

第4回環境研究機関連絡会研究交流セミナー
環境ビッグデータとその活用
2022年12月1日

人工知能学習のための 病害虫被害画像ビッグデータの収集

農研機構農業環境研究部門
岩崎亘典

※ 農研機構（のうけんきこう）、は、国立研究開発法人 農業・食品産業技術総合研究機構のコミュニケーションネーム（通称）、です。

NARO

• 環境ビッグデータの整理

• データ取得とメタデータの付与

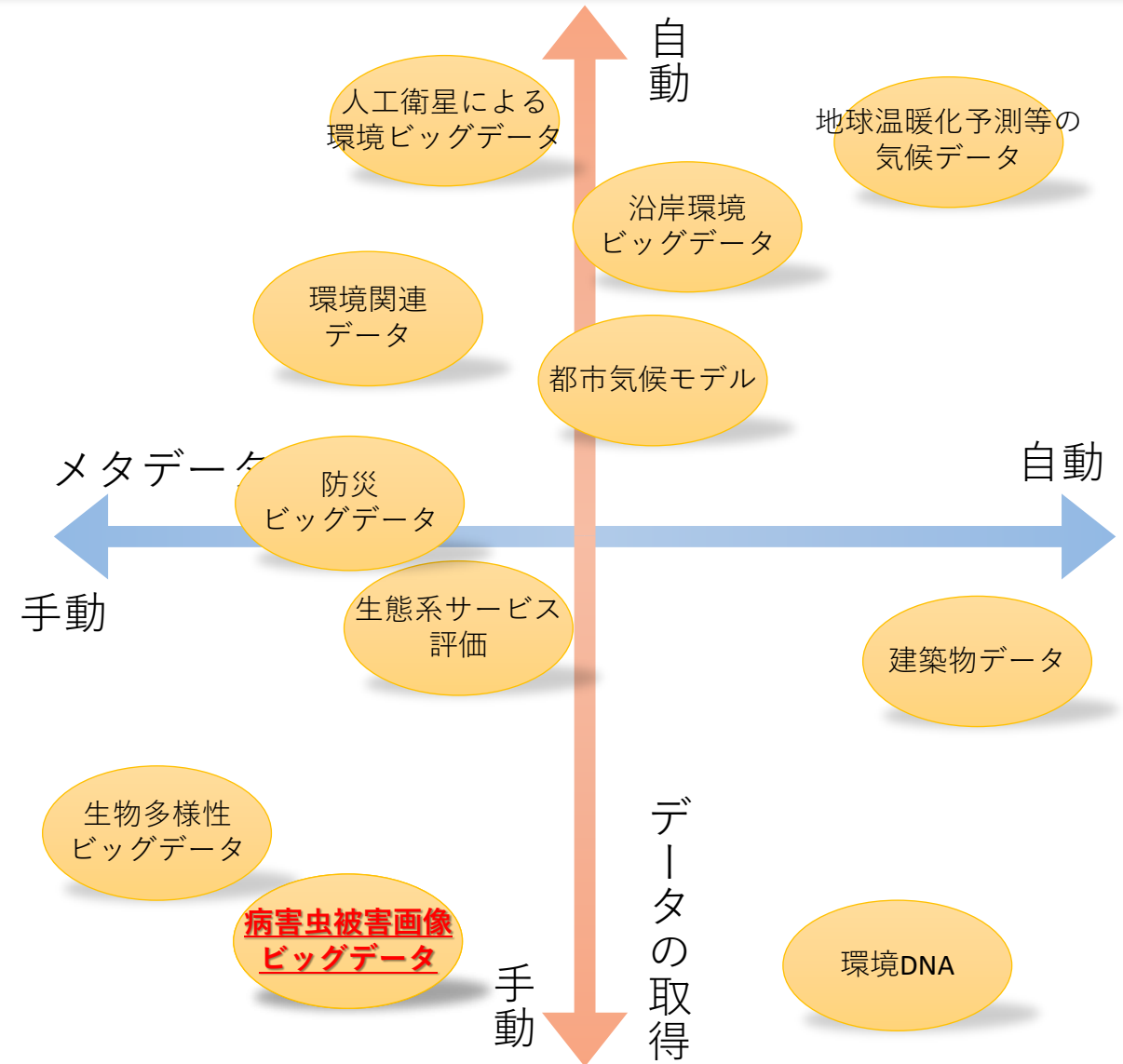
- ビッグデータ活用はメタデータ整備が必要

• メタデータ：データのデータ

- 収集したデータの対象、収集方法、使用機器、単位、ライセンス、etc...

