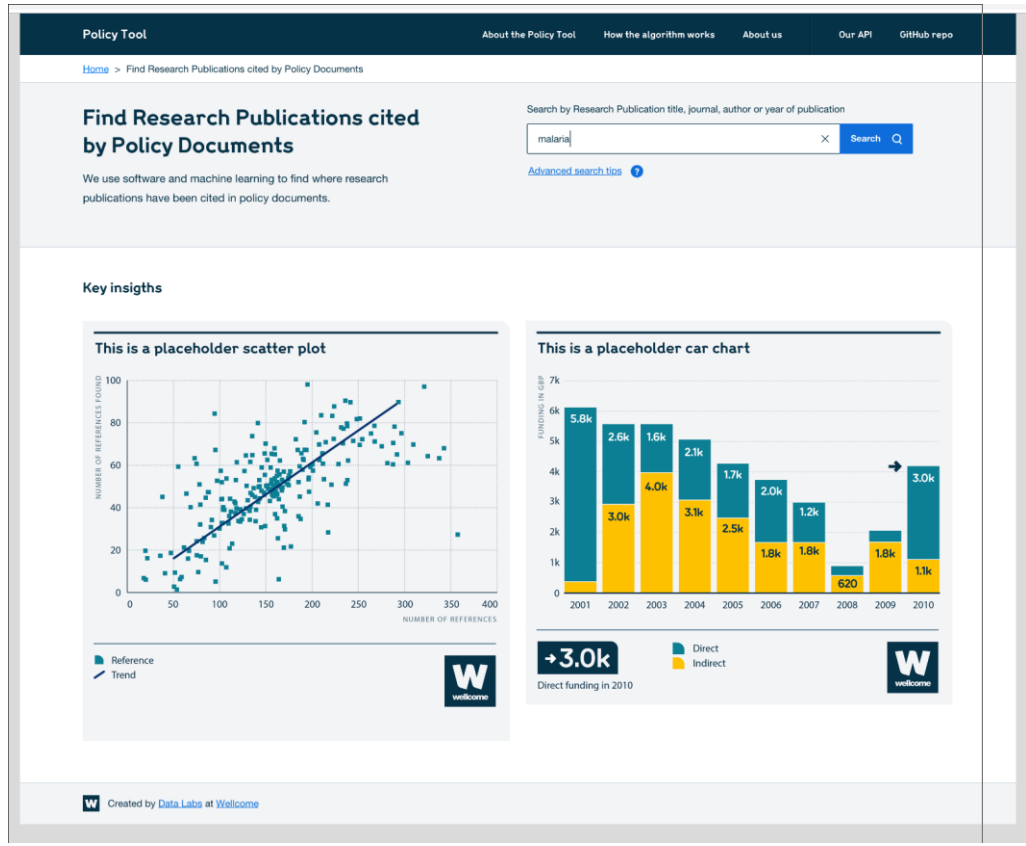
- We are working to build a productionised data pipeline that will automatically clean the data coming out of the Reach tools analysis to remove duplicates
- We are working to increase the accuracy of the 4 stages of ML and have already increase accuracy of the Split step from 57% to 67%. We expect to get this to c. 85% accuracy with current ML approach
- In a parallel track another member of the team is working on replacing the current Naïve Bayes approach with Neural Nets for all steps, which we will expect to push accuracy of all steps above 90-95%

Adding Functionality

- Add all text analysis function to find e.g. researchers and institutions using Lucene Search
- Add ability to analyse against sub-categories, such as disease types and document types
- Expand to other Funders and other sources of policy documents and other types of documents like white papers, conference proceedings and panel proceedings
- Integrate with patent and publications analysis tools
- Create a usable front end interface and make tools available as an open service

