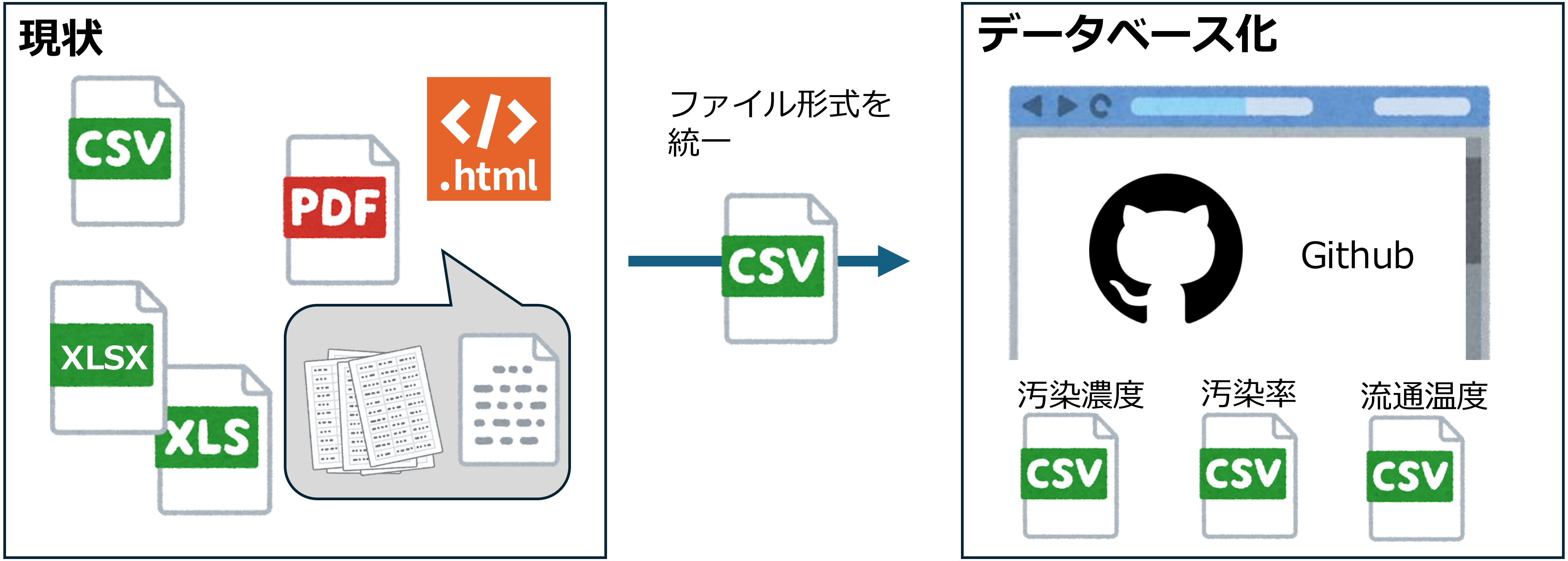
食品の細菌汚染実態の調査から見えたデータ活用の課題



北海道大学大学院農学研究院　小山健斗

背景：曝露評価で用いるデータの管理

現状：どこに何があるのか不明瞭、**信頼性の担保ができない**

デジタル技術を用いる前に、データの整理が必要

参考事例：人工知能AIによる肺線維症の解明