• 必用なメタデータは、データや利用目的により異なる

- 発表者の理解の範囲なので、誤りはご容赦ください



- 農水省委託プロ「AIを活用した病害虫診断技術の開発」(H29～R3)として実施

- 環境の変化

- 多様な環境と営農形態、病害虫被害
- 種苗等の輸入拡大、温暖化等による新規病害虫の増加

- 農業に関わる社会情勢

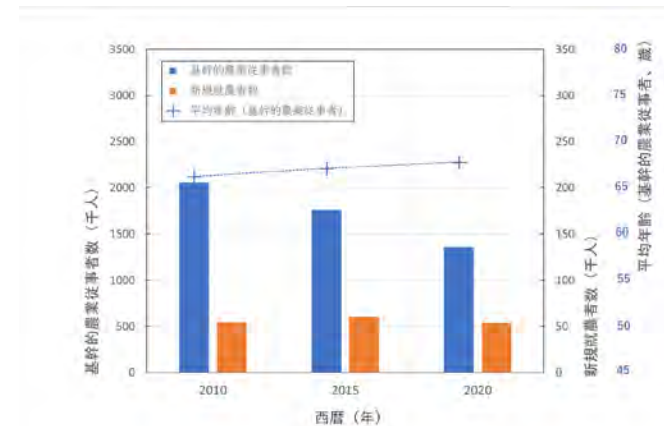
- 農家の高齢化と新規就農者の増加
- 法人経営による非熟練作業者の増加
- 植物防疫担当者の二極化

