



Why do we need an update to the Data Strategy?

Time to take stock

In 2016 we created our data strategy which set out our vision, principles, goals and action plan for the future. Three years on, it's a good time to review our progress; to look at where we are, confirm we're on the right track, and reaffirm our commitment to the original strategy. You can find the full document here: <https://www.food.gov.uk/document/data-strategy>.

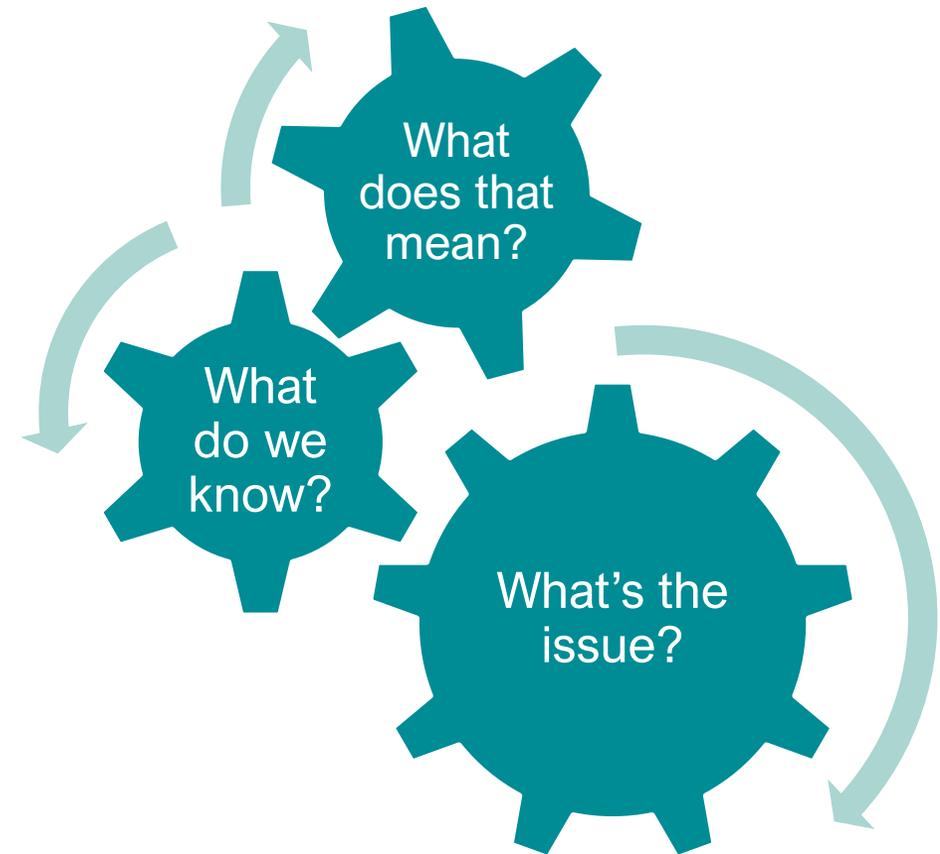
This update describes the work we've done towards achieving those goals. Our roadmap for delivery has six key themes:



Introduction

Using data to make a difference

In today's world, there are virtually no areas of government or business that don't depend on data, and very few jobs that can be done without either using or generating it. That said, we know that not everyone is interested in data for its own sake; so in this section we discuss a few of the ways data benefits the Agency at a strategic and operational level. The more detailed technical descriptions of what we're doing can be found later on in this document.



Helping everyone get the most from their data

Phrases like ‘data literacy’ and ‘data driven culture’ are popping up in strategy documents more and more now; what do they really mean? Just as literacy is about people learning to derive meaning from the written word, data literacy provides a set of basic skills around deriving meaning from data, and being able to turn it into reliable, useful information.

This year for the first time the Agency will have access to enterprise-quality reporting and analytics tooling, so our users can access reliable data that is accurate, documented and automatically refreshed.

At a more advanced level, Agency experts need to keep pace with the changing technical landscape of data innovation and explore how to take advantage of new technologies and data handling techniques.

Connecting, sharing and working well with others

We've focused on adopting (and in some cases creating) defined data standards and interoperability so we can share data with partners more efficiently, or mash up their data with ours, and our shared understanding of what's happening is enriched as a result. We can work together better on issues that affect us jointly. We discuss this in more detail in the most recent Chief Scientific Officer's report: <https://www.food.gov.uk/about-us/csa-science-reports> .

