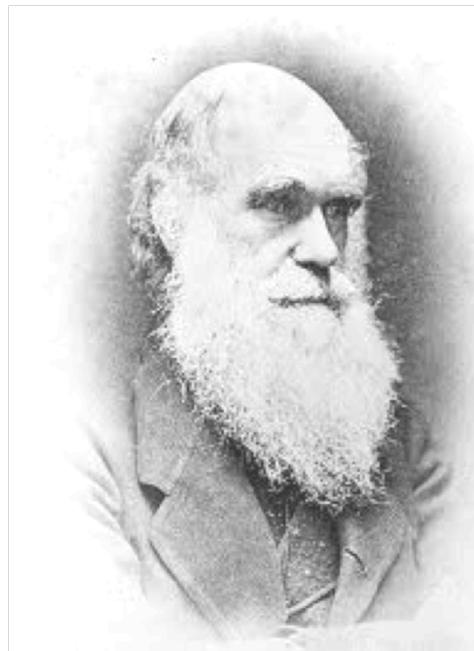


Darwin Core extension for genebanks



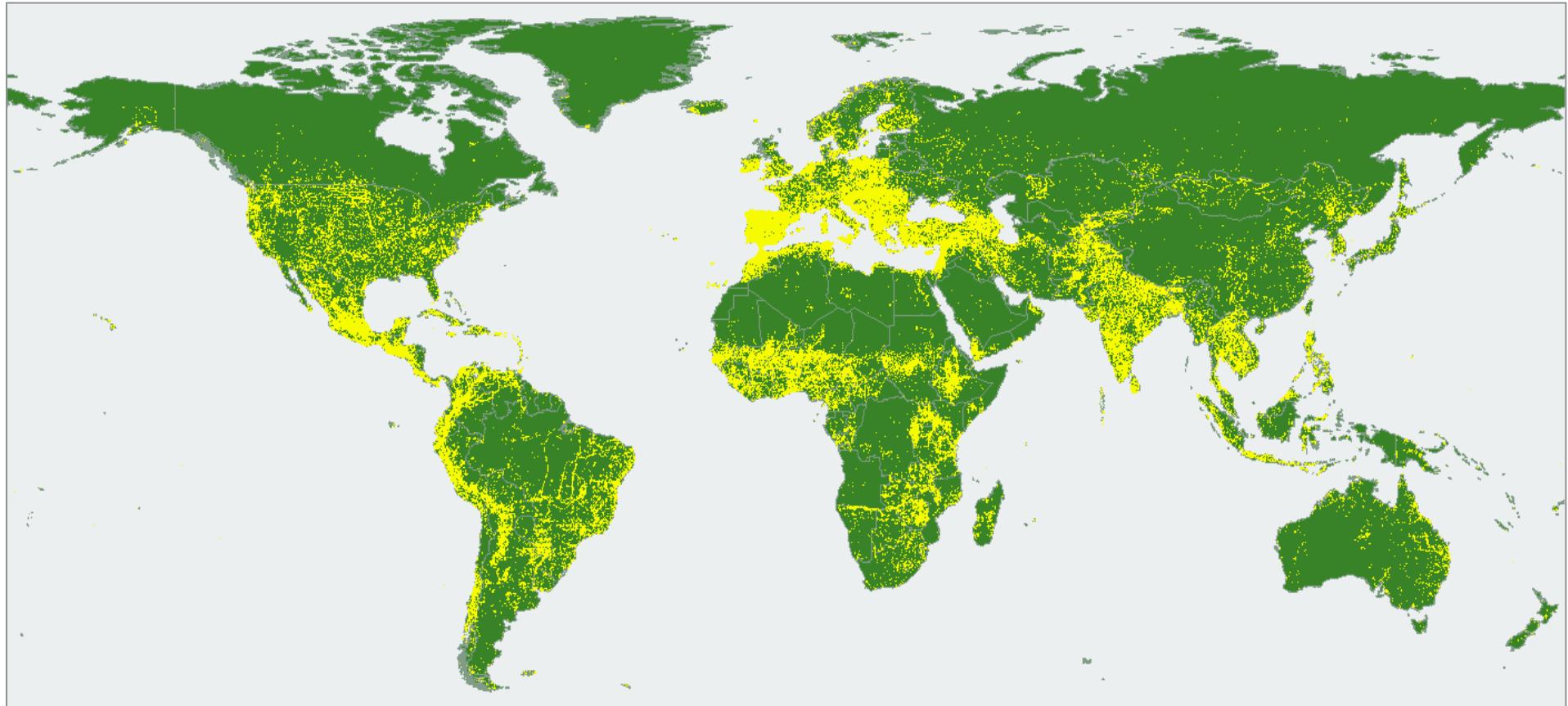
Hi!
/

Hi!
\\



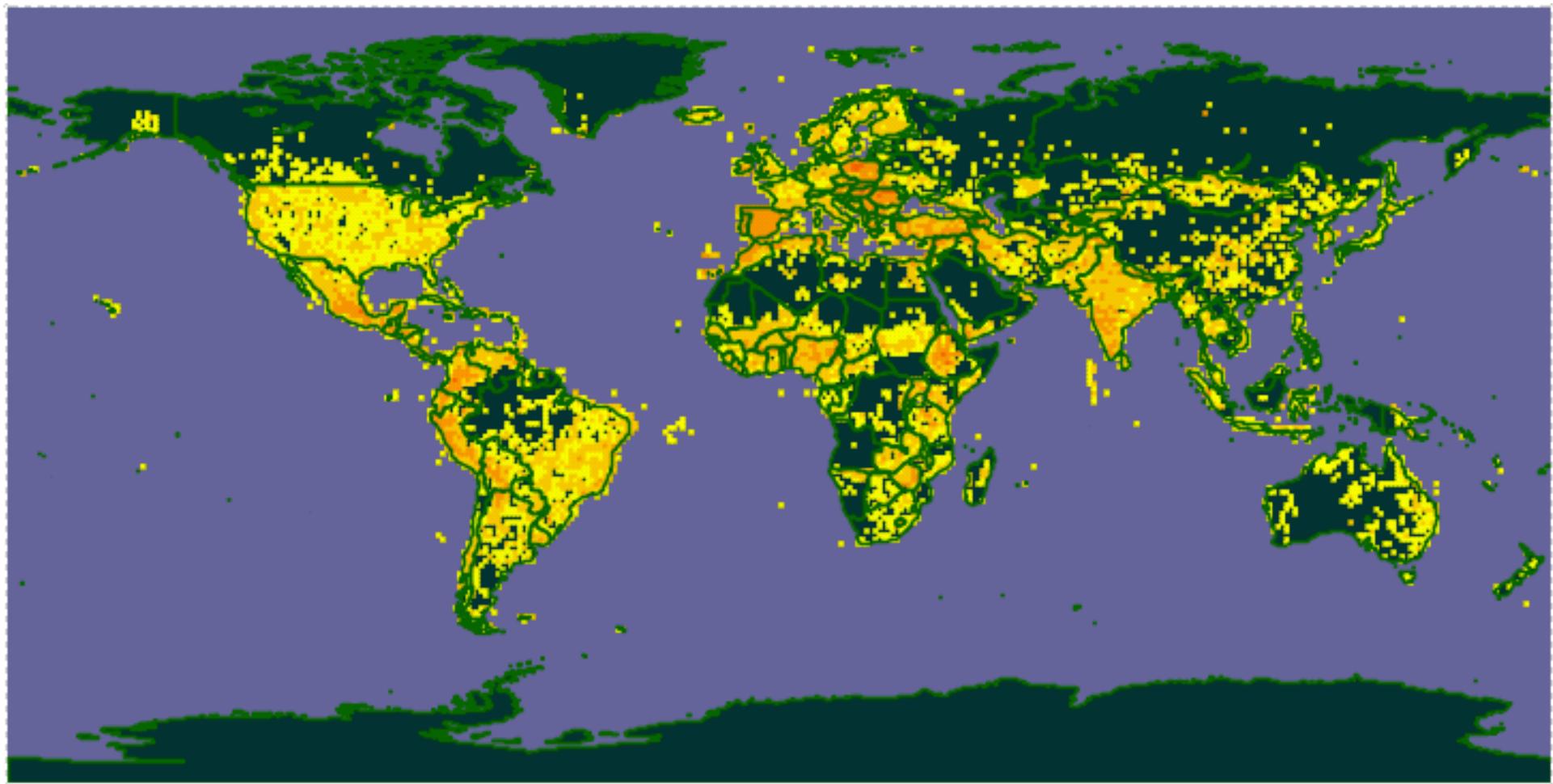


genesys-pgr.org



The GENESYS gateway to genetic resources provides access to