→ 非熟練者でも可能な、病害虫診断技術の開発が急務

- 深層学習 (Deep Learning) による 病害虫診断AI を開発

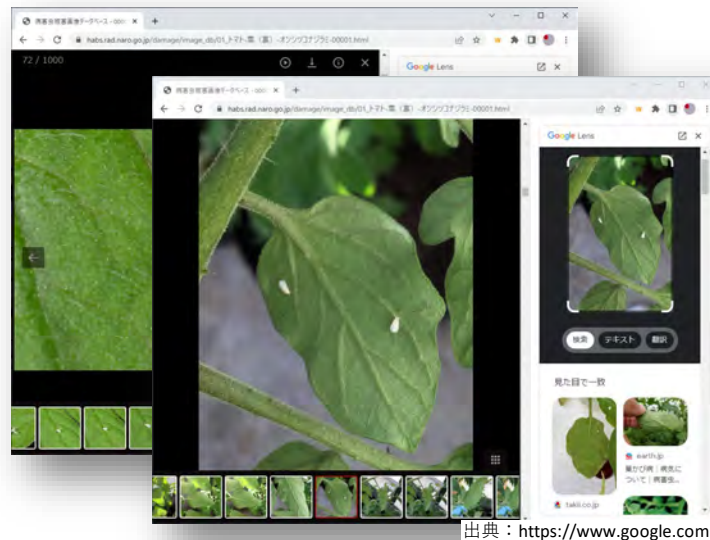


多様な営農形態、環境、病害虫

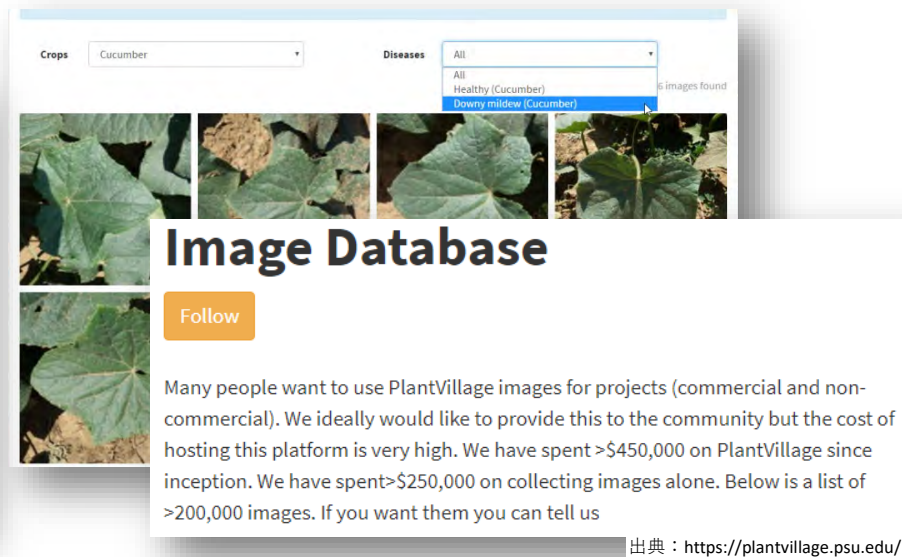


高齢化や新規就農者割合の増加等により、防除技術の円滑な継承が困難

- 病害虫識別AIの開発と病害虫被害ビッグデータの課題
 - 汎用的な画像識別AIでは、病虫害診断が困難
 - 深層学習用画像ビッグデータが整備されていない
 - 対象、枚数が限定的。「図鑑」とは異なる情報の収集が必要
 - 深層学習には大量かつ多様な画像ビッグデータが必要