http://www.imed3.med.osaka-u.ac.jp/research/r-resp09.html

目的：曝露評価で用いるデータの管理

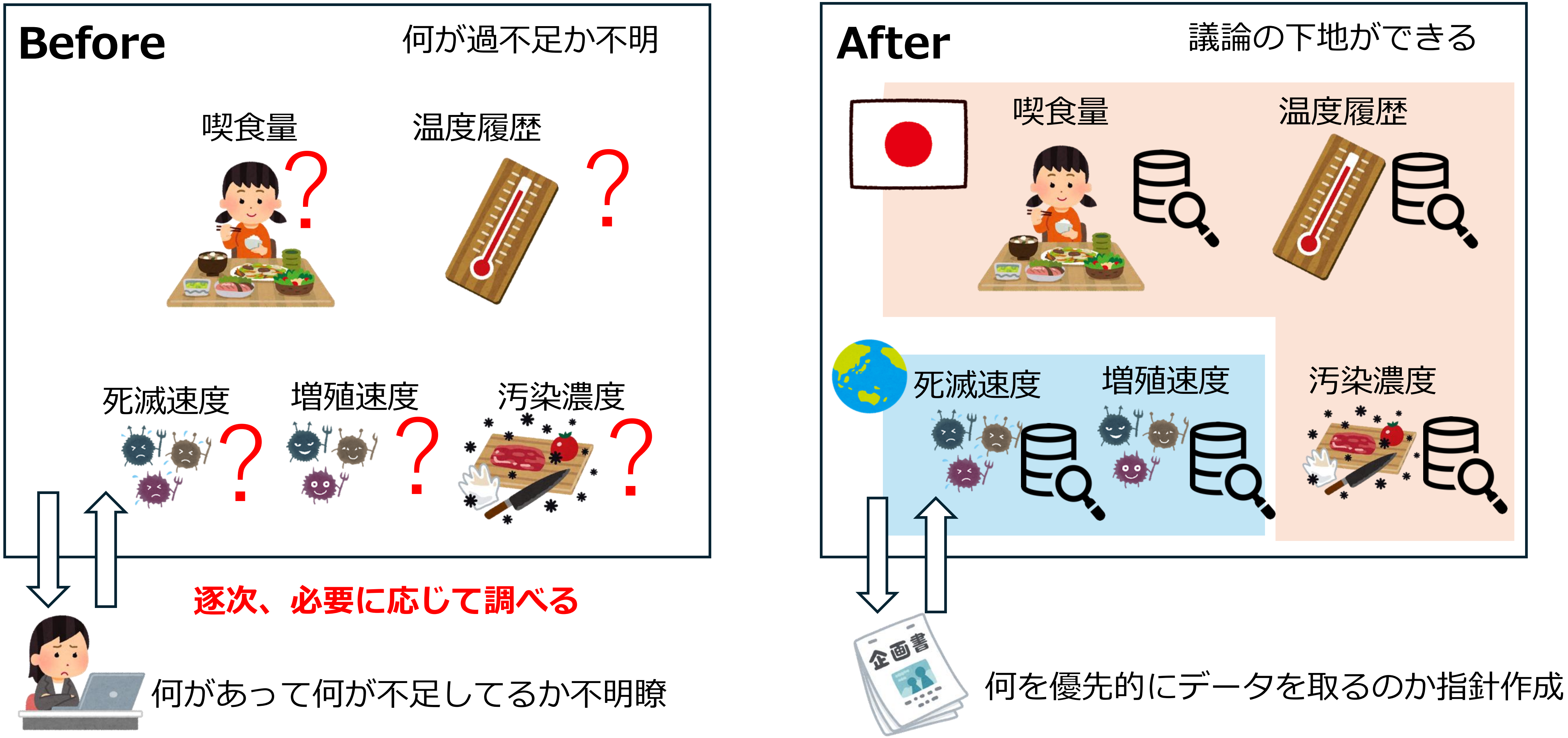
リスク評価のデジタルデータの基盤を作る。

だれでも途中から参画、運用できる場を作る。

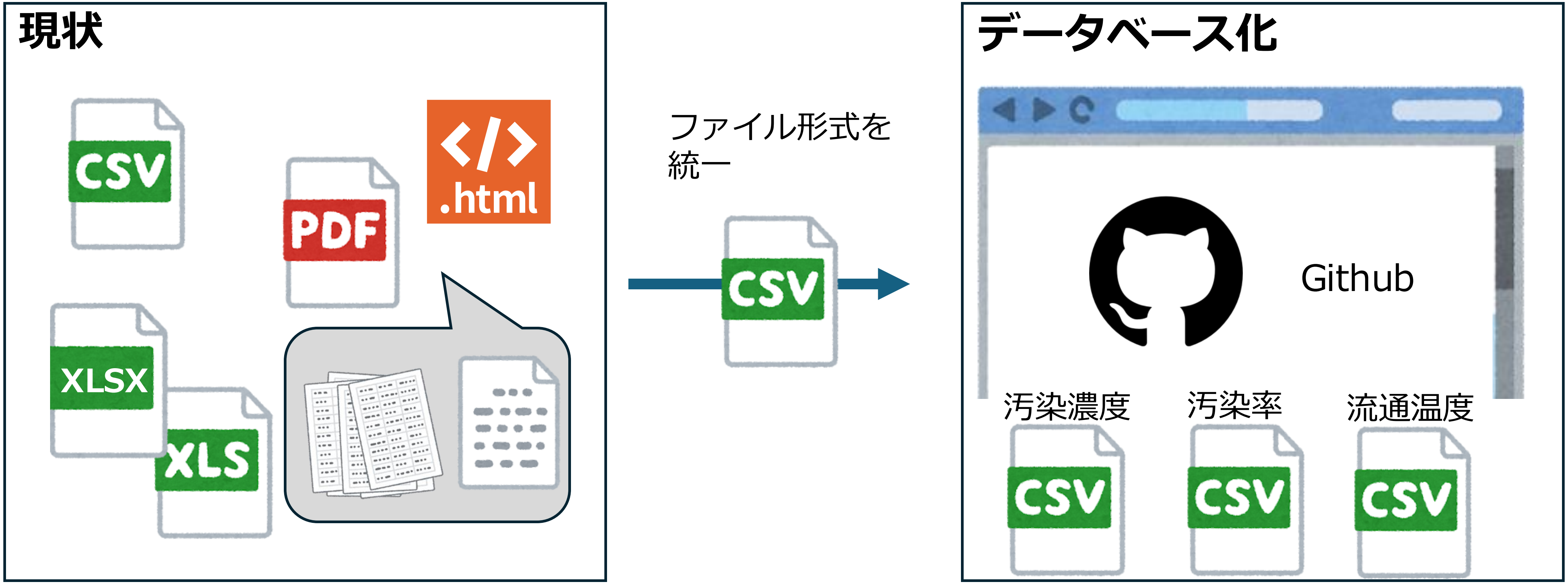