information on more than 2.3 million genebank accessions, <http://www.genesys-pgr.org/>

Potential of the GBIF technology

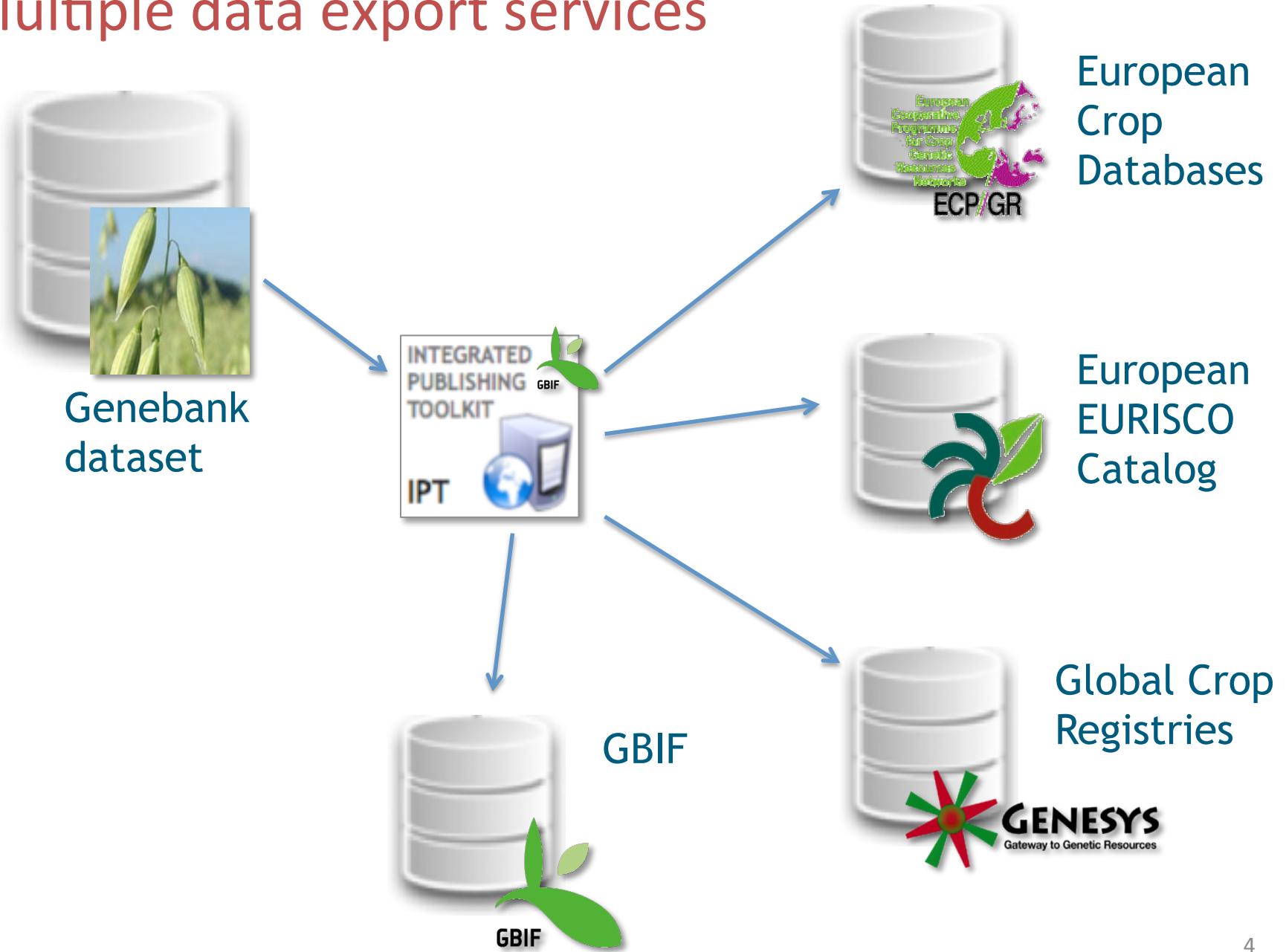


This map only shows records with coordinates (**552,772** records from a total of **2,040,886** records).

