



The Biodiversity  
Information  
Standards (TDWG):  
Opportunities for  
Collaboration and  
Participation



**cni**

Coalition for Networked Information



Biodiversity  
Information  
Standards  
T D W G

Martin R. Kalfatovic  
TDWG Fundraising and Partnerships  
Subcommittee, Chair

Contributions from:

- Wouter Addink  
TDWG Europe Regional  
Representative
- Dimitris Koureas  
TDWG Chair
- James A Macklin  
TDWG Deputy Chair

12 December 2017  
Washington, DC

---

# TDWG Mission



Image by : [opensource.com](https://opensource.com)

**Develop, adopt and promote standards** and guidelines for the recording and exchange of data about organisms

**Promote the use** of standards through the most appropriate and effective means and

Act as a **forum for discussion** through holding meetings and through publications

# What TDWG does ...

TDWG Standards assist biodiversity organizations and research institutions around the world to:

- ▶ Increase data collection and collation efficiency
- ▶ Enhance institutional profiles through Internet discovery
- ▶ Further research, conservation and management goals by raising data visibility and integration
- ▶ Enhance opportunities for collaboration on international biodiversity projects
- ▶ Cost-effectively manage and share natural history data
- ▶ Minimize 'reinvention of the wheel'
- ▶ Provide a stimulating environment for the development of leading edge Internet tools and techniques

# What is a standard?

- ▶ In common English:
  - *A flag*
  - *An upright pole or beam*
  - *A backing for currency*
  - *American automobile*
  - *A bush on a long stalk*
  - *An ideal to be judged against*
  - *Model of authorisation of excellence*
  - *A basis for comparison*
  - *1,980 board feet of wood*
  - *A newspaper*
  - *An established norm*

A technical **standard** is an established norm or requirement in regard to technical systems. It is usually a formal document that establishes uniform engineering or technical criteria, methods, processes and practices. (Wikipedia)