コナジラミの識別例。一匹であれば識別できるが、二匹になると不可



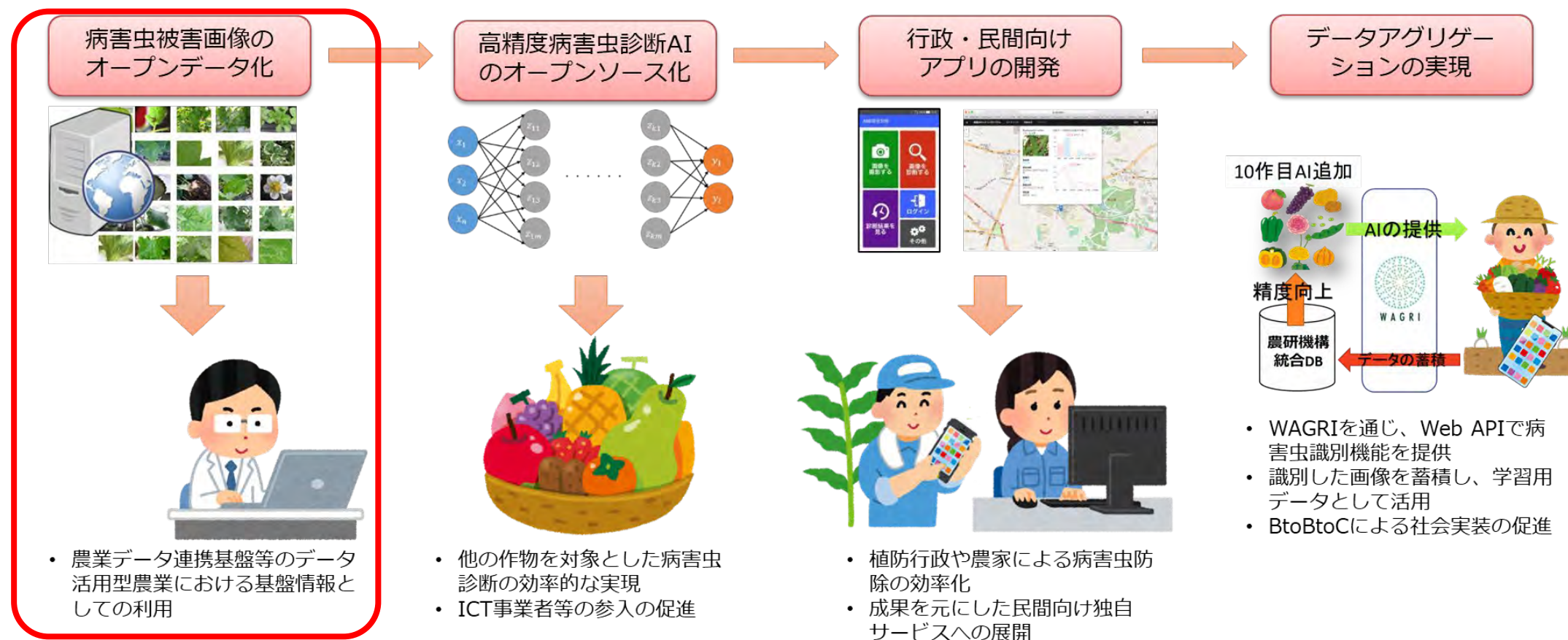
公開されていたデータベースの例。
現在ダウンロードできない。運用コストの問題



既存のWeb図鑑の例。人にとって利用しやすい情報が整備されているが、AIの学習には適さない

プロジェクトの概要：以下の4つのテーマについて実施

- ビッグデータ収集、AIの開発、民生用アプリの開発
- R1年度より内閣府PRIMS予算にて社会実装を促進



-
- | 野菜の種類 | 数量 |
|--------|-------|
| かぼちゃ | 9,991 |
| なす | 847 |
| トマト | 2,182 |
| いちご | 1,617 |
| きゅうり | 1,396 |
| れんこん | 940 |
| ほうれんそう | 1,013 |
| ピーマン | 1,020 |
| ピーナッツ | 1,070 |
| だいこん | 1,140 |
| きんぴら | 1,140 |
| きゅうり | 1,396 |
| いちご | 1,617 |
| トマト | 2,182 |
| なす | 847 |
| かぼちゃ | 9,991 |

調査実施府県

トマト
イチゴ
キュウリ
ナス

0 100 200 300 400 500 km

農研機構撮影

高知県撮影

農研機構撮影

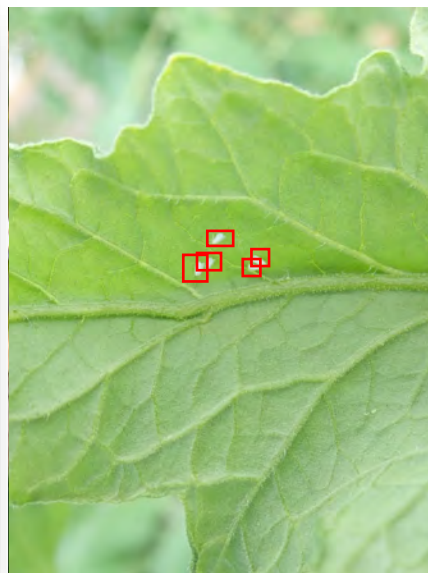
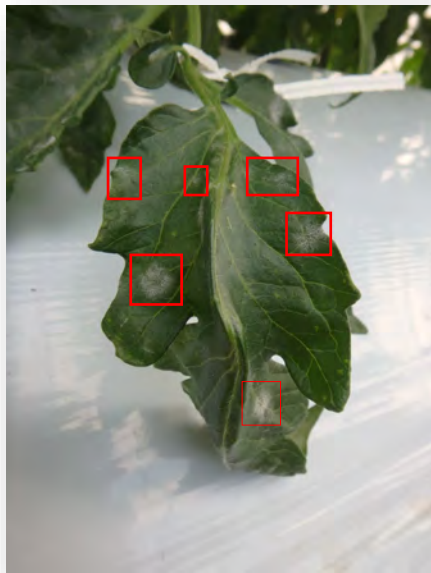
5

- 撮影対象となる 重要病害虫の選定
 - 各地で問題となる重要病害虫を四作目累計で 病害36、虫害34
 - 画像収集対象の決定
 - 被害が発生しやすい病虫害・部位を選定：トマト 49、イチゴ 38、キュウリ 47、ナス 52
 - 虫害については、虫体、被害・生活痕の両方の画像を収集
- 5年間で 累計 71.3 万枚の画像を収集
 - 全てにメタデータを付与する必要がある

作物別の画像収集枚数

	病害	虫害	健全	合計
トマト	98,431	57,752	37,377	196,072
イチゴ	44,686	56,266	51,057	156,851
キュウリ	80,351	49,678	47,379	181,461
ナス	61,157	69,515	43,106	178,779
合計	284,625	233,211	178,919	713,163