We actively engage and share learning with the U.K. food industry, other government departments (e.g. APHA, HMRC) and academia. For instance, working on the Internet of Food Things (IoFT) initiative led by Lincoln University, we are looking to connect the data with the physical foodstuff. On an international stage we are engaging with the Global Food Safety Initiative and the U.S. Food and Drug Administration.

Understanding the ecosystem we operate in

Mapping our data ecosystem gave us a more complete picture of the food and feed supply chain and the food business landscape, so we're in a much better place as an effective modern regulator. We know where to look for risk and impacts, and we understand the factors we need to watch. We can decide where we use resources to be most effective; where we should act, where we should influence, and where we should monitor.

Making use of our expanding Data Science capacity, we can also use sophisticated methods of predictive modelling to identify potential incidents before they arise, with a better understanding of the factors involved. We can explore complex relationships using weather, climate, demographic and economic data to name just a few of the options available to us.



.....

DATA ECOSYSTEM

.....

We said

We want to understand our data ecosystem.

We will create a unified view of food businesses.

We will develop our metadata management capability.

We did

We developed a map to show the processes that create, manage and maintain our data assets.

We cleaned, enhanced and aggregated key data sets.

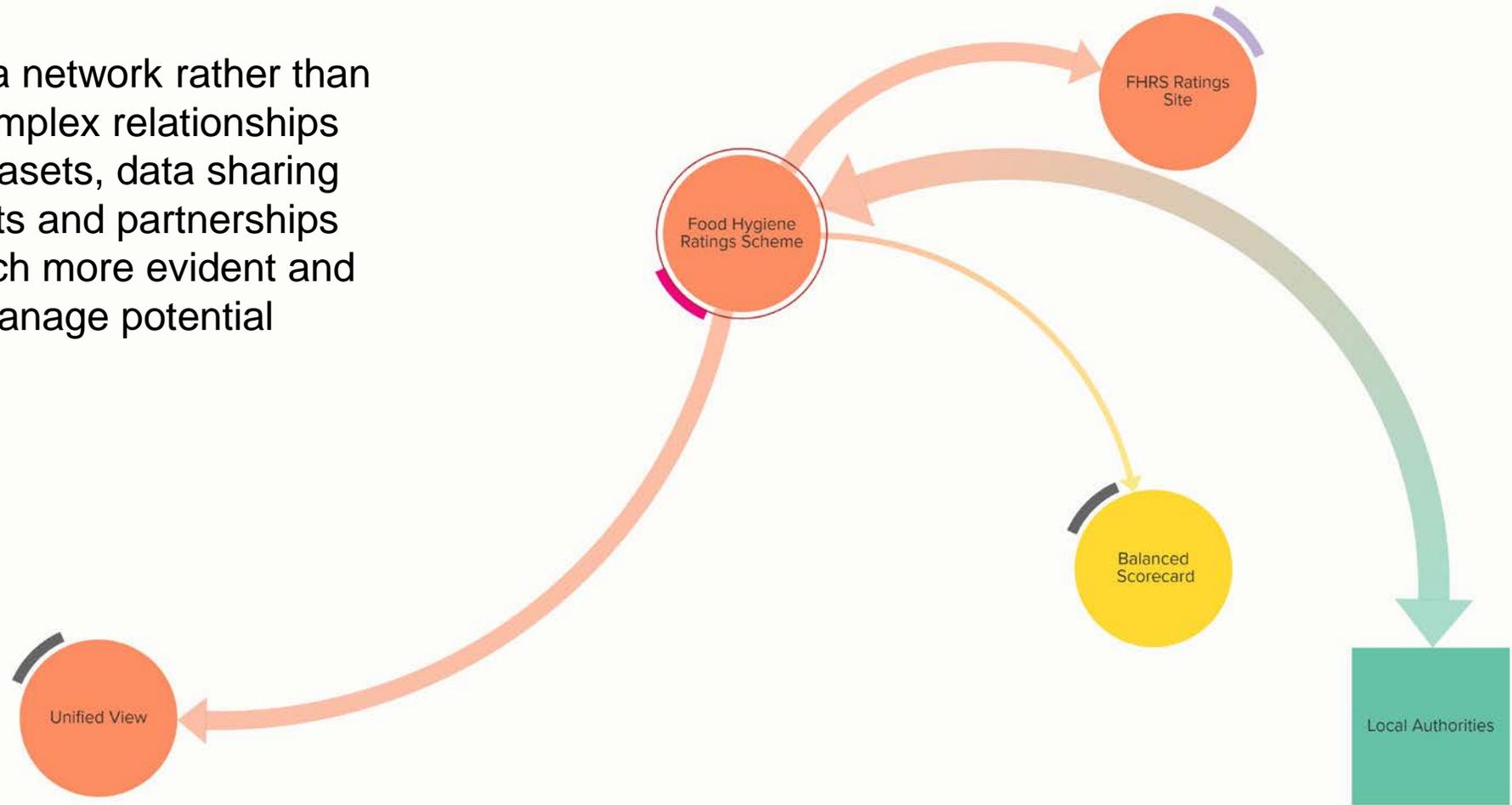
We improved our internal cataloguing and external publishing.