# TDWG Standards Categories

## **Technical specification**

*Protocol, service, procedure, format*

## **Applicability statement**

*How a technical specification might be applied*

## **Best current practice**

*A description for good behavior*

## **Data standard**

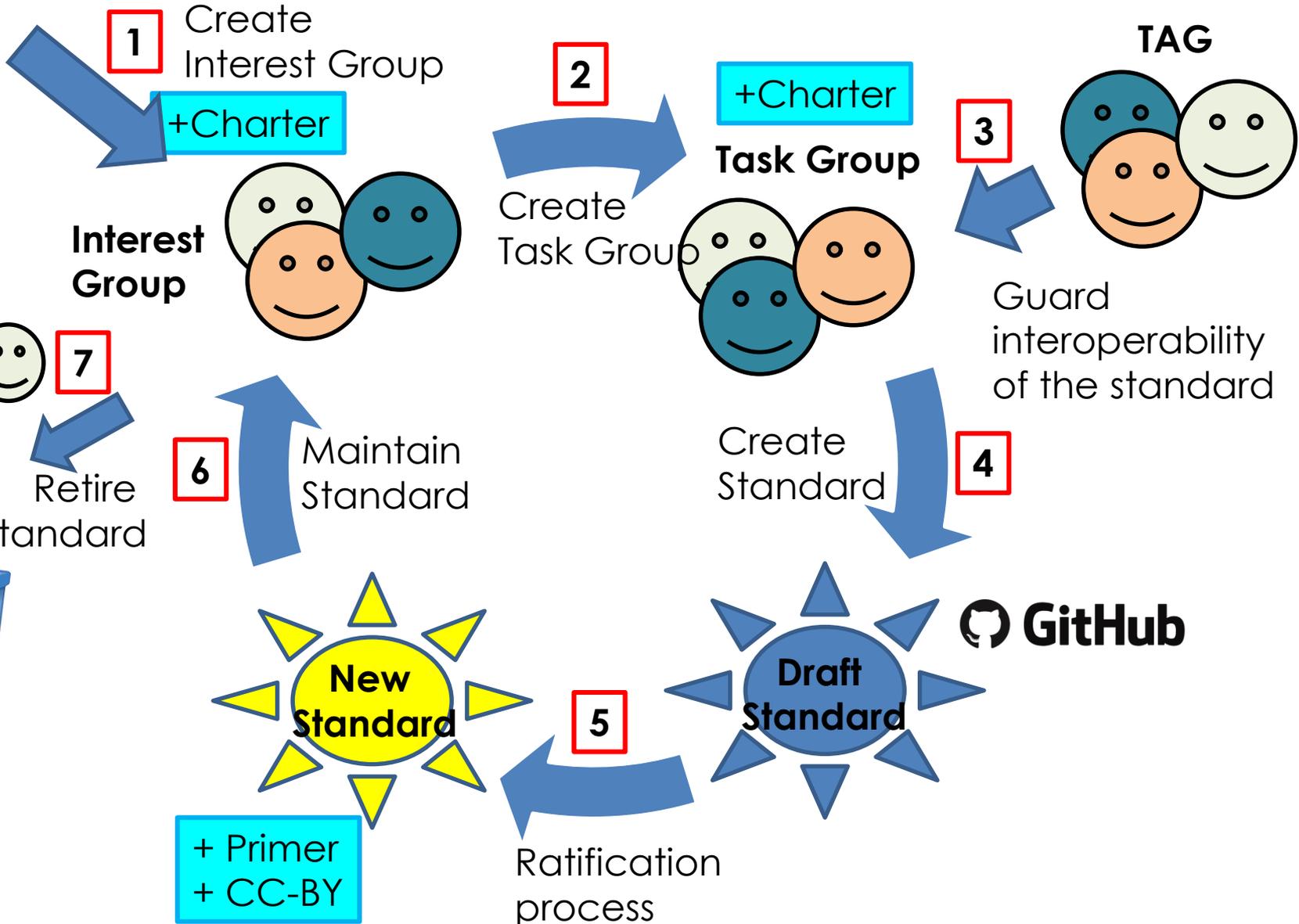
*Content or controlled vocabularies*

# TDWG Standards Process

▶ TDWG standards are:

- **Consensus**
- **Community wide** (+/-)
- **Voluntary** (But we have a TDWG Community Support Fund to support working meetings or dissemination of results)

# How we create a standard

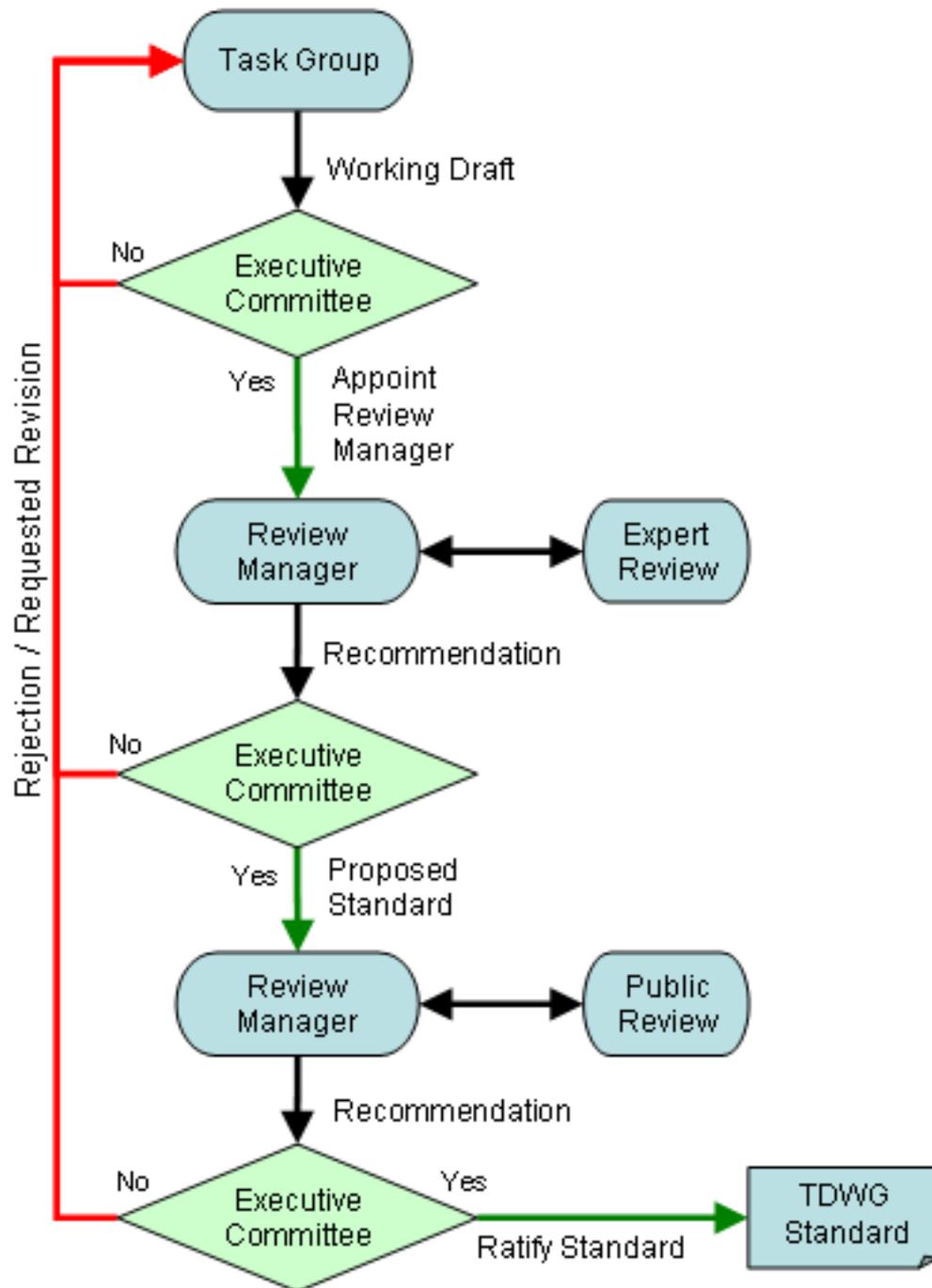


# Group forming and collaboration

- ▶ Interest Groups or Task Groups are formed by creating a **Charter**, that defines the convenor (first contact), core members, motivation, guidance how to become involved and a summary of what standard or guideline the group wants to achieve.
- ▶ **Github** (<https://github.com/tdwg>) is the main collaboration platform (adopted in 2014) and used to host and version the standards. TDWG also provides a few mailinglists for communication