Releasing the Wellcome Reach Tool as an Open Service





Practical ethics in product development

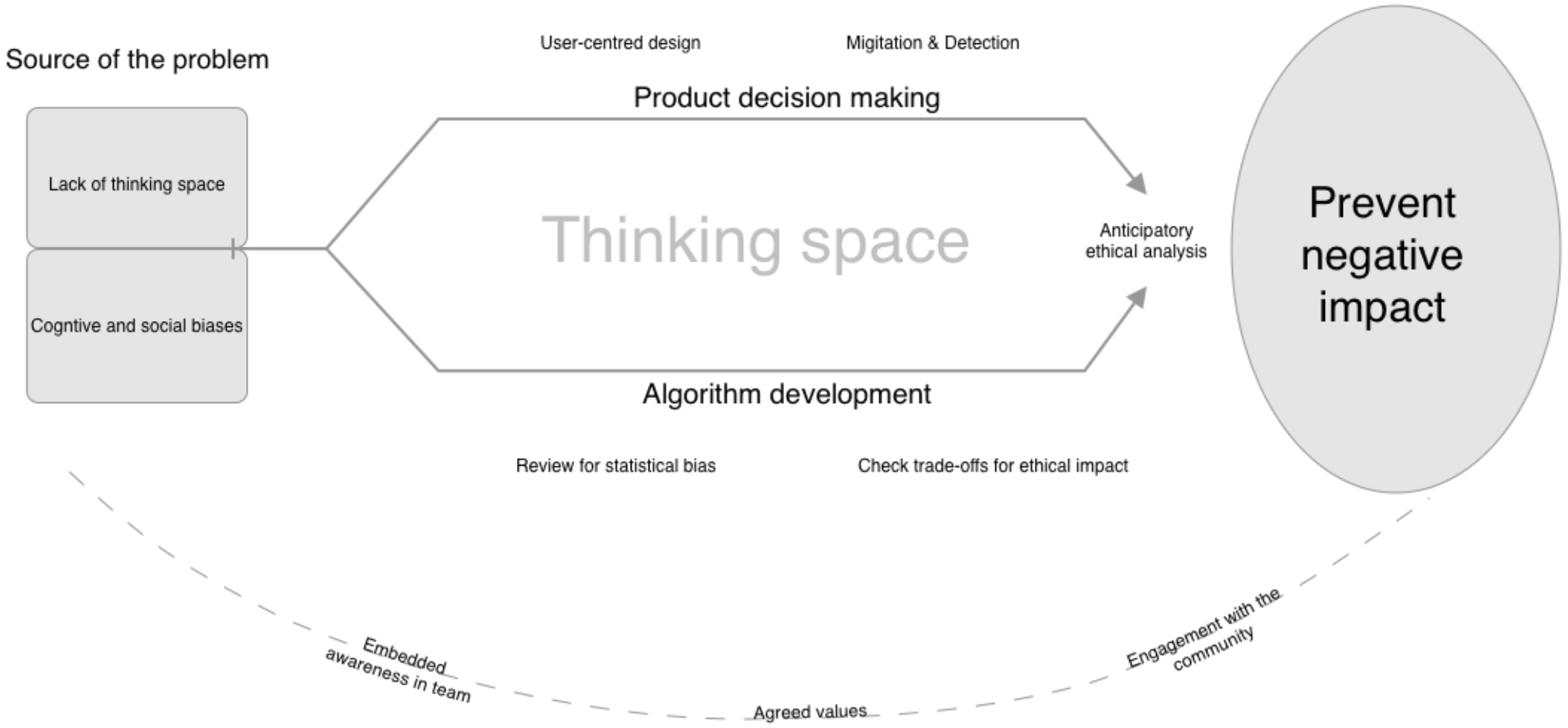
Our ambition

Blend ethical analysis with data science in practical, lean way that keeps up to speed with product development.

The starting point:

- User researcher collaborates with data scientists to uncover ethical risks.
- Objectivity and deeper layer of analysis is established with the oversight of an external social scientist.
- A working group provides a wider perspectives and keeps things in check.
- Transparency and openness a guiding principle.