Mapping the data ecosystem

Mapped as a network rather than a list, the complex relationships between datasets, data sharing arrangements and partnerships become much more evident and help us to manage potential issues.

Legend

- Data Flow
- User
- Service
- Process
- Application
- Publication
- Report
- Microservice
- Dataset



Data Ecosystem – Next Steps

- We will improve the data ecosystem map to be more detailed, and link it to business and technical documentation for our processes and services.
- We will standardise the definitions and data models for the key data entities in our business.
- We will effectively communicate those entities to the business e.g. what do we mean when we say “establishment” or “premises”?



DATA FLOWS

We said

We will define and agree organisational data standards.

We will explore and adopt modern tooling for manipulating and managing the flow of data.

We will make it easier to collaborate and share data.

We did

We created templating to define and share our data standards.

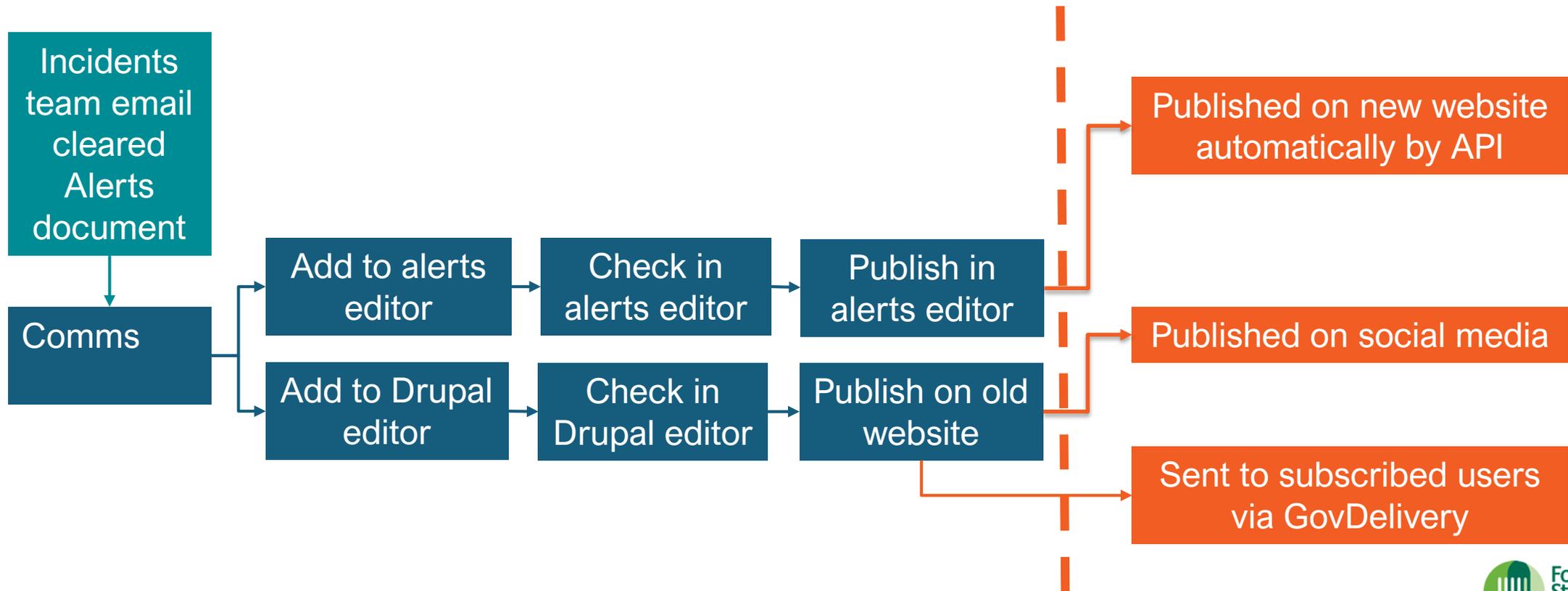
We piloted the use of cloud based tooling for extract, transform and load of data.