Current TDWG groups:  
<https://github.com/orgs/tdwg/teams>

# TDWG Standard Ratification Process



# TDWG Standard Statuses

- ▶ **Current Standard**
  - Standards that TDWG **recommends for use**.
- ▶ **Current Standard (2005)**
  - Standards that TDWG **recommends for use**, but not (yet) compliant with the standards development process adopted in 2006.
- ▶ **Draft Standard**
  - Standard under review
- ▶ **Prior Standard**
  - Standards that were ratified prior to 2005 and that are not currently being promoted for ratification under the post 2006 ratification process
- ▶ **Retired Standard**
  - Standards that have been ratified by TDWG in the past but that are **no longer recommended**.

# Current TDWG Standards

- ▶ **Darwin Core** (DwC) - occurrences
- ▶ **Access to Biological Data v2.6** (ABCD) - occurrences
- ▶ **Audubon Core** – media metadata
- ▶ **TDWG Access Protocol for Information Retrieval** (TAPIR) – exchange protocol
- ▶ **Structured Descriptive Data** (SDD) – taxonomic descriptions
- ▶ **Taxonomic Concept Transfer Schema** (TCS) – taxon concepts
- ▶ **Delta Format** – taxonomic descriptions

# ...and Standards to support the standardisation process...

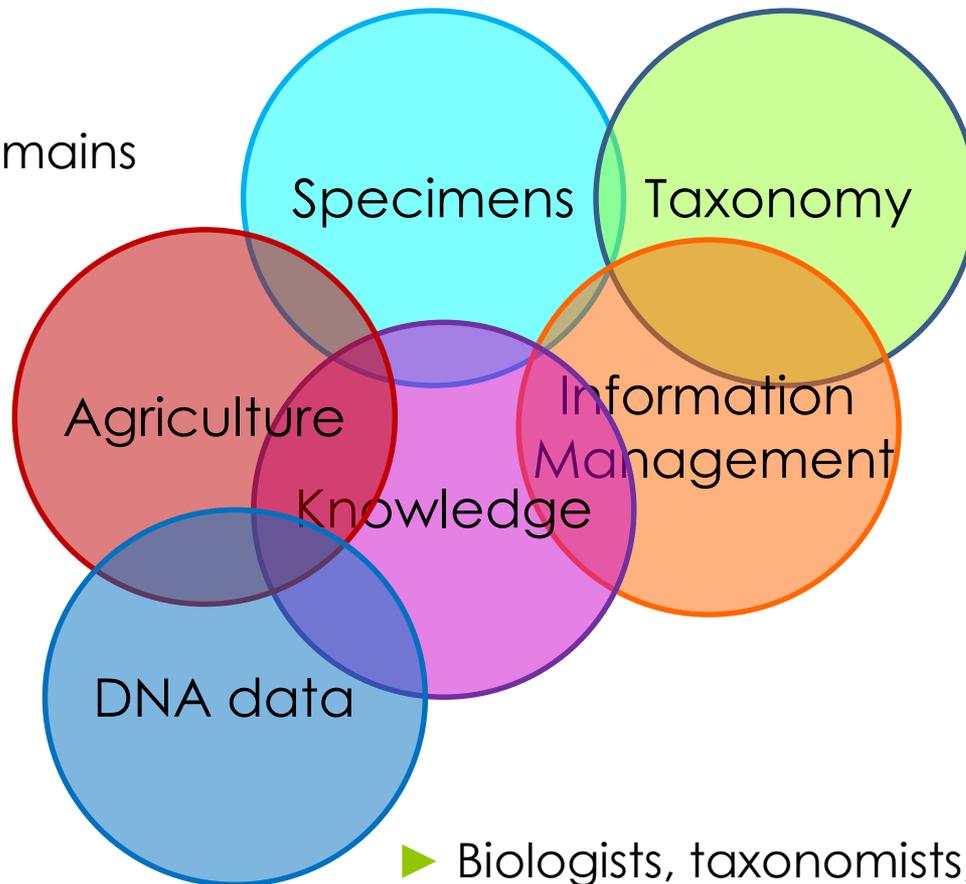
- ▶ **TDWG Standards Documentation Standard**
- ▶ **Vocabulary Maintenance Standard**

# Who is the TDWG community?



# Who is the TDWG community?

- Intersection of many domains



- Biologists, taxonomists, computer scientists, each with something to offer to each other's domains