- 管理および学習に必要なメタデータを検討
 - 項目が多い方がいい
 - 気象条件、周囲の病虫害の発生、過去の履歴、etc...
 - メタデータではないが、Bounding box (bbox) も学習に有効
 - 多すぎるとメタデータ付与が負担となる
 - 撮影枚数の減少が懸念。必要最低限を選定。



bboxの付与イメージ

学習用画像に付与したメタデータ

メタ情報名	説明	備考
FileName	ファイル名	試験IDと撮影日時等から一意的に生成
Model	撮影カメラ名	
CreateDate	撮影日時	
GPSLatitude	緯度	GPSまたは、QRコードにより付与
GPSLongitude	経度	GPSまたは、QRコードにより付与
CropName	対象作物名	
ExaminationOrganization	試験実施機関名	
ExaminationID	試験ID	接種試験毎に付与し、試験条件を実施機関で記録
ExaminationEnvironment	試験を行う環境。	露地、施設、グロースチャンバー
ShootingPart	撮影対象となる部位	葉（表）、花、果実等
PestDiseaseClassification1	撮影対象の病害、虫害、健全の区別	
PestDiseaseName1	撮影した病害虫名	
PestDiseaseIdentification1	撮影した病害虫診断の正確度	実験室で判定または簡易判定
CropDamageLevel1	病害虫による作物の被害程度	4段階
PestBoddyPresence1	画像内の虫体の有無	
PestBoddyClassification1	虫体が写っている場合の区分	個体または群体
(2～5まで入力可能)		
Copyright	撮影した写真の著作権者	試験実施機関とする。
License	利用許諾条件	標準的ライセンスであるCC BY 4.0を採用
LicenseURL	利用許諾URL	https://creativecommons.org/licenses/by/4.0/deed.ja

人工知能学習のための病害虫被害画像ビッグデータの収集

- メタデータ付与の効率化・自動化
 - 手入力は「必ず」エラーが起こる
 - 表記揺れ、別名表記、省略、etc...
 - ヒューマンエラーは「0」にはならない
 - 手入力を最小限にし、自動化する必要がある

	G	H	I	J
2	病害虫の別	病害虫名	部位	被害程度
1818	虫害	モモアカカブラムシ	葉(裏)	小
1819	虫害	モモアカカブラムシ	全体	小
1820	虫害	モモアカカブラムシ	葉(表)	小
1821	虫害	モモアカカブラムシ	葉(表)	小
1822	虫害	モモアカカブラムシ	葉(表)	小
1823	虫害	モモアカカブラムシ	葉(裏)	小
1824	虫害	モモアカカブラムシ	全体	小
1825	虫害	モモアカカブラムシ	葉(表)	小
1826	虫害	モモアカカブラムシ	葉(表)	小
1827	虫害	モモアカカブラムシ	葉(表)	小
1828	虫害	モモアカカブラムシ	葉(裏)	小
1829	虫害	モモアカカブラムシ	全体	小
1830	健全	健全	全体	無接種・被害なし
1831	健全	健全	葉(表)	無接種・被害なし
1832	健全	健全	葉(表)	無接種・被害なし
1833	健全	健全	葉(表)	無接種・被害なし
1834	健全	健全	葉(裏)	無接種・被害なし

誤字 (モモアカ**カ**ブラムシ)

黄化えそ病(MYSV)

MYSV

表記ゆれ

アザミウマ類

(スリップス類)

別名表記

炭疽病

炭疽病(萎凋症)
炭疽病(葉枯症)