We developed different tools to consume and share data, including a look at how we can use APIs*

* Application Programming Interface

Improving data flows: using our Alerts API as an example

The Alerts service provides a public notification of issues like allergens, contamination or product recalls via multiple social media platforms. This process is very manual, and involves the Comms team having to update information in several different places once they receive an alert document from the Incidents team.



Improving data flows: Simplifying and automating the process

The new process still requires the incident information to be checked and added manually, but once that's done and published to the Alerts Editor the alert is pushed to all the designated platforms using our API.



Data Flows – Next Steps

- We will expand our data tooling offering to colleagues in Data Science
- We will design data standards as part of end to end service design
- We will identify and address capability gaps (i.e. be able to add new sources to Unified View)
- We will work with other government departments to improve sharing and interoperability

.....
INSIGHT FROM OUR DATA
.....

We said

We will make our data more discoverable for internal and external users.

We will develop and share data stories to promote our work and lessons learned.

We will make it easier to create accurate, repeatable and trusted reports across the organisation.

We did

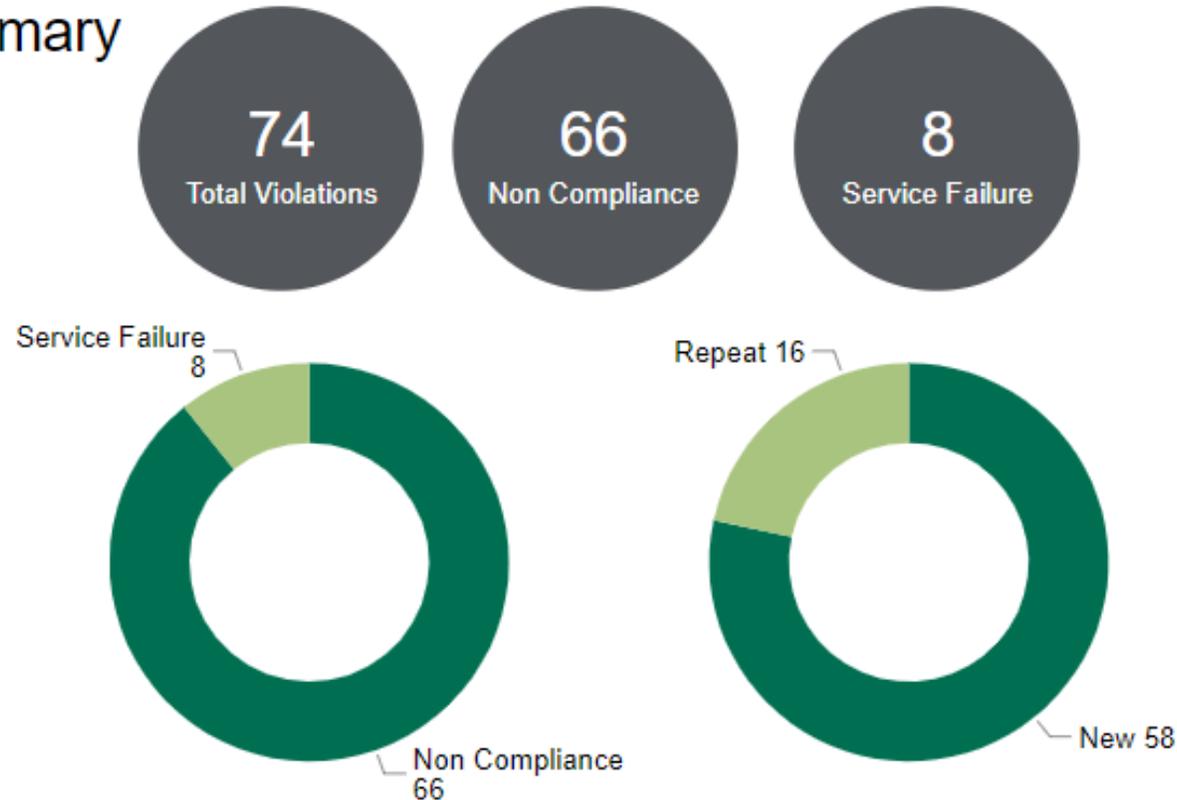
We improved our internal data cataloguing and external publishing

We have written data stories about our data science projects, shared through blog posts and interactive dashboards.