現在：データが散在　　データベース　　AIなどの活用

参考事例：人工知能AIによる肺線維症の解明
http://www.imed3.med.osaka-u.ac.jp/research/r-resp09.html

方法：曝露評価で用いるデータの中で何を管理対象にする？



方法：曝露評価で用いるデータの管理

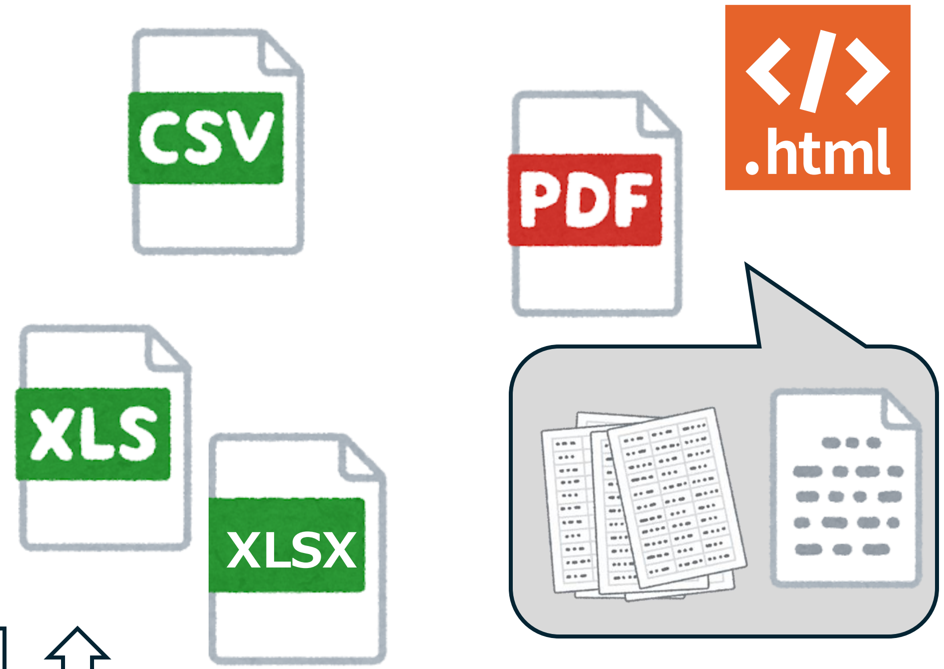


国民に対しての説明できる素材をオープンにする

方法：曝露評価で用いるデータの管理

Before

どこにあるのか不明

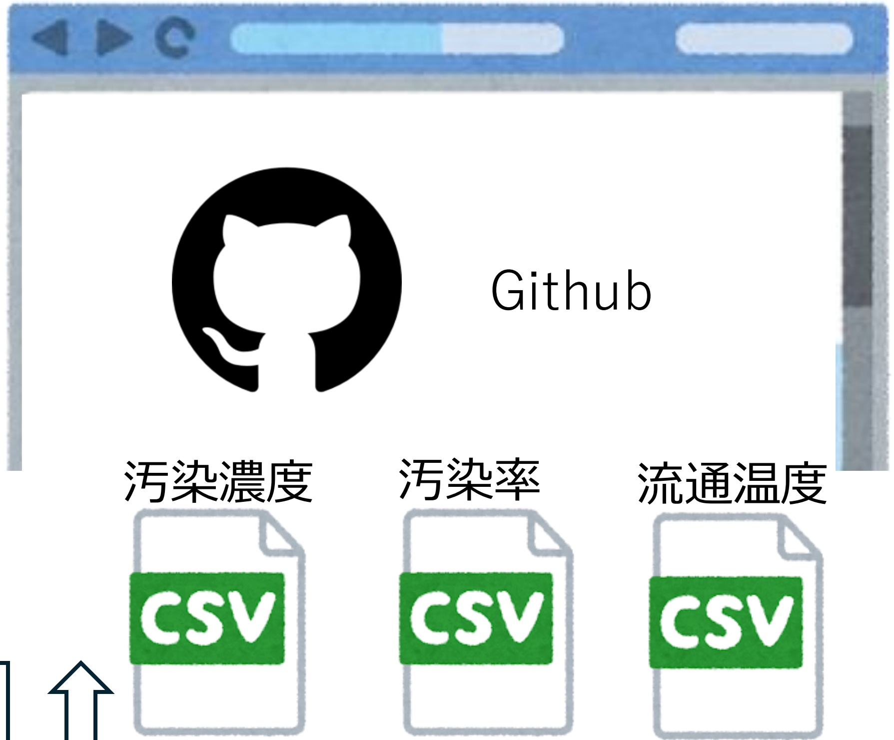


厳しいセキュリティ、雑多なデータ

長く業界にいる方のみが理解できる

After