# Who is the TDWG community really?

- ▶ Frustrated taxonomists
  - *Looking for a better way*
  - *Largely self taught*
- ▶ Bored computer scientists
  - *Looking for excitement, challenge*
- ▶ Misfits and visionaries
  - *In search of a 'Brave New World'*
- ▶ Egomaniacs
  - *In search of glory, fame, power, riches*

# Members and Users

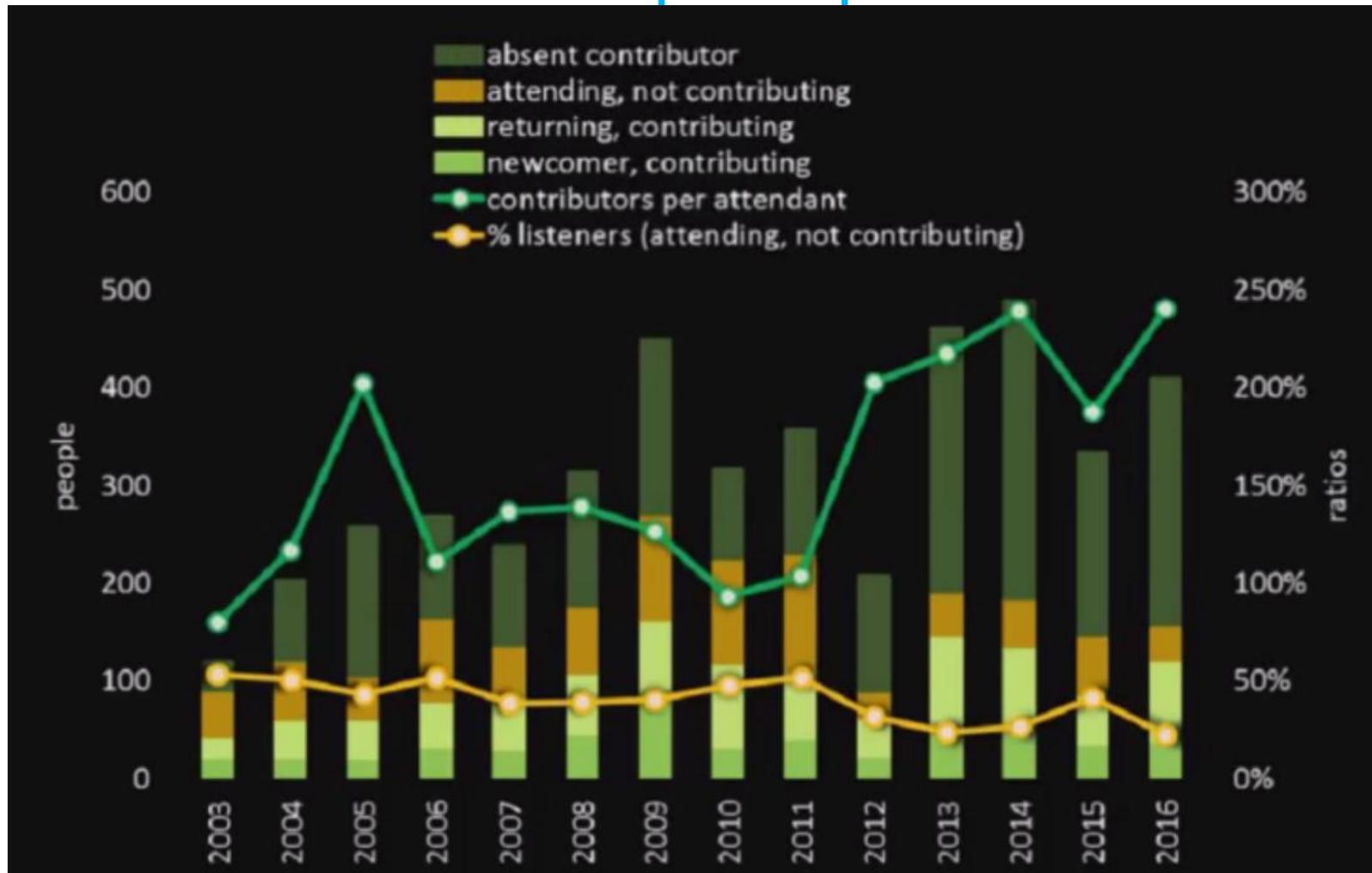
*Our members are a diverse mix of biologists, computer scientists, bioinformaticians, taxonomists, ecologists, librarians, geoscientists...*