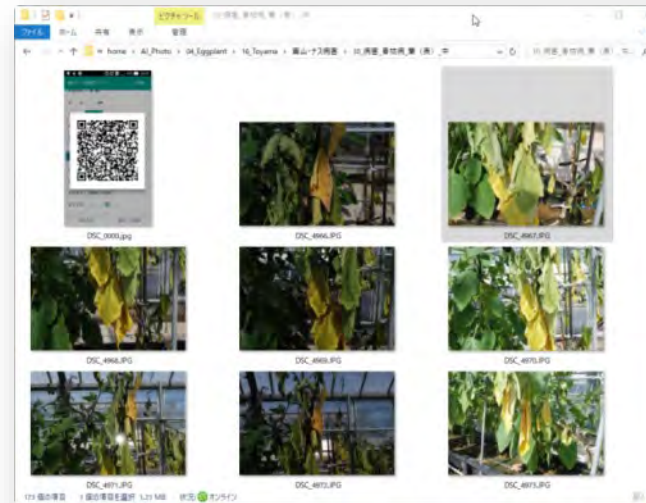
コナジラミ類

オンシツコナジラミ
タバココナジラミ

区分できない

プロジェクト初期に見られエラーの例

- 専用アプリを用いメタデータの付与
 - データの手入力を避け、人為的ミスを減らす

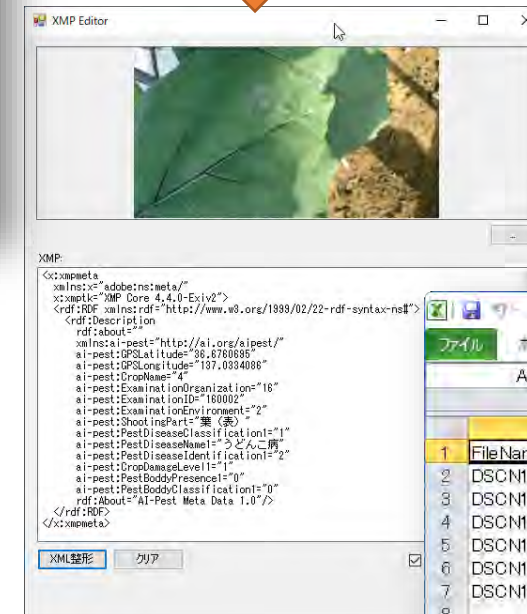


QRコードを画像と
同じフォルダに配置

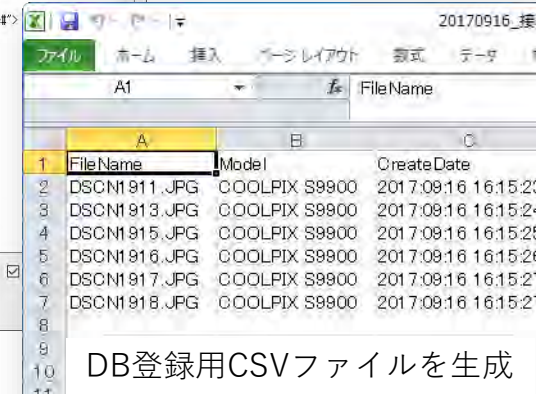
Androidアプリでメタデータを含
むQRコードを生成



QRコードをメタデータに画像に付与
するスクリプトを生成



XMP領域にメタ
データを付与



File Name	Model	Create Date
DSCN1911.JPG	COOLPIX S9900	2017:09:16 16:15:20
DSCN1913.JPG	COOLPIX S9900	2017:09:16 16:15:24
DSCN1915.JPG	COOLPIX S9900	2017:09:16 16:15:28
DSCN1916.JPG	COOLPIX S9900	2017:09:16 16:15:26
DSCN1917.JPG	COOLPIX S9900	2017:09:16 16:15:27
DSCN1918.JPG	COOLPIX S9900	2017:09:16 16:15:27