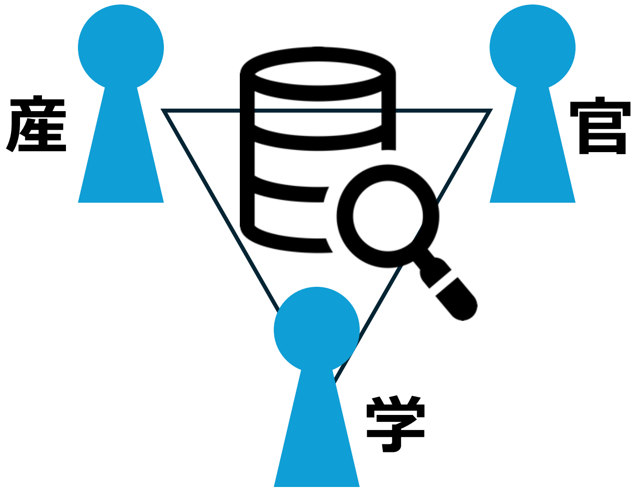
ここに行けばわかる！



APIの利用（外部からのアクセス可）

情報科学分野の方が参入しやすい

方法：曝露評価で用いるデータの管理



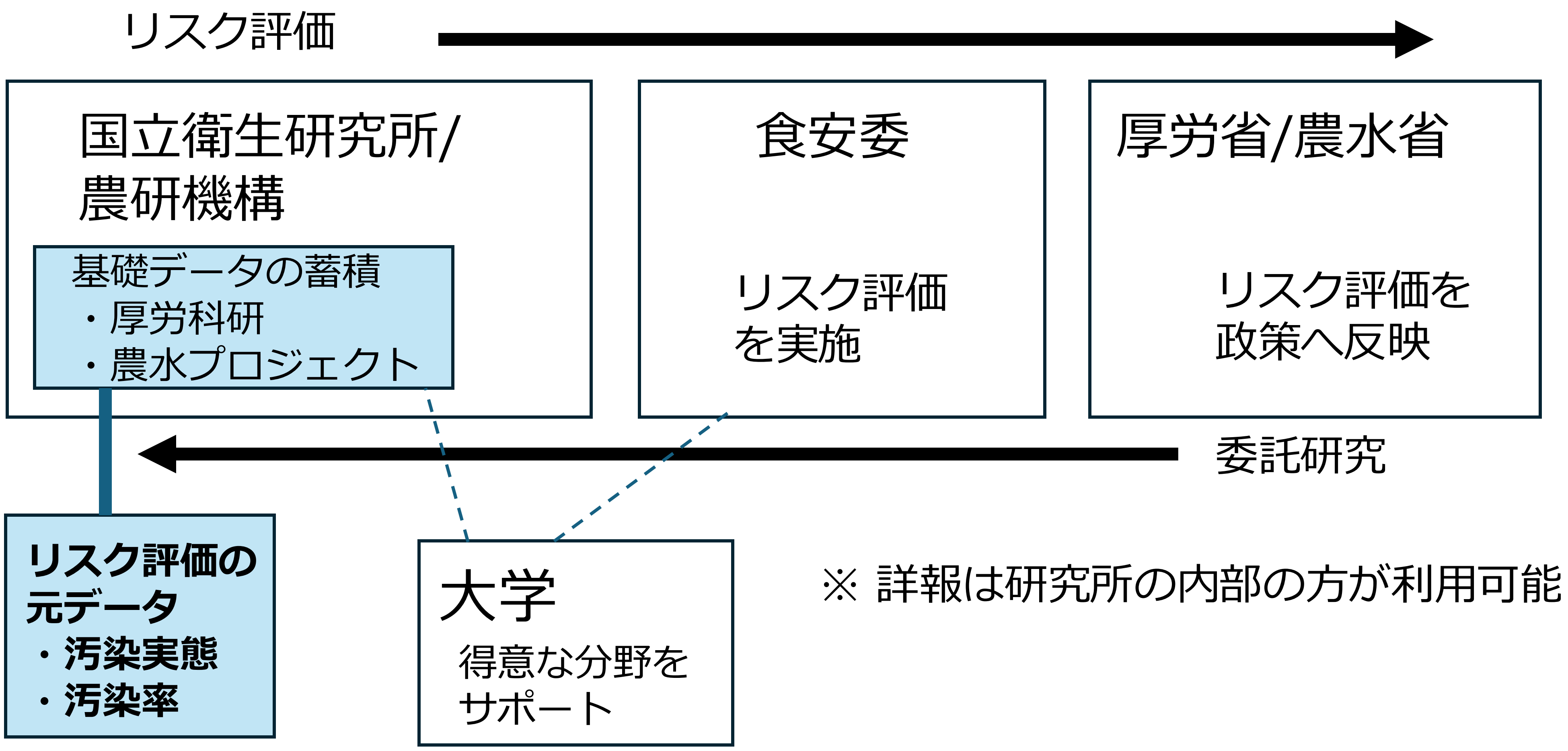
優先事項：
・情報系の方々が参画しやすい。
・途中からでも管理しやすい。
・手間なく管理、閲覧できる。