Ethical Framework of Thinking for a Product Team

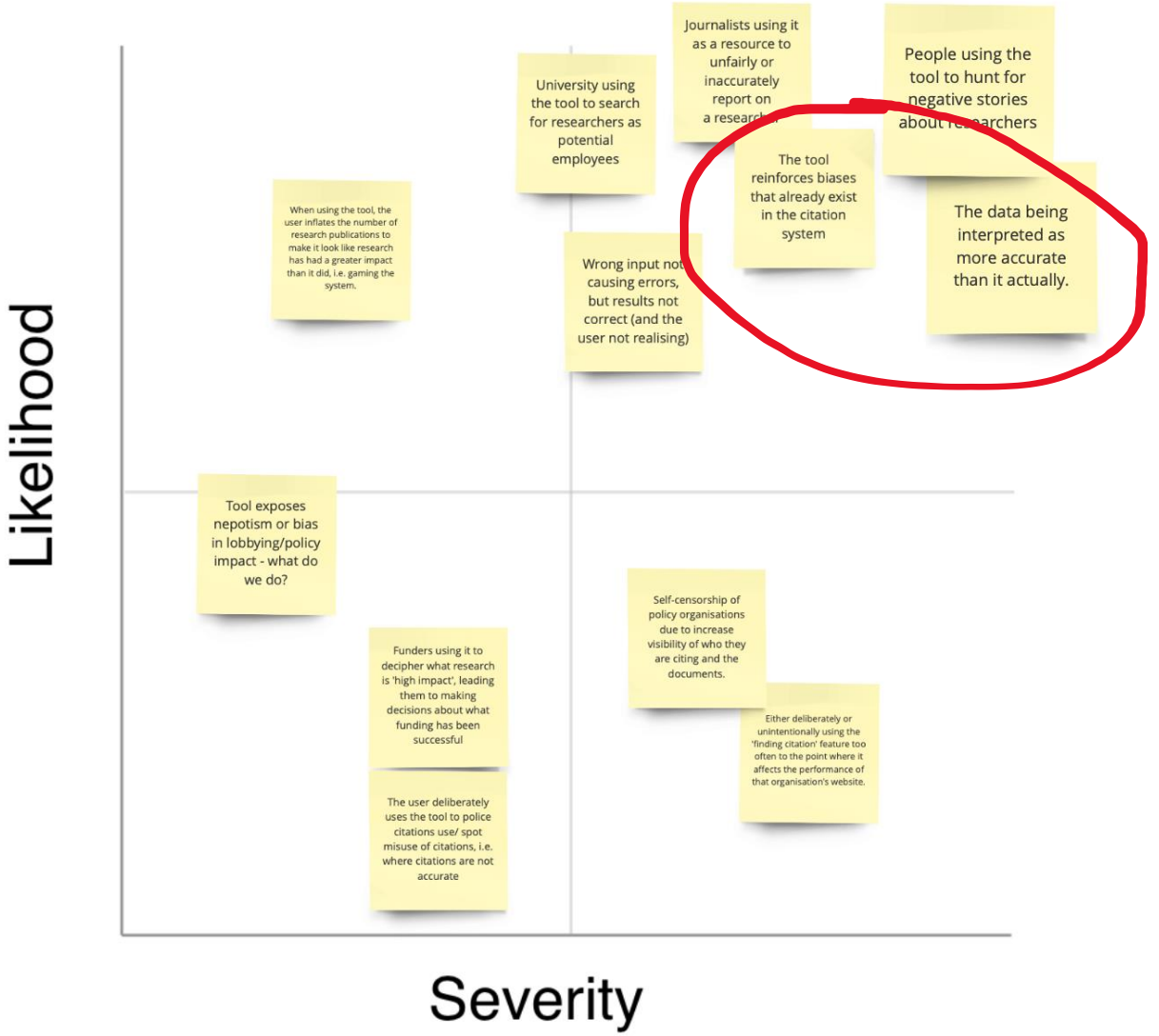


What have we achieved?