DB登録用CSVファイルを生成

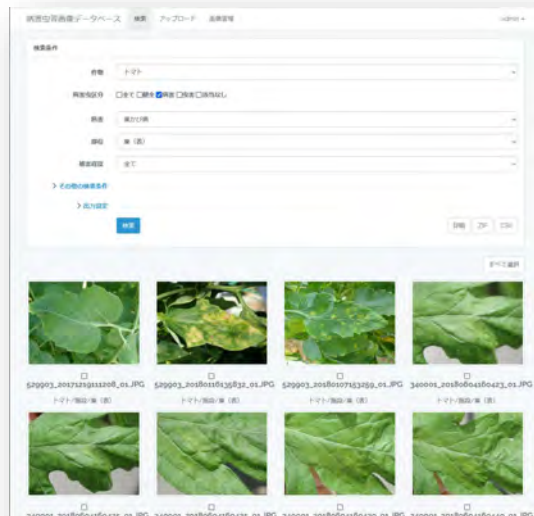
• 被害画像ビッグデータの管理/公開

- 画像管理用DB：内部向け。プロジェクト期間のみ運用
- 公開用ページ：外部向け。静的Webページとして公開
 - ライセンスはCC BY 4.0 国際

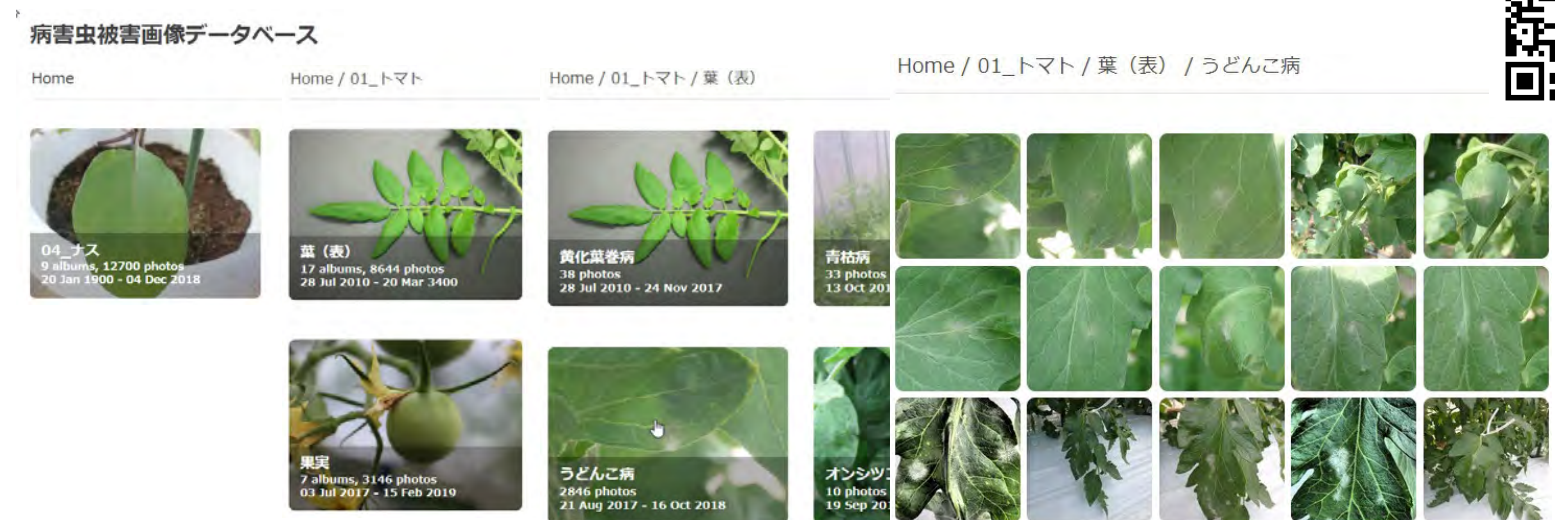
• “as is”での公開

- 既存ツールでフォルダ単位のデータをWebページへ変換
- システムの寿命とデータの寿命
 - 持続可能な運用を優先
 - 公開にコストをかけない
- <https://habs.rad.naro.go.jp/damage/#!index.md>