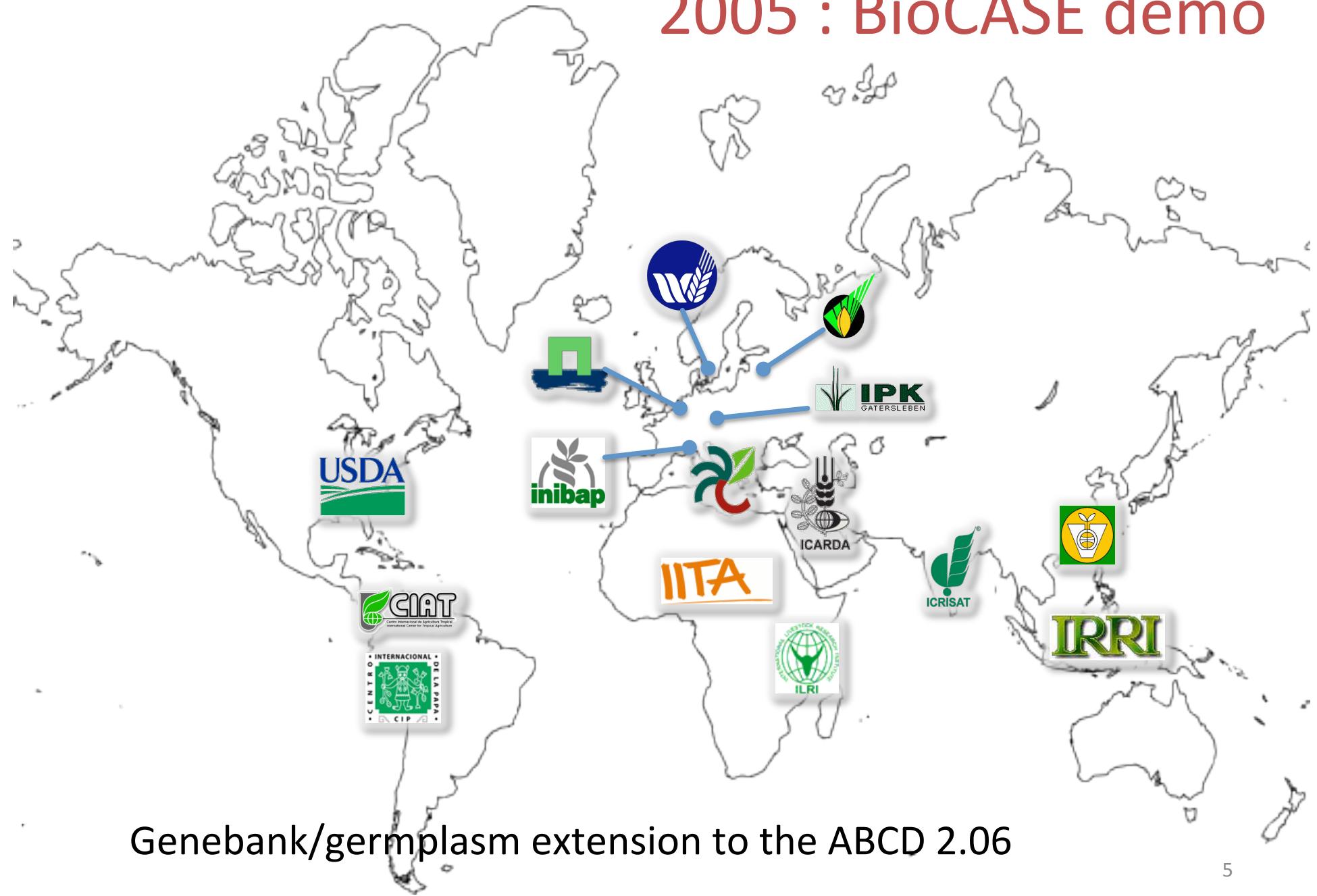
Disclaimer: Maps depict density of data registered within the GBIF network index and not necessarily true species occurrence density gradients. It includes records from all datasets shared by this data publisher.