Applying creative thinking to generate anticipatory analysis

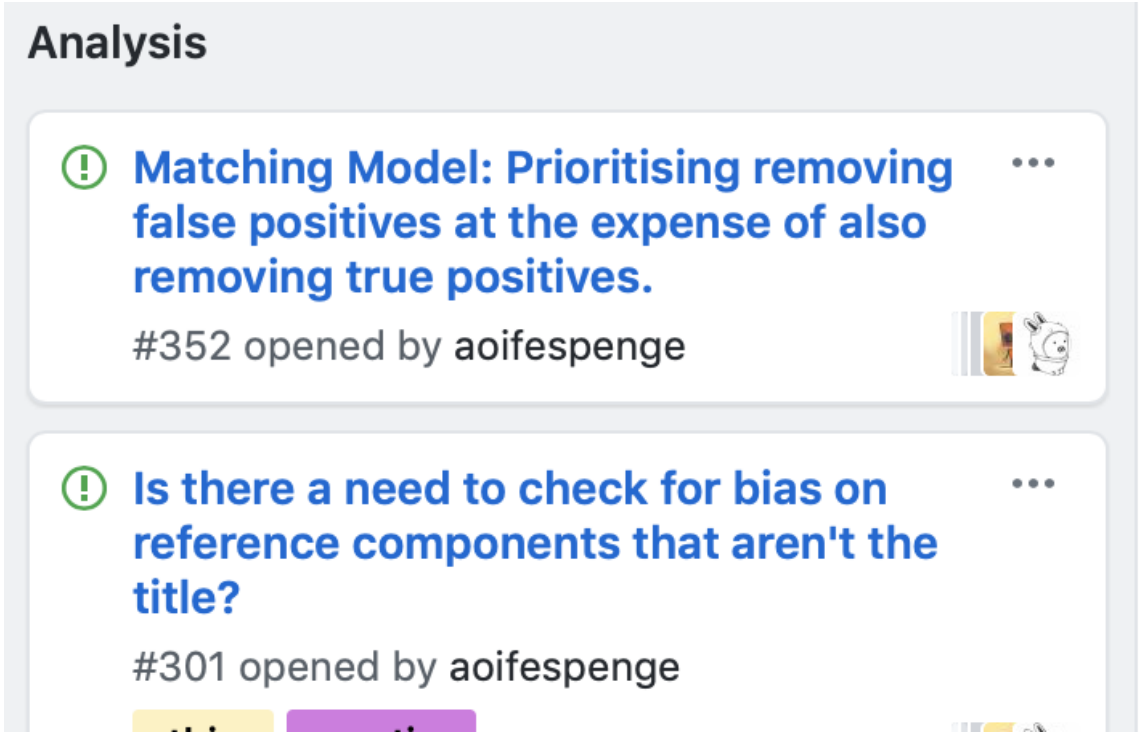






Using product prioritisation techniques to narrow on highest risks

Online and offline routines to create space for ethical thinking

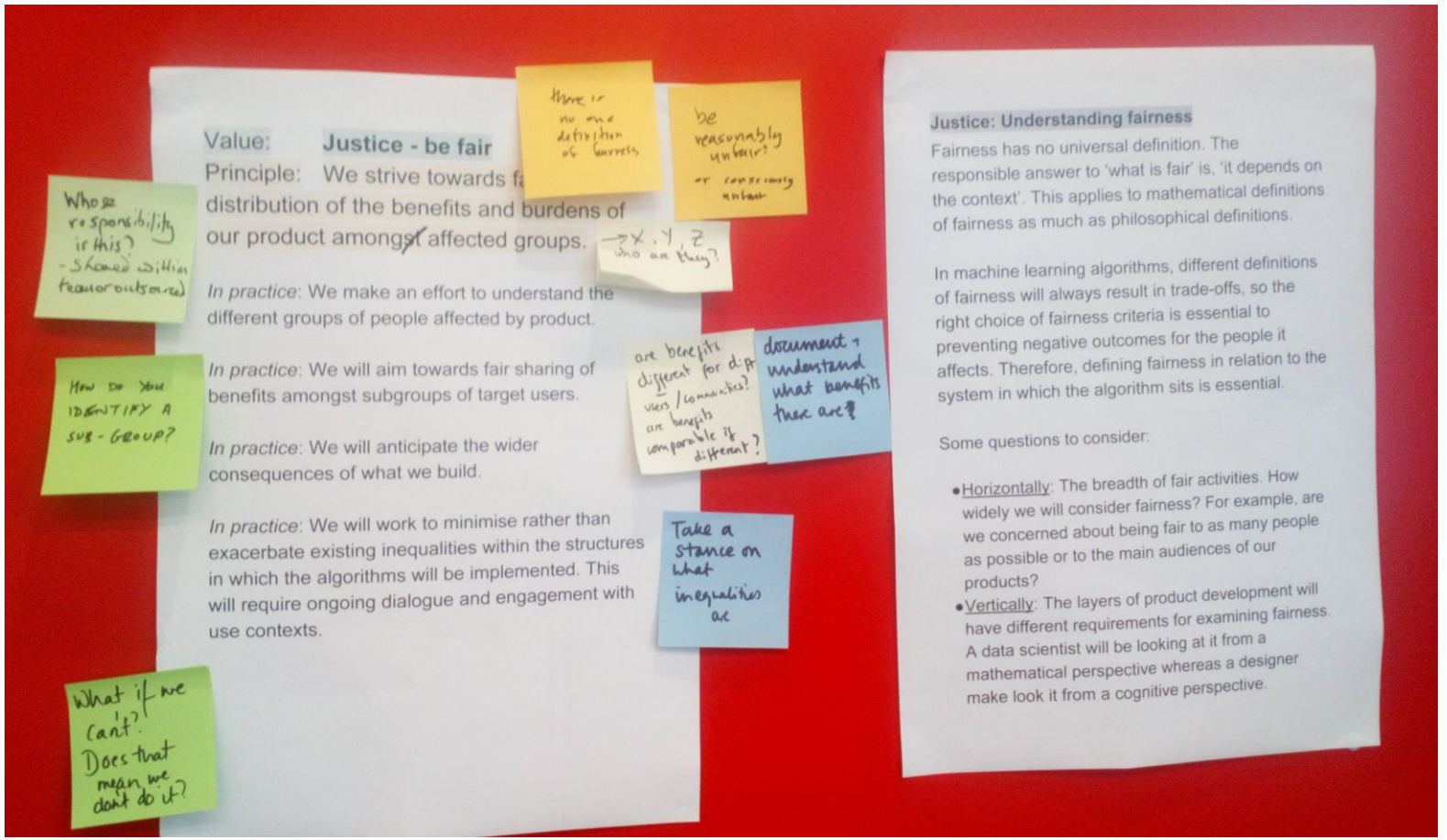


Monthly meetings to review our efforts.

Working group meetings every three months.

Using Github to discuss and prioritize ethical concerns.

Agreeing practical ethical values



Agreed set of ethical values and principles.

Aware of tensions and using them to focus our attention on solving them.

The challenges

Effort versus reward

How do we balance improving the product (eg faster product development; greater accuracy using neural nets) with assuring ourselves that we have sufficiently thought about unintended consequences?

While the point of effort versus reward is always pertinent, the difficult case we have is that we don’t know what we don’t know.

By assuming that research citations is a low risk area (eg compared to finding tumours in medical images) and therefore not putting as much time into anticipating unintended consequences, we run the risk of missing an ethical concern that was not obvious. However, pragmatically we have to draw the line somewhere.

Uncovering bias

It is impossible to remove all bias that could lead to negative consequences, but possible to align the appropriate fairness criteria with the community of people the algorithm system will affect.
It is for this reason we have learned that it is important to think of algorithms as one step in a larger system that affects people and build them transparently. Understanding the entire user experience and community map in which the algorithm sits is the key to understanding how best to optimise the algorithm ethically.
Are there opportunities for feedback and correction if the algorithm is found to be inaccurate? How do we communicate to the user what a responsible interpretation of the outputs are?

The issue of responsibility

The question of responsibility for misuse, has been one of the most difficult to answer. If you release a public tool you lose control over how it can be used by others.

For example, if the Reach tool is used to find and target researchers who work in politically controversial areas, like animal testing, which is not an intended use case, what is the responsibility of us as developers to anticipate this?

With thanks to Dr Chonnettia Jones, Jessica Romo, Nick Sorros, Dr Elizabeth Gallagher, Marta Moratti, Sam Depardieu, Hunter Blanks, Aoife Spengeman, Dawn Duhaney, Natalie Leach and Andy Lee.

Please check out our blog at
medium.com/wellcome-data-labs



wellcome.ac.uk