病害虫被害画像管理DB



病害虫被害画像公開用Webページ



- 11

• 病害虫被害画像収集の課題

• 新規収集画像のビッグデータ化

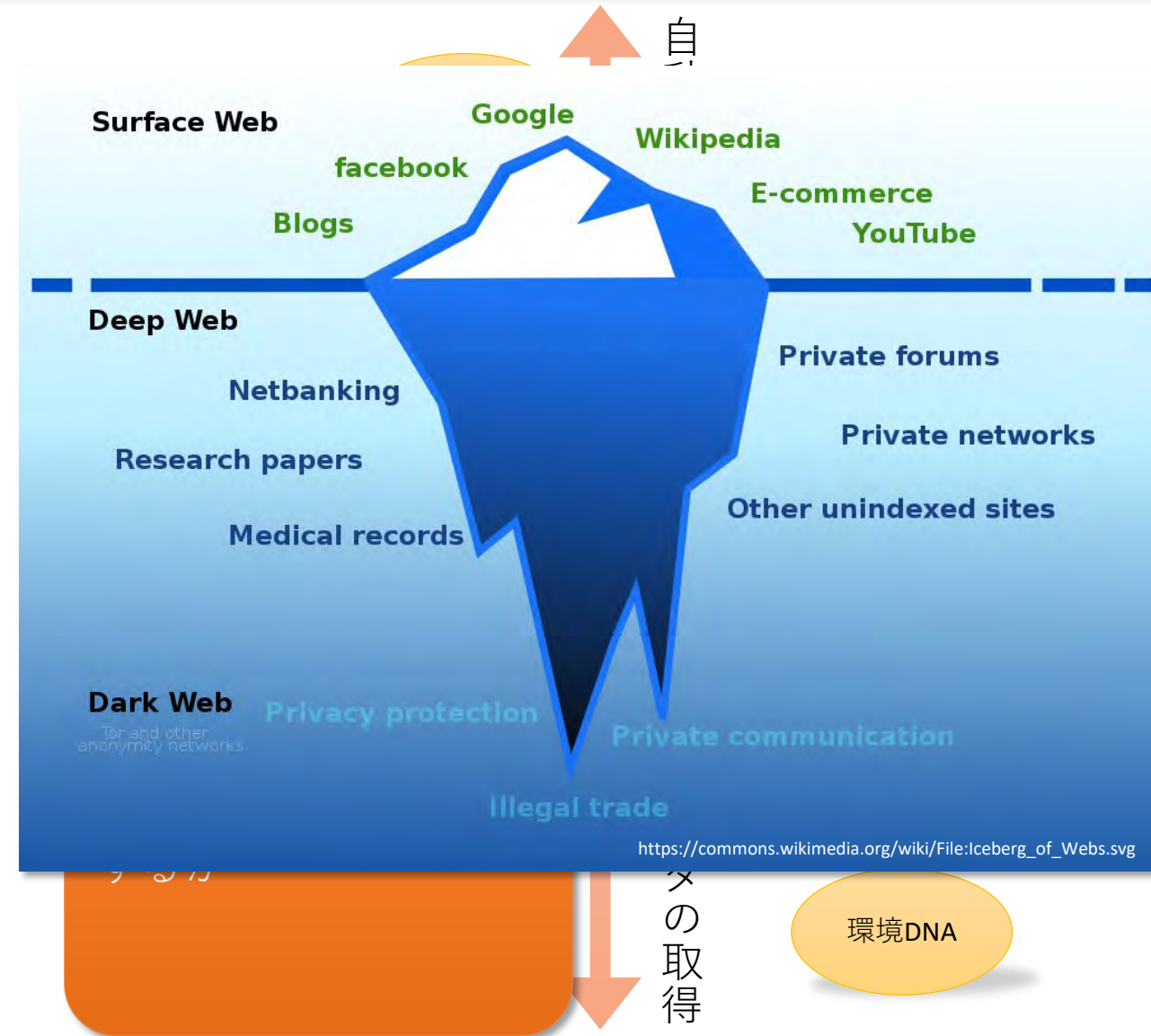
- PRISM課題にてシステムを構築
- 農業データアグリゲーションスキーム

• 過去の蓄積のビッグデータ化

- 大学、行政、研究機関、個人等により所有
- 利用できない大量のデータが存在
 - 「Deep Web」と同様
 - 現在は「Surface Web」領域のみ

• 多分野で共通の課題

- データが遺失する前の取り組みが必要
- 分野横断的検討が必要



- 謝辞

- 本研究は、農林水産省委託プロジェクト研究「A I を活用した病害虫診断技術の開発」JP17935051の補助を受けて行った。
- 本研究は、官民 研究開発投資拡大プログラム（PRISM） 予算 を活用 して 行った。

- 以下の参画機関の皆様のご協力、ご尽力にも、心から感謝します。

- 岩手県農業研究センター、宮城県農業・園芸総合研究所、福島県、茨城県農業総合センター、栃木県農業試験場、群馬県農業技術センター、埼玉県、新潟県農業総合研究所、富山県、山梨県、長野県野菜花き試験場、岐阜県農業技術センター、静岡県農林技術研究所、愛知県、三重県、京都府農林水産技術センター、兵庫県立農林水産技術総合センター、広島県立総合技術研究所、香川県、高知県、長崎県、宮崎県総合農業試験場、鹿児島県農業開発総合センター、国立大学法人名古屋大学、学校法人法政大学、株式会社ノーザンシステムサービス、株式会社N T T データ、株式会社N T T データ C C S、日本農薬株式会社、農研機構・植物防疫研究部門、野菜花き研究部門、農業環境研究部門