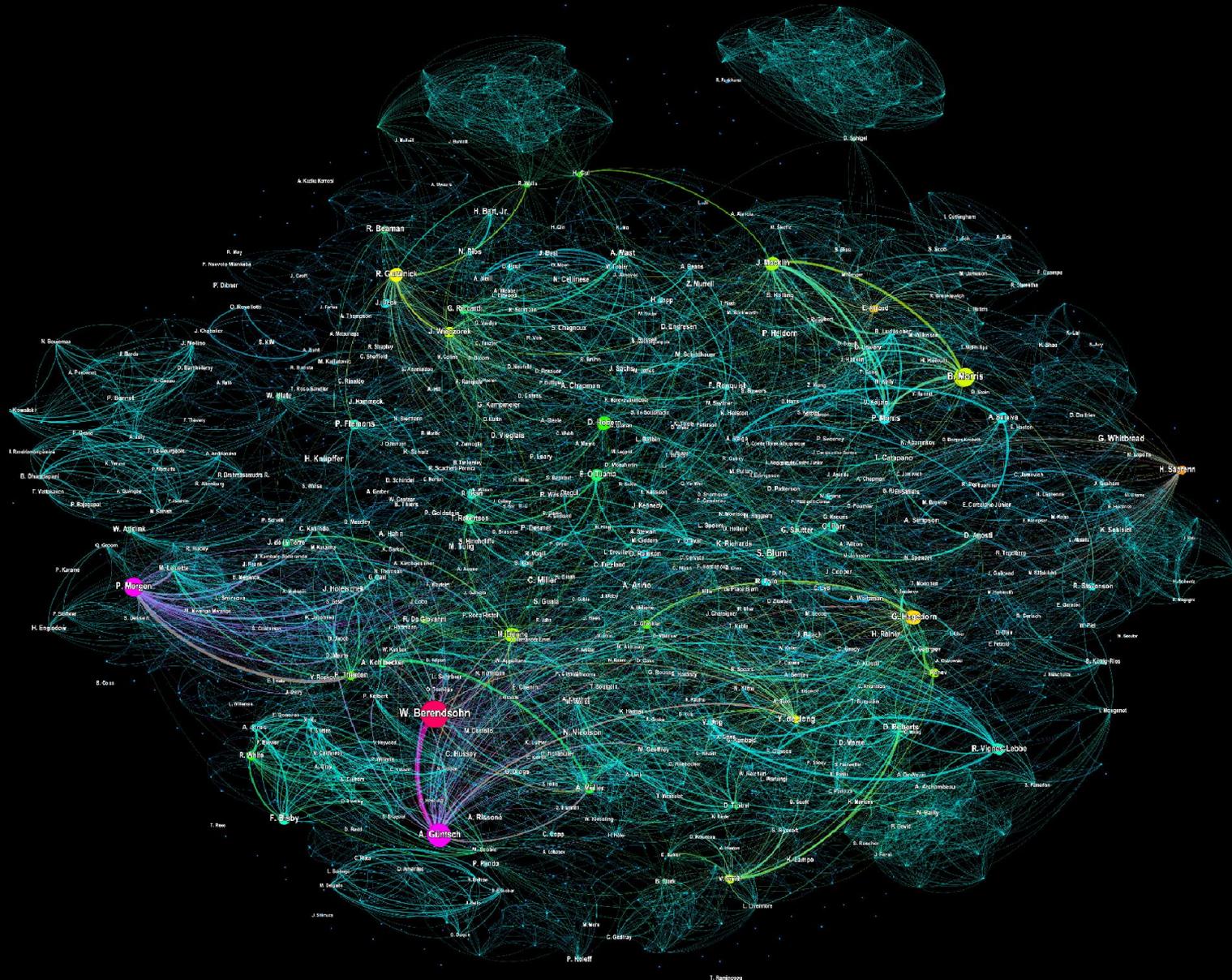
**125+ individual members**  
**20 institutional members**



Our standards are used by museums, botanical gardens, research departments, government agencies and organizations who work with any type of biological data.

# TDWG: A growing community of highly collaborative people





# Interest Groups

- ▶ Biodiversity Data Quality
  - ▶ -Framework
  - ▶ -Tools, Services & Workflows
  - ▶ -Use Case Library

## Biodiversity Informatics Curriculum

- ▶ Annotations, Attribution

## Natural Collections

- ▶ Descriptions

## Genomic Biodiversity , Phylogenetics

- ▶ Imaging

- ▶ Invasive Species

## Biological Descriptions, Species information

- ▶ Literature

- ▶ Taxon Names & Concepts

- ▶ Observation and Specimen Records

- ABCD
- DarwinCore (DwC)
- Observations(OBS)

- ▶ Citizen Science

- ▶ Paleobiology

# TDWG's History

early history

- ▶ A not-for-profit organization that started as the **Taxonomic Database Working Group** in **1985**, Frank Bisby as its first chair.
- ▶ Initially formed to establish international collaboration among biological database projects, to promote the wider and more effective dissemination of information about the world's heritage of biological organisms.
- ▶ It was accepted as a Commission of the International Union of Biological Sciences (**IUBS**) in October 1988
- ▶ Initially focusing on plant taxonomic databases, it expanded its scope to over all taxonomic databases in 1994.