<http://data.gbif.org/datasets/network/2>

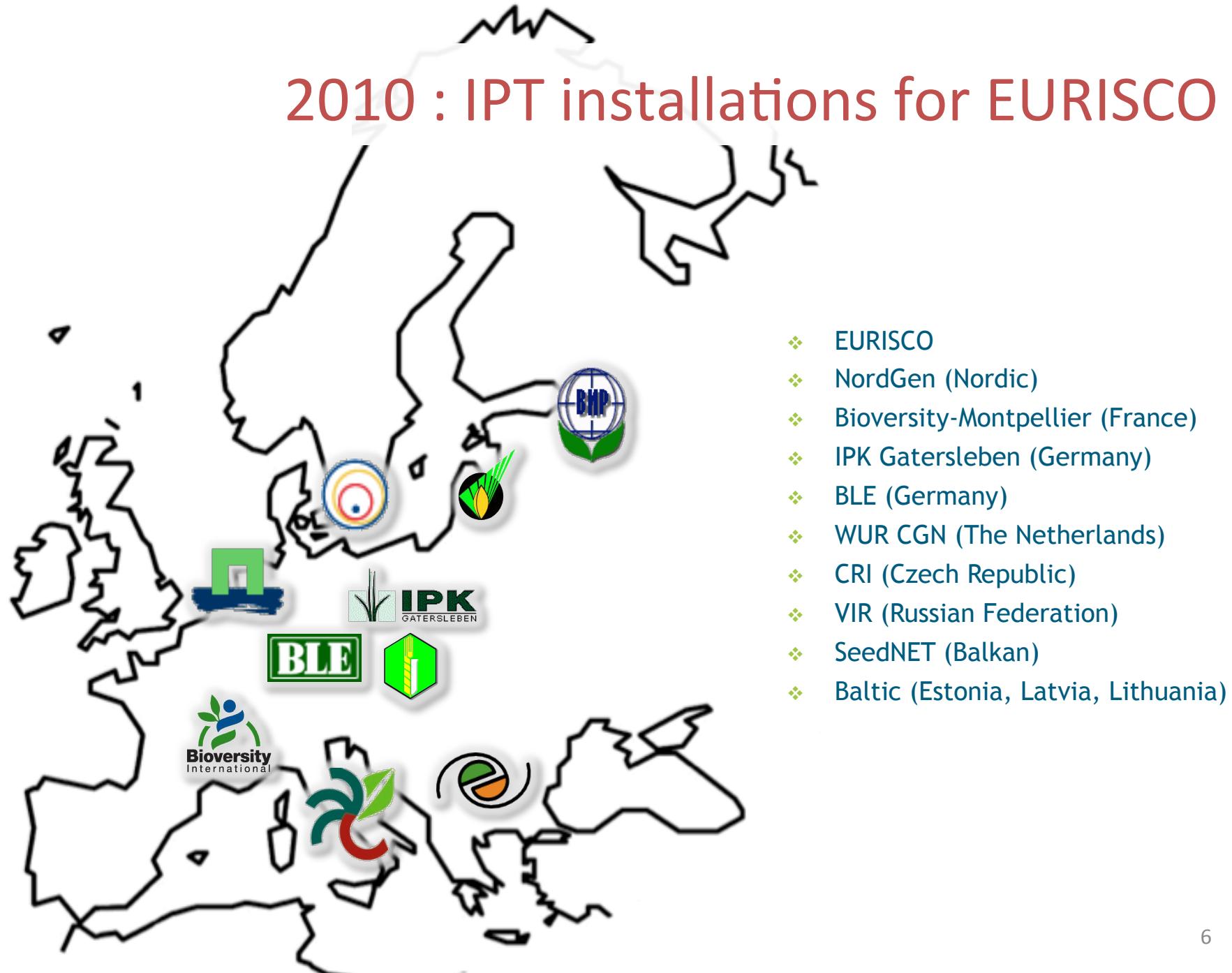
Multiple data export services



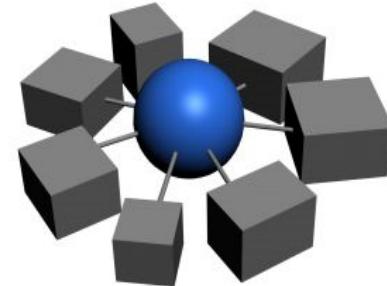
2005 : BioCASE demo



2010 : IPT installations for EURISCO



Darwin Core



The purpose of DwC terms is to facilitate data sharing

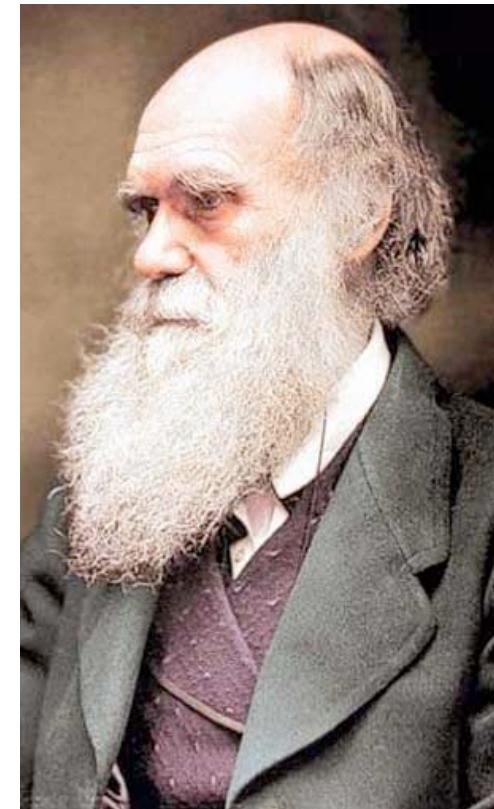
- a well-defined standard core vocabulary
- a flexible framework to maximize re-usability
- approved as TDWG standard 2009

"The Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, and samples, and related information."

<http://rs.tdwg.org/dwc/>

The Darwin Core can be extended by new terms to share additional information.

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7(1): e29715. doi: 10.1371/journal.pone.0029715



Darwin Core extension for genebanks



DwC Germplasm : DRAFT 0.1 : August 26, 2009

- “MCPD in Darwin Core”
- Additional terms to describe germplasm samples
- Includes terms from the breeding/cultivation event
- Includes additional terms for crop trait experiments
- Includes terms for international crop treaty regulations



<http://code.google.com/p/darwincore-germplasm>

<http://rs.nordgen.org/dwc/> (draft version)

<http://purl.org/germplasm/terms#> (coming soon)

FAO/IPGRI MULTI-CROP PASSPORT DESCRIPTORS
December 2001

The list of multi-crop passport descriptors (MCPD) is developed jointly by IPGRI and FAO to provide international standards to facilitate germplasm passport information exchange. These descriptors aim to be compatible with IPGRI crop descriptor lists and with the descriptors used for the FAO World Information and Early Warning System (WIEWS) on plant genetic resources (PGR).

For each multi-crop passport descriptor, a brief explanation of content, coding scheme and *suggested* fieldname (in parentheses) is provided to assist in the computerized exchange of this type of data. It is recognized that networks or groups of users may want to further expand this MCPD List to meet their specific needs. As long as these additions allow for an easy conversion to the format proposed in the multi-crop passport descriptors, basic passport data can be exchanged worldwide in a consistent manner.