We developed our capability to provide clean, automated data as managed content to colleagues.

Deriving insight from trusted data

Summary



Section of a prototype dashboard for Operations

We can serve up validated, current and securely managed data to colleagues so they can produce the management information and reporting they need to see. A combination of understanding what data we have, building the infrastructure to manage the data efficiently, and making the data accessible to our corporate management information tools, all combine to create a pipeline for trusted information to reach decision makers.

Insight from data – next steps

- We will share more of our data stories, internally and externally
- We will share our experiences collaboratively and freely with our colleagues and stakeholders
- We will increase our capability and capacity to deliver robust data sources for reporting, and test that by getting value from less structured data sources

.....

**CAPABILITY AND SHARING
KNOWLEDGE**

.....

We said

We need to increase the overall data literacy across the Agency.

We need to develop our Data Science function and find ways to increase its impact.

We want to promote the culture of self-sufficiency in data work across the Agency.

We did

We fostered communities of practice around coding, report building and data analysis.

We expanded the team to give the business more access than ever to data science resource.

We trained and supported colleagues to maintain and publish key datasets.

Capability and sharing knowledge – next steps

- We will develop an action plan for improving data literacy across the Agency - supporting colleagues to place data in context, using effective analysis and visualisation tools and techniques, and understanding the data landscape around them including ethical considerations
- We will share findings from our data initiatives with a wider range of stakeholders through various forums (e.g. lightning talks)
- We will support colleagues to use opportunities from digitisation to reduce data duplication (collect it once)
- We will start the internal conversations about data ethics and the work we need to do to understand the issues

.....

DATA AS A SERVICE

.....

We said

We will design a future state data architecture fit for a modern regulator.

We will deliver an architecture that is in line with government and Agency design principles, and publish open data.

We will collaborate with colleagues across Openness, Data and Digital to create the best data components to support effective digital services.