recent

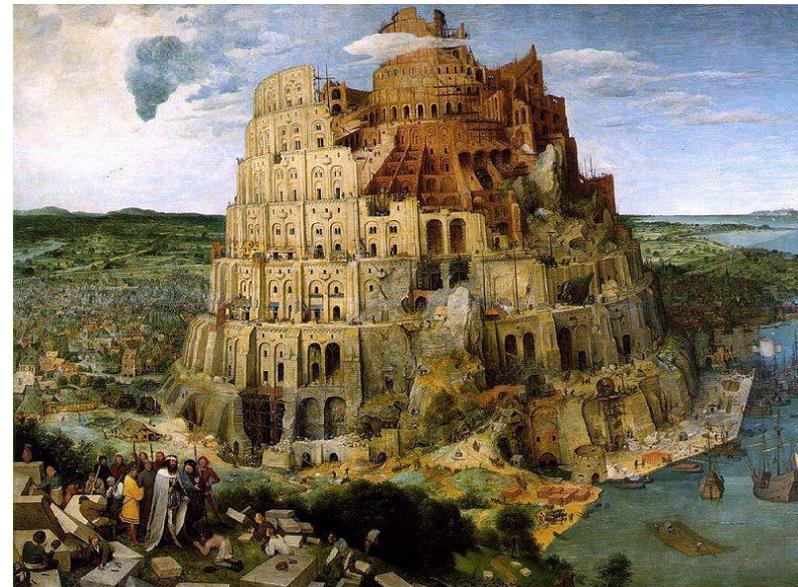
- ▶ In 2005-2007 the infrastructure for supporting standards development was modernized with funding from the Gordon and Betty Moore Foundation.
- ▶ In 2006 the scope was changed into developing **standards for sharing** biodiversity data
- ▶ In 2017, nearing completion of the process of becoming an official non-profit organization in the US and Europe.

# TDWG development phases

- ▶ Phase 0 (1985)
  - *Seemed like a good idea at the time*
- ▶ Phase 1 (first decade)
  - *Data dictionaries, data models*
- ▶ Phase 2 (second decade)
  - *ER models, DIGIR, DwC, XML, etc.*
- ▶ Phase 3 (third decade)
  - *Schemas, ontologies, RDF*
- ▶ Phase 4 (Now)
  - Re-evaluating prior standards and developing infrastructure to support documentation and vocabulary maintenance of existing standards
  - TDWG is partnering with **Global and regional Data Infrastructures** (GBIF, iDigBio, ALA and others) which can aid the standards development providing leadership, expertise and tools

# What drives the TDWG community?

- ▶ Untangle the 'Biodiversity Babel'
- ▶ Develop common communication
- ▶ Harness efficiency of collaboration
- ▶ Economic pressures to reduce duplication



# Governance

## Executive Committee

### Core Officers

- Chair
- Deputy Chair
- Secretary
- Treasurer

### Co-opted Members (Non-Voting)

- Conference Program Committee Co-Chair
- TDWG Coordinator

### Subcommittee Chairs

- Technical Architecture Group
- Fundraising and Partnerships
- Infrastructure
- Outreach and Communications
- Time and Place

### Regional Representatives

- Africa
- Asia
- Europe
- Latin America
- North America
- Oceania

# Governance

## Executive Committee

- Core Officers
- Co-opted Members
- Subcommittee Chairs
- Regional Representatives

## Functional Subcommittees

- Fundraising and Partnerships
  - Infrastructure Subcommittee
  - Outreach and Communications
  - Time and Place Subcommittee
- Technical Architecture Group (TAG)