Github

・データ・コードを共有が容易　→　だれでも参加できる
・更新履歴あり　→　だれでもプロジェクトの歴史を辿れる
・URLを公開　→　だれでも見られる

将来構想：データ運用に向けて

デジタル化プロジェクトは研究所が将来的に担うのがよいか



将来構想：データ運用に向けて

海外でも政府データの解析は研究所が担っている。

Dashboard for Interactive Analysis of Mycotoxin Occurrence in Human Foods Data: Streamlining Annual Report Generation and Data-Sharing

Tabitha J. Miller[1], Anthony Adeuya[1], Ernest Kwegyir-Afful[2], Lauren Posnick Robin[1]
1. US Food and Drug Administration, Center for Food Safety and Applied Nutrition, Office of Food Safety 2. US Food and Drug Administration, Center for Food Safety and Applied Nutrition, Office of Food Additive Safety

Abstract

Fungal growth on agricultural commodities not only decreases the value of the product, but also presents a health concern. These fungi produce mycotoxins, a toxic metabolite produced by certain species of fungi. The FDA has generated a large amount of data through the Compliance Program for Mycotoxins in Domestic and Imported Human Foods regarding the occurrence of mycotoxins in foods. Each fiscal year, an internal annual report is generated summarizing the data collected via the compliance program, but no external-facing report has been generated to summarize and discuss the data. In order to create an overview of the data, find trends, and facilitate report development, we created a data dashboard using Amazon Web Services Quicksight. The new dashboard has increased our ability to review and summarize mycotoxin data as well as respond to questions more rapidly.

Introduction

Mycotoxins are toxic metabolites produced by specific fungi that can infect and grow on various agricultural commodities. Aflatoxins, fumonisins, dexynivalneol, patulin and ochratoxin A are the main mycotoxins of concern for humans. Foods commonly contaminated with mycotoxins are those derived from cereal grains (wheat, corn, etc.), peanuts and tree nuts, legumes, and certain fruits. Mycotoxin occurrence is influenced by environmental factors such as temperature, humidity, and rainfall during growth and harvest stages. Deleterious health effects due to mycotoxins range from gastrointestinal upset to long-term consequences including cancer. Because mycotoxin occurrence is not entirely avoidable, small amounts may be legally permitted in foods. Therefore, to protect consumers, FDA has issued action levels and guidance levels for certain mycotoxins. FDA staff review mycotoxin findings on a case-by-case basis, considering the action/guidance levels and other factors, to identify adulterated samples. About 2000 samples of susceptible human food products (domestic and imports) are collected and tested yearly for one or more mycotoxins each year. The CFSAN Office of Food Safety develops an internal report each year summarizing the findings, but an external-facing report would provide added value for stakeholders. In addition, there is also a need to support FDA work on international standard setting in Codex Alimentarius. The dashboard was developed to address these needs.

Methods

To prepare data for the dashboard, it was first downloaded from the Office of Regulatory Affairs' (ORA) FACTS database. Manual cleanup steps were conducted, including ensuring correct product codes, consistency in the units used for data reporting, and removing any non-human foods data. The data was then uploaded to Quicksight through a direct query data connection upload portal developed by collaborators in OFAS. Once the data was on Quicksight, additional data fields were generated to facilitate categorization, calculations, and visualizations.