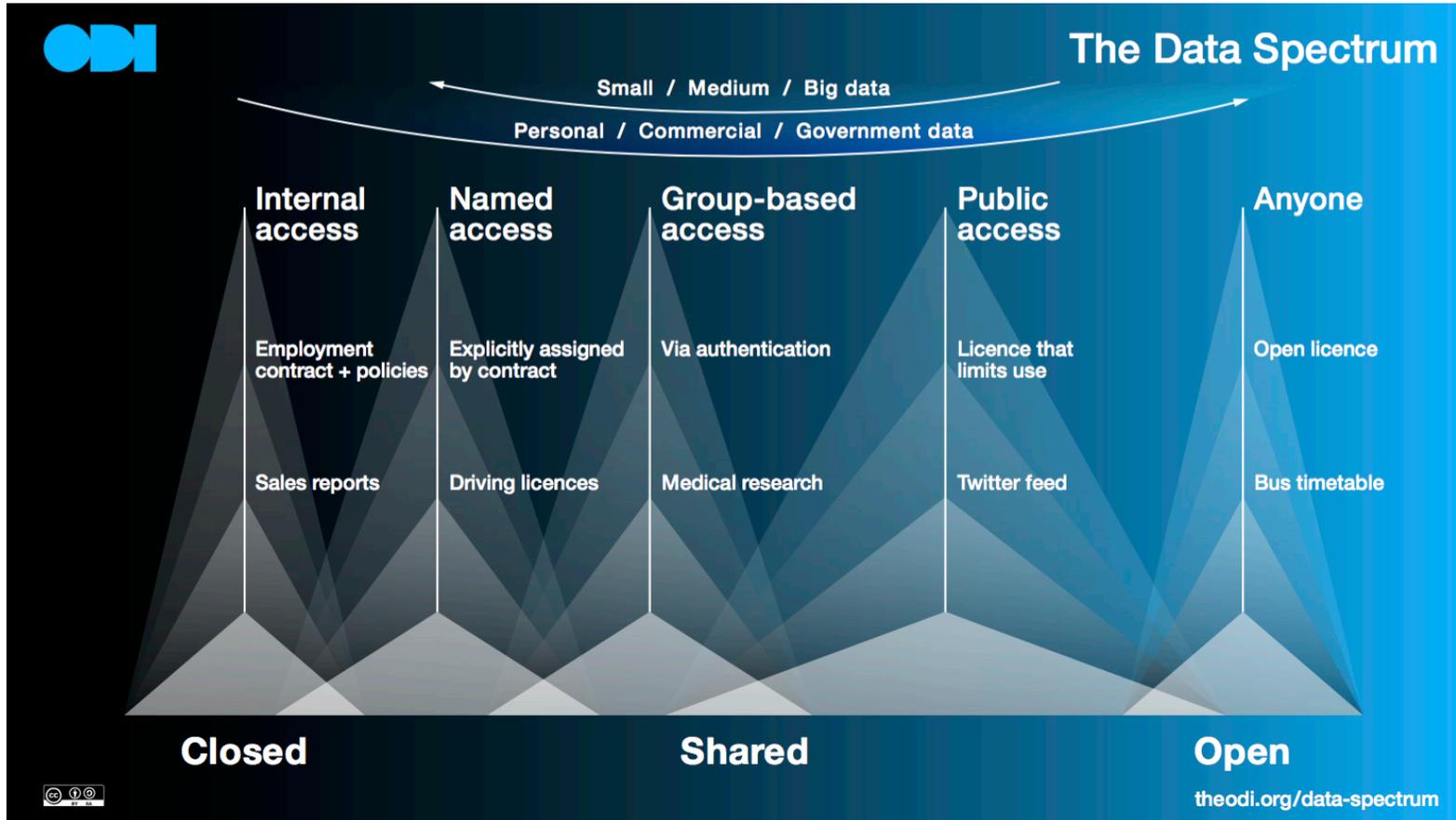
We did

We combined some of our key data sources into one unified view of all food businesses.

We developed how we model reference data to deliver reusable, accurate and trusted reference data, and a catalogue of open data.

We agreed some common tooling for data science colleagues to improve the effectiveness of prototype delivery.

The data spectrum



Data can rarely be categorized as 'completely open' or completely closed'; the data spectrum model explores how different types of data need different levels of access management based on the sensitivity and the inclusion of specific attributes e.g. personal data. Designing data services, we need a robust understanding of who is accountable or responsible for a dataset and who should be able to access it under what circumstances.

Open data

We make all of our open datasets available on our data catalogue, free for anyone to use. A dataset can take many forms (tables, spreadsheets, database reports) but is essentially a structured set of data. We support colleagues to publish their own datasets through the catalogue editor tool, and monitor activity and usage through the editor's API.

Building and managing data services

As we design services to make data more accessible, we need to understand and document several crucial elements for the datasets the service calls on.

- Who is accountable or responsible for a dataset?
- Who maintains and updates the data, and can provide context when needed?
- Who should be able to access and use the data, and under what circumstances?

These concepts are very similar to the levels of control we have for records management, and are well understood when it comes to things like information security and safeguarding personal data.

As we progress along the data maturity scale, we can use the same processes to create secure, managed, appropriate access to data for Agency colleagues, partners and the general public.

Data as a Service – next steps

- Continue to develop the tooling, governance and capability to deliver and maintain internal and external registers that are robust enough to be used in services
- Expand the available sources in Unified View and develop our internal capability to model and design data ingestion into the service
- Building on the user research we commissioned on the next phase of the data catalogue and publishing workflow, we will begin implementing the improvements we've identified.
- Actively participate in cross government and cross sector data and standards initiatives

.....
DATA INNOVATION
.....

We said

We will pilot innovative data technologies to improve food traceability and use data from the food supply chain effectively

A system that allows us to identify food and feed risks in a predictive way will, for example, enable better targeted and more cost effective sampling

We did

We have run two pilots of Distributed Ledger Technology (Blockchain) for Collection and Communication of Inspection Results (CCIR) and Export Health Certification

We have run 9 sprints so far in our strategic surveillance programme, using data science techniques to answer tricky questions from business users

Surveillance: using open data to predict risk

Problem statement

- If we can relate known hazards to climate, can we predict unknown hazards by looking for similar climates?

Solutions Highlights

- We built a predictive model using aflatoxins in figs from Turkey as an example.
- We were able to model aflatoxins in figs from countries with similar climates.

Key findings

- Climate risks can be quantified.
- Climate data is highly correlated with some hazards.
- Known risk and climate can be used to predict unknown risk.