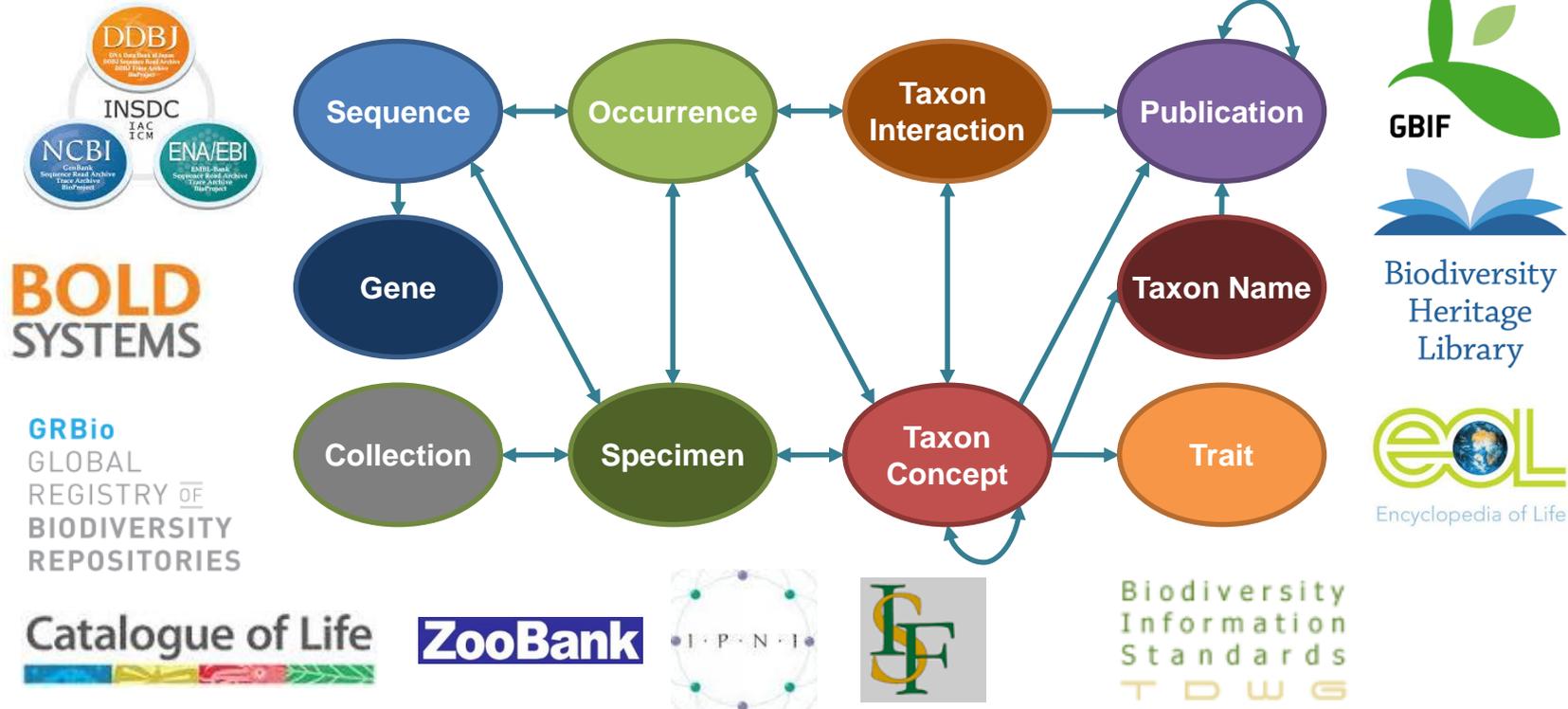
## Interest Groups (long term)

- Biodiversity Data Quality (BDQ)
- Species Information
- Biodiversity Informatics Curriculum (BDIC)
- ...

## Task Groups (temporary)

- BDQ Framework
  - BDQ Tools, Services & Workflows
  - BDQ Use Case Library
- Plinian Core
  - Invasive Species

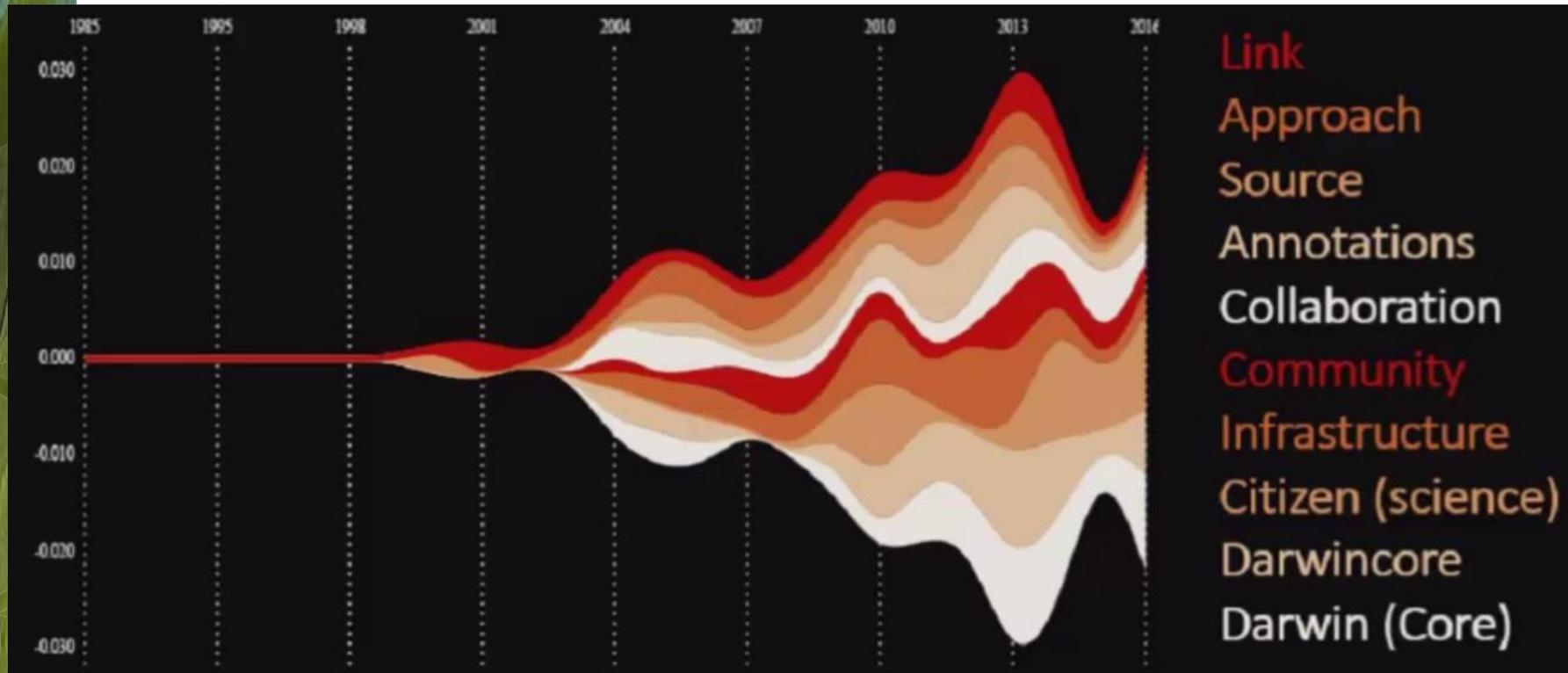
# Data Infrastructures need TDWG standards...