General comments:

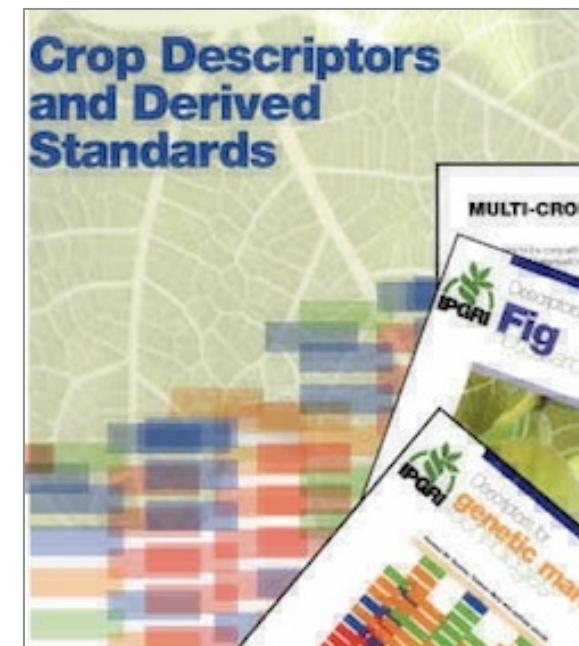
- If a field allows multiple values, these values should be separated by a semicolon (;) without space(s), (i.e. Accession name:Rheinische Vorgebirgstrauben;Emma;Avlon).
- A field for which no value is available should be left empty (i.e. Elevation). If data are exchanged in ASCII format for a field with a missing numeric value, it should be left empty. If data are exchanged in a database format, missing numeric values should be represented by generic NULL values.
- Dates are recorded as YYYYMMDD. If the month and/or day are missing this should be indicated with hyphens. Leading zeros are required (i.e. 197506-, or 1975--).
- Latitude and longitude are recorded in an alphanumeric format. If the minutes or seconds are missing, this should be indicated with hyphens. Leading zeros are required.
- Country names: Three letter ISO codes are used for countries. The ISO 3166-1: Code List and the Country or the Country or area numerical codes added or changed are not available on-line, but can be obtained from IPGRI [ipgri-mcpd@cgiar.org]
- For institutes the codes from FAO should be used. These codes are available from <http://apps3.fao.org/wiews/> for registered WIEWS users. From the Main Menu select: 'PGR' and 'Download'. If new Institute Codes are required, they can be generated online by national WIEWS administrators, or by the FAO WIEWS administrator [Stefano.Diulgheroff@fao.org].
- The preferred language for free text fields is English (i.e. Location of collecting site and Remarks).

MULTI-CROP PASSPORT DESCRIPTORS

1. Institute code	(INSTCODE)
Code of the institute where the accession is maintained. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a number. The current set of Institute Codes is available from the FAO website (http://apps3.fao.org/wiews/).	
2. Accession number	(ACCENUMB)
This number serves as a unique identifier for accessions within a genebank collection, and is assigned when a sample is entered into the genebank collection.	
3. Collecting number	(COLLNUMB)
Original number assigned by the collector(s) of the sample, normally composed of the name or initials of the collector(s) followed by a number. This number is essential for identifying duplicates held in different collections.	

Alercia, A., S. Diulgheroff, T. Metz (2001). FAO/IPGRI Multi-crop passport descriptors, December 2001. International Plant Genetic Resources Institute (IPGRI) / Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

Available at http://apps3.fao.org/wiews/mcpd/MCPD_Doc2001_EN.pdf



DwC Germplasm (1)

GermplasmDataset

- dwc:datasetID [\[1\]](#) ↗
- NICODE (missing term) - Discussion at: [NationalInventoryCode](#)
- UPLOADERCODE (epgris3-trait:UPLOADERCODE)
- DATASET_REMARKS (epgris3-trait:DATASET_REMARKS)



Taxon

- dwc:scientificNameID [\[8\]](#) ↗
- dwc:genus [\[9\]](#) ↗ (mcpd:GENUS)
- dwc:specificEpithet [\[10\]](#) ↗ (mcpd:SPECIES)
- dwc:infraspecificEpithet [\[11\]](#) ↗ (mcpd:SUBTAXA)
- dwc:scientificNameAuthorship [\[12\]](#) ↗ (mcpd:SPAUTHOR)* (mcpd:SUBAUTHOR)*
- dwc:vernacularName [\[13\]](#) ↗ (mcpd:CROPNAME)

DwC Germplasm (2)

GermplasmSample

- dwc:occurrenceID [2]  (epgris3-trait:GENOTYPE_NUMBER)
- GermplasmID (redundant by dwc:occurrenceID) [3] 
- dwc:catalogNumber [4]  (mcpd:ACCENUMB)
- GermplasmIdentifier (mcpd:ACCENAME)
- dwc:institutionCode [5]  (mcpd:INSTCODE)
- BiologicalStatusOfSample
- BiologicalStatusOfSampleCode (mcpd:SAMPSTAT)
- TypeOfStorage (mcpd:STORAGE)
- dwc:otherCatalogNumbers [6]  (mcpd:OTHERNUMB)
- dwc:occurrenceRemarks [7]  (mcpd:REMARKS)



SampleAcquisition

- SampleAcquisitionID
- DonorInstitute (eurisco:DONORDESCR)
- DonorInstituteCode (mcpd:DONORCODE)
- DonorsSampleIdentifier (mcpd:DONORNUMB)
- SampleAcquisitionDate (mcpd:ACQDATE)
- SampleAcquisitionSource (mcpd:COLLSRC)
- SampleAcquisitionRemarks

DwC Germplasm (3)

CollectingEvent

- dwc:recordNumber [14] ↗
- CollectingInstituteCode (mcpd:COLLCODE)
- dwc:recordNumber [15] ↗ (mcpd:COLLNUMB)
- dwc:decimalLatitude [16] ↗ (mcpd:LATITUDE)
- dwc:decimalLongitude [17] ↗ (mcpd:LONGITUDE)
- dwc:minimumElevationinMeters [18] ↗ (mcpd:ELEVATION)
- dwc:eventDate [19] ↗ (mcpd:COLLDATE)
- dwc:locality [20] ↗ (mcpd:COLLSITE)
- dwc:countryCode [21] ↗ (mcpd:ORIGCTY)
- dwc:eventRemarks [22] ↗



BreedingEvent

- BreedingEventID
- BreedersSampleIdentifier (mcpd:ACCENAME)**
- BreederInstitute (eurisco:BREDDDESCR)
- BreederInstituteCode (mcpd:BREDCODE)
- BreederPerson
- BreedingCountry
- BreedingCountryCode
- BreedingYear
- AncestralData (mcpd:ANCEST)
- PurdyPedigree
- BreedingEventRemarks