New technologies: Using Blockchain to improve food traceability

Problem statement

- Fraud and error in the food supply chain impacts on industry and consumer trust
- Data is used ineffectively across the food supply chain

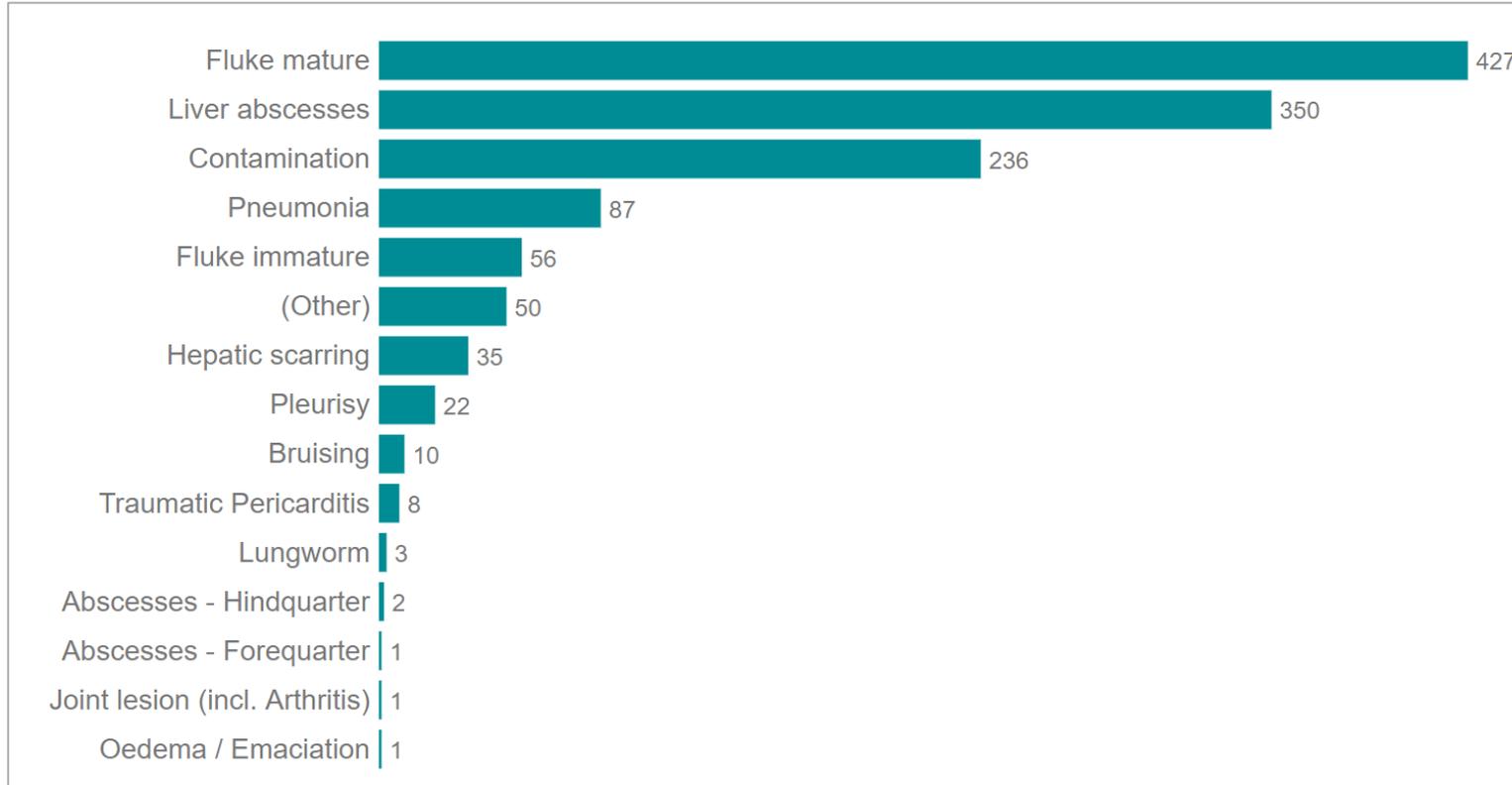
Solutions Highlights

- Two pilots trialled using blockchain (distributed ledger technology) to improve traceability by working with industry
- Data stored securely on blockchain with role based access controls

Key findings

- Improved use of data to reduce operational burdens
- Improved ability to trace an individual animal with conditions