- Maintain catalogues of class instances
- Identify and unite related/identical instances
- Communicate to resolve/harden links between instances
- Serve as hubs for curation and annotation

... To deliver linked global biodiversity knowledgebase

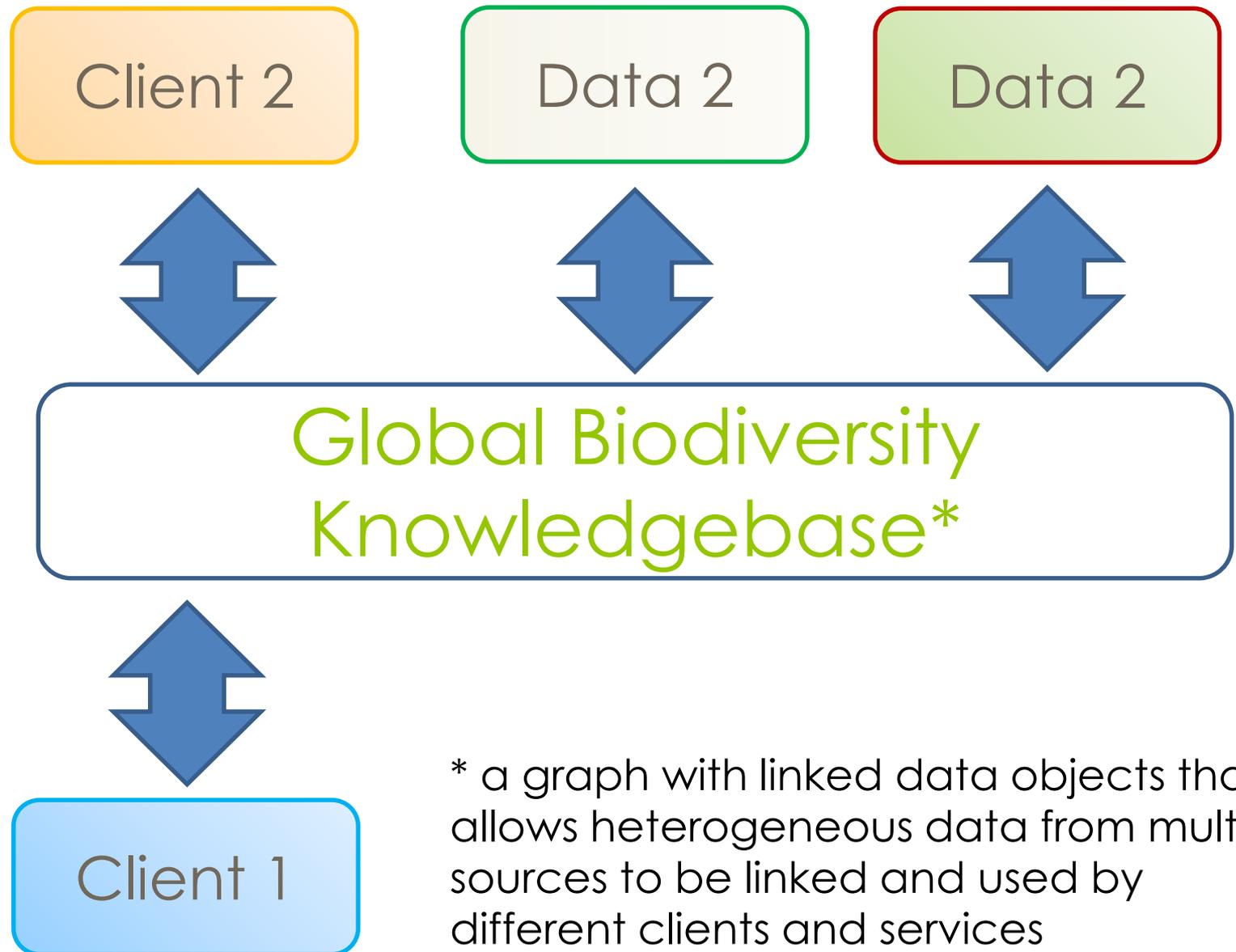
# A rich, loose environment of shifting topics



A.H. Ariño et al: TDWG Now and Then, TDWG , Costa Rica, 7-XII-2016

## Emerging Terms

TDWG Standards help in the development of the  
Global Biodiversity Knowledgebase