Table 1. Mycotoxins included in the FDA Compliance Program.

Mycotoxin	Foods Commonly Contaminated	Current FDA Regulatory Level

Results and Discussion

The resulting dashboard allows for quick data tabulation and display of statistical measures of mycotoxin findings. Subcategories (specific years, mycotoxins, types of foods, etc.) can be highlighted to tease out factors that might affect mycotoxin adulteration, such as region of production. The functionality of a dashboard also provided an opportunity to design pages with specific work functions.

The first page gives a big picture outlook on the occurrence of mycotoxins in foods. Adulteration rates by mycotoxin are displayed, and statistical data about the number of samples tested overall for each mycotoxin are displayed. Additionally, we added a geospatial display for quickly displaying a heat map of adulteration rates by country of origin. This page also contains the controls that allow the user to focus in on specific subsets of data (e.g., a specific country of origin or specific food category).

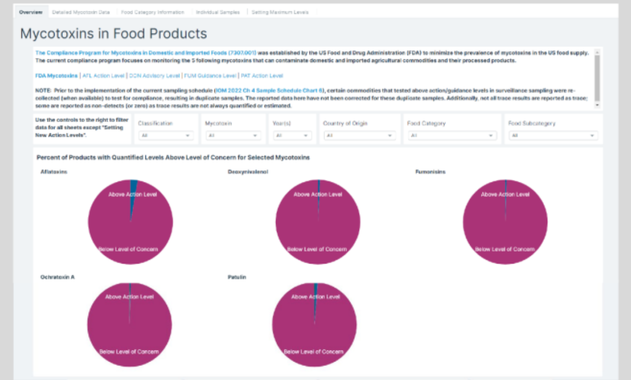


Figure 1. The dashboard has various pages. Shown here is the overview page and some

The third- and fourth-pages breakdown the data even further, looking at the data by food category and food subcategories, as well as individual samples. These two pages serve to help quickly identify foods which should be looked at more carefully due to higher positive rates, as well as to assist in the generation of annual reports.

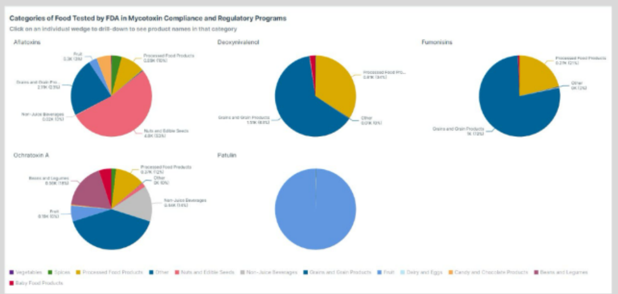
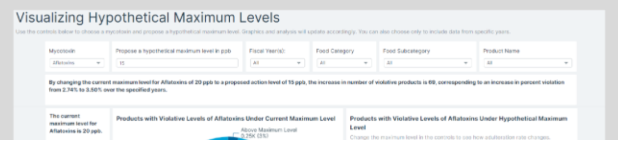


Figure 3. Dashboard visualizations showing the breakdown of food categories tested by mycotoxin.

The last page on the dashboard allows a user to test the effects of hypothetical maximum levels on violation rates. As the level is changed, graphics show how the maximum level affects the violation rate and the food supply. This page was specifically designed to support future work in Codex Alimentarius on setting new maximum levels of mycotoxins in foods.



将来構想：データ運用に向けて

Github

国衛研の研究者の協力が必要

現在、プロジェクト10ヶ月目
継続2年で完成させたい。

現在：データが散在

2～3年

データベース

2～3年

AIなどの活用

将来構想：データ運用に向けて

学術的な成果
・成果の論文化
・国際的なデータ活用に協力

現在：データが散在

2〜3年

データベース

2〜3年

AIなどの活用