New technologies: Using Blockchain to improve food traceability



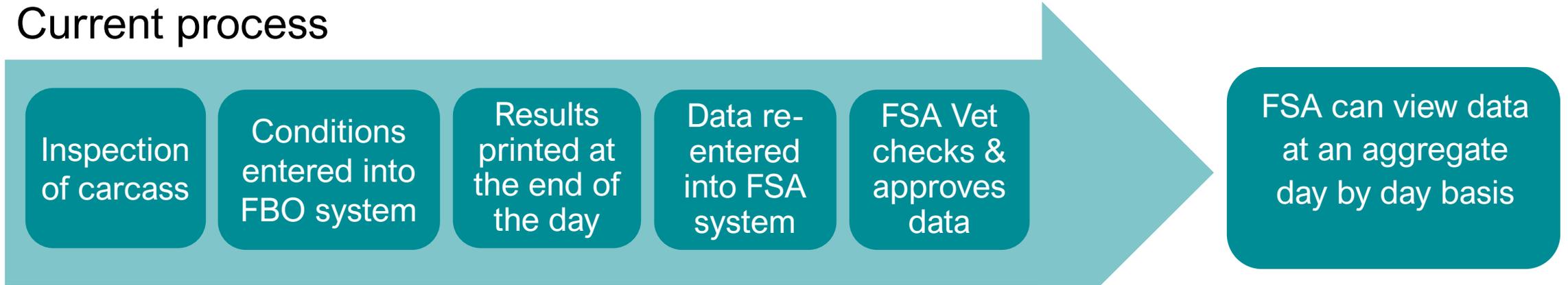
This very detailed data exists in meat processing plant systems, but when they report it for our purposes staff have to manually calculate tallies and enter them by hand into a different system. Aside from the time spent, this also introduces the possibility of errors and breaks traceability.

Excerpt from report: Number of animals found to have the condition at inspection.

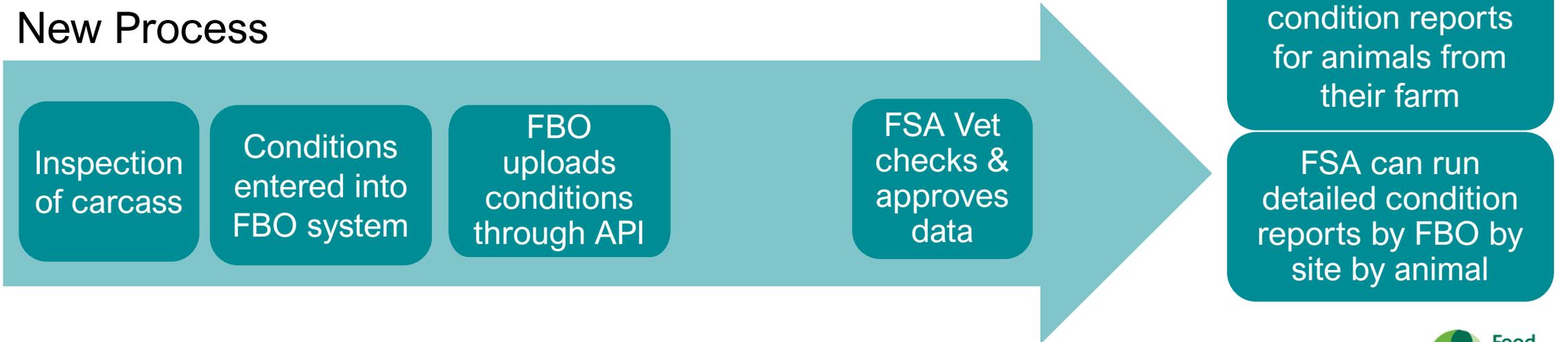
.....

New technologies: Using Blockchain to improve food traceability

Current process



New Process



Data innovation – next steps

- We will link physical sampling with our data to develop a strategic surveillance approach.
- We will align with cross-government work on the National Data Strategy.
- We will work with the science council to explore opportunities to use Artificial Intelligence to improve food system risk management.
- We will work with other government departments, industry and academia to share learning and identify opportunities from our Blockchain pilots.
- We will continue to exploit open data sources using data science capabilities to identify risks and inform interventions through strategic surveillance.
- We will increase our use of novel data sources, e.g. less structured data, to develop horizon scanning capability.
- We will enrich our data services with data from other sources (e.g. GS1 barcodes for recalls).