DwC Germplasm (4)

LocationOfSafetyDuplication

- SafetyDuplicationID
- SafetyDuplicationDate
- SafetyDuplicationInstitute
- SafetyDuplicationInstituteCode (mcpd:DUPLSITE)



GermplasmTreatiesAndRegulations

- GermplasmRegulationID
- TreatyOrRegulationName
- TreatyOrRegulationGoverningBody



DwC Germplasm (5)

GermplasmTrait

- GermplasmTraitID
- GermplasmTraitIdentifier
- GermplasmTraitClass
- dwc:measurementType [23] ↗
- dwc:measurementMethod [24] ↗
- GermplasmTraitSource
- GermplasmTraitScale
- GermplasmTraitRemarks

GermplasmTraitExperiment

- GermplasmExperimentID
- GermplasmExperimentIdentifier
- GermplasmExperimentReport
- GermplasmExperimentYear
- GermplasmExperimentRemarks



GermplasmTraitMeasurement

- dwc:measurementID [25] ↗
- dwc:measurementValue [26] ↗
- dwc:measurementDeterminedBy [27] ↗
- MeasurementByInstituteCode (*redundant by new term dwc:measurementDeterminedBy*)
- MeasurementGrowthStage



Germplasm vocabulary of terms (RDF/SKOS)

...

```
<skos:Concept rdf:about="http://rs.nordgen.org/dwc/germplasm/0.1/terms/SampleAcquisitionID">
  <skos:prefLabel xml:lang="en">donors accession ID</skos:prefLabel>
  - <skos:definition xml:lang="en">
    An identifier for the source germplasm material that was included in your genebank collection.
  </skos:definition>
  - <skos:example xml:lang="en">
    http://data.gbif.org/occurrences/164390781/; URI; DOI; LSID or similar
  </skos:example>
  - <skos:note xml:lang="en">
    Local identifiers used by the donor is best reported using the DonorsSampleIdentifier term.
  </skos:note>
  <skos:scopeNote rdf:resource="http://code.google.com/p/darwincore-germplasm/wiki/SampleAcquisitionID"/>
  <dcterms:issued>2009-08-24</dcterms:issued>
  <dcterms:modified>2012-03-27</dcterms:modified>
  <vann:termGroup rdf:resource="http://rs.nordgen.org/dwc/germplasm/0.1/terms/AcquisitionEvent"/>
  <rdfs:isDefinedBy rdf:resource="http://rs.nordgen.org/dwc/germplasm/0.1/terms/" />
  <skos:inScheme rdf:resource="http://rs.nordgen.org/dwc/germplasm/0.1/terms/" />
  <vs:term_status>unstable</vs:term_status>
  <dcterms:isReplacedBy rdf:resource="http://purl.org/germplasm/terms#donorsID"/>
  <dcterms:source rdf:resource="http://apps3.fao.org/wIEWS/mCPD/MCPD_Dec2001_EN.pdf"/>
  <rdfs:range rdf:resource="dcterms:identifier"/>
  <skos:relatedMatch rdf:resource="dwc:otherCatalogNumbers"/>
</skos:Concept>
```

...

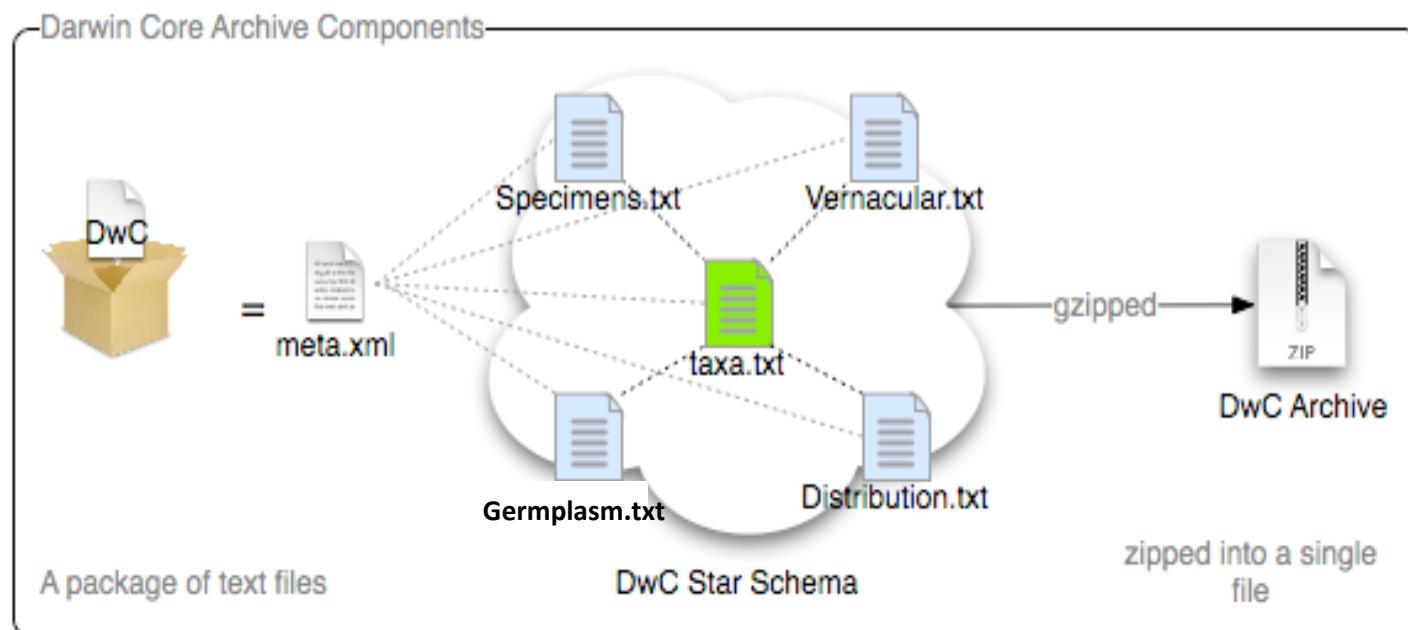
http://rs.gbif.org/sandbox/terms/germplasm/germplasm_01.rdf

<http://purl.org/germplasm/> (in preparation)

<http://kos.gbif.org/wiki/Germplasm> (wiki forum)

Darwin Core Archive (DwC-A)

- ❖ DwC-A publish dwc records including extensions
- ❖ Simple text based format
- ❖ Zipped single file archive



Darwin Core Archive extension (XML)

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="/style/human.xsl"?>
<extension xmlns="http://rs.gbif.org/extension/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dc="http://purl.org/dc/terms/"
  xsi:schemaLocation="http://rs.gbif.org/extension/ http://rs.gbif.org/schema/extension.xsd"
  dc:title="Germplasm (0.1b)"
  dc:description="The Darwin Core extension for genebanks (DwC germplasm) was developed in 2009 for sharing genebank datasets with t
  dc:relation="http://rs.nordgen.org/dwc/"
  name="GermplasmAccession"
  namespace="http://rs.nordgen.org/dwc/germplasm/0.1/terms/"
  rowType="http://rs.nordgen.org/dwc/germplasm/0.1/terms/GermplasmAccession">

<!--

Application schema for the GBIF IPT data provider software
based on the Darwin Core extension for genebanks (DwC germplasm) version 0.1
Dag Endresen (NordGen), 2009-05-15, 2009-08-21, 2009-08-26, 2011-10-06, 2012-03-14

The Multi-Crop Passport (MCPD) standard was published jointly by FAO and IPGRI (now Bioversity International) in 2001,
and is available online from http://www.bioversityinternational.org/index.php?id=19&amp;user_bioversitypublications_pil[showUid]=219
--&gt;

&lt;!-- GermplasmAccession --&gt;
&lt;property name="GermplasmID"
  namespace="http://rs.nordgen.org/dwc/germplasm/0.1/terms/"
  qualName="http://rs.nordgen.org/dwc/germplasm/0.1/terms/GermplasmID"
  dc:description="An identifier for the genebank accession. Could be an unique and stable identifier specific to the genebank coll
  type="string"
  examples="URI, DOI, LSID or similar; http://data.gbif.org/occurrences/215969415/"
  dc:relation=""
  group="GermplasmAccession"
  required="true"/&gt;
&lt;property name="BiologicalStatusOfSample"
  namespace="http://rs.nordgen.org/dwc/germplasm/0.1/terms/"
  qualName="http://rs.nordgen.org/dwc/germplasm/0.1/terms/BiologicalStatusOfSample"
  dc:description="(MCPD term 20). Biological status of sample. Recommended to follow a regulated vocabulary, or preferable the SAM
  examples="Wild; Crop wild relative (CWR); Primitive; Traditional cultivar/landrace; Breeding line; Genetic stock; Modern cultiva
  type="string"
  dc:relation=""
  group="GermplasmAccession"
  required="false"/&gt;
&lt;property name="BiologicalStatusOfSampleCode"
  namespace="http://rs.nordgen.org/dwc/germplasm/0.1/terms/"
  qualName="http://rs.nordgen.org/dwc/germplasm/0.1/terms/BiologicalStatusOfSampleCode"
  thesaurus="http://rs.gbif.org/vocabulary/germplasm/0.1/BiologicalStatusOfSample.xml"
  dc:description="MCPD term 20. Biological status of sample, SAMPSTAT. SAMPSTAT codes distinguish wild and primitive cultivars fro
  examples="100; 110; 120; 200; 300; 400; 410; 411; 412; 413; 414; 415; 420; 500; 999"
  type="integer"
  ...</pre>
```

<http://rs.gbif.org/extension/nordgen/0.1/germplasm.xml>



Germplasm Accession

[View](#) [Workflow](#)

Field	Content
Description	The GermplasmAccession is an application schema or IPT extension to describe genebank accessions. This "extension" includes terms for the unit or rowType of the Genebank Specimen (Accession).
Name	GermplasmAccession
Namespace	http://rs.nordgen.org/dwc/germplasm/0.1/terms/
Web Service	
Publisher	Nordic Genetic Resources Center (NordGen)

Properties

A list of properties associated with this extension.

Group: AcquisitionEvent

Name	Namespace	Datatype	Required	Vocabulary
SampleAcquisitionID	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
SampleAcquisition	http://rs.nordgen.org/dwc/germplasm/0.1/terms/		false	
SampleAcquisitionDate	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	Date	false	
DonorInstitute	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
DonorInstituteCode	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
DonorsSampleIdentifier	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
SampleAcquisitionSource	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	Integer	false	
SampleAcquisitionRemarks	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	

Group: BreedingEvent

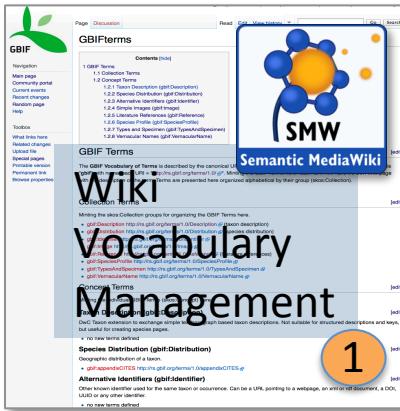
Name	Namespace	Datatype	Required	Vocabulary
BreedingEventID	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
BreedersSampleIdentifier	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
BreederInstitute	http://rs.nordgen.org/dwc/germplasm/0.1/terms/	String	false	
BreederInstituteCode	http://rs.nordgen.org	String	false	

<http://vocabularies.gbif.org/node/163947>

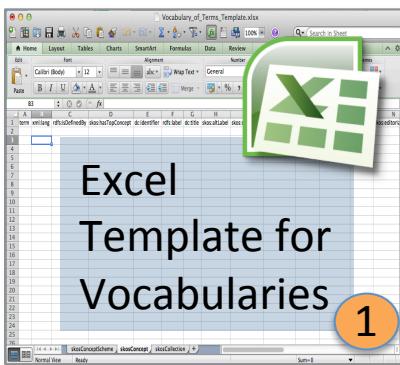
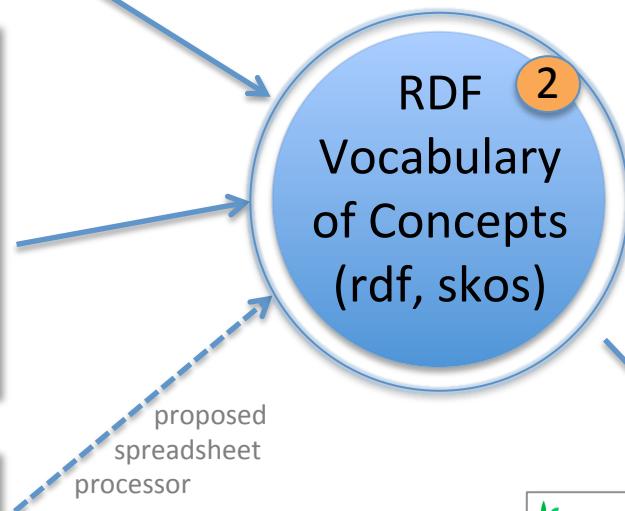
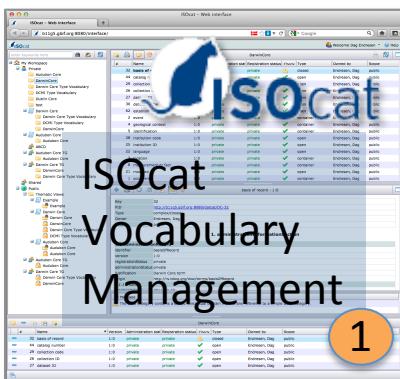
User login

Username: *

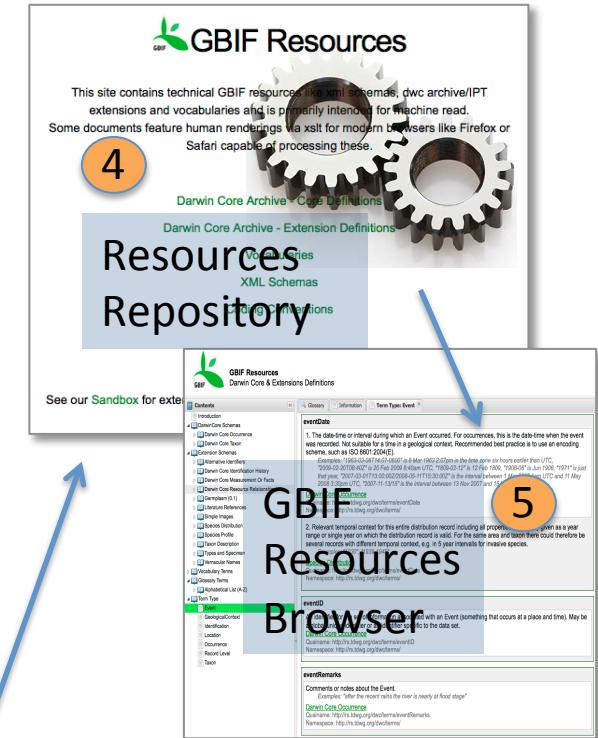
Password: *



1. Mint and maintain concepts and terms, in domain-expert working groups.
2. Release final version as a **RDF Vocabulary**.
3. **REUSE** terms from published RDF vocabularies and ontologies when designing new DwC-A extensions, controlled value vocabularies (and new Ontologies).
4. Publish at the GBIF Resources Repository.
5. Browse at the GBIF Resources Browser.

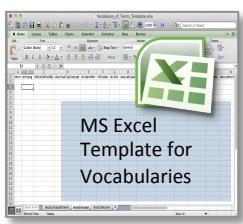
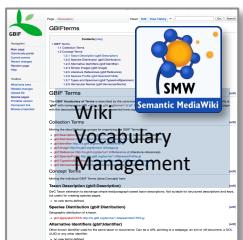


Collaborative management tools



GBIF Vocabularies as a collaborative management tool for Darwin Core Archive extensions and controlled vocabularies.

Evaluation of various tools for collaborative management of RDF vocabularies.



proposed
spreadsheet
processor

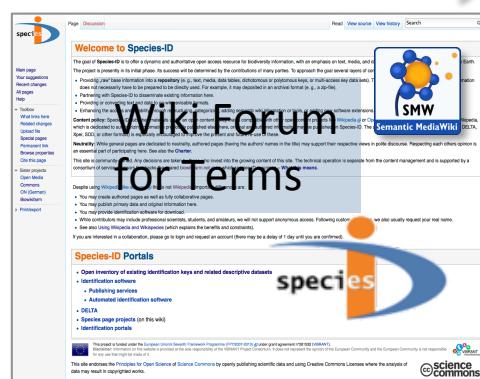
RDF Vocabulary of Concepts (rdf, skos)

GBIF Resources Repository



See our [Sandbox](#) for extensions & vocabularies under development. Click to change.

DwC-A Extensions & Controlled vocabularies



Wiki forum for terms as an open community platform for description of new and (reused) existing terms.



Vocabulary Management

Vocabulary Management for the description of biodiversity information resources

 [Subscribe to feed](#) [Report this](#) [Bookmark this](#) [Share |](#) [Edit group](#)[Invite users](#)[Membership approval](#)[Mail Members](#)[Group discussion](#)[Group pages](#)[Group files](#)[Group albums](#)[Group blog](#)[Group bookmarks](#)

Group members



Vocabulary Management



Description:

The Vocabulary Management Task Group (VoMaG) will address the best practices for the collaborative development and maintenance of terms for the description of biodiversity information resources. The biodiversity vocabulary management routines contributes to the development of a Knowledge Organization System for biodiversity resources. VoMaG plan to establish itself as a Task Group under the Technical Architecture Group (TAG) of the Biodiversity Information Standards (TDWG).

Brief description: Vocabulary Management for the description of biodiversity information resources

Tags: vocabulary, vocabulary of terms, terms, vomag, tdwg

Website: <http://kos.gbif.org>

<http://community.gbif.org/pg/groups/21382/>

Latest discussion

No topics have been created.

Group files



Biodiversity Knowledge Organization System: Proposed Architecture; Version 0.2 February 2012

Dag Endresen 97 days ago, Comments (1)



Vocabulary Management Task Group Charter A Task Group of the TAG Interest Group, Draft version February 2012

Group pages



About VoMaG

Last updated 50 days ago by Dag Endresen



Species interactions

Last updated 56 days ago by Dag Endresen

[More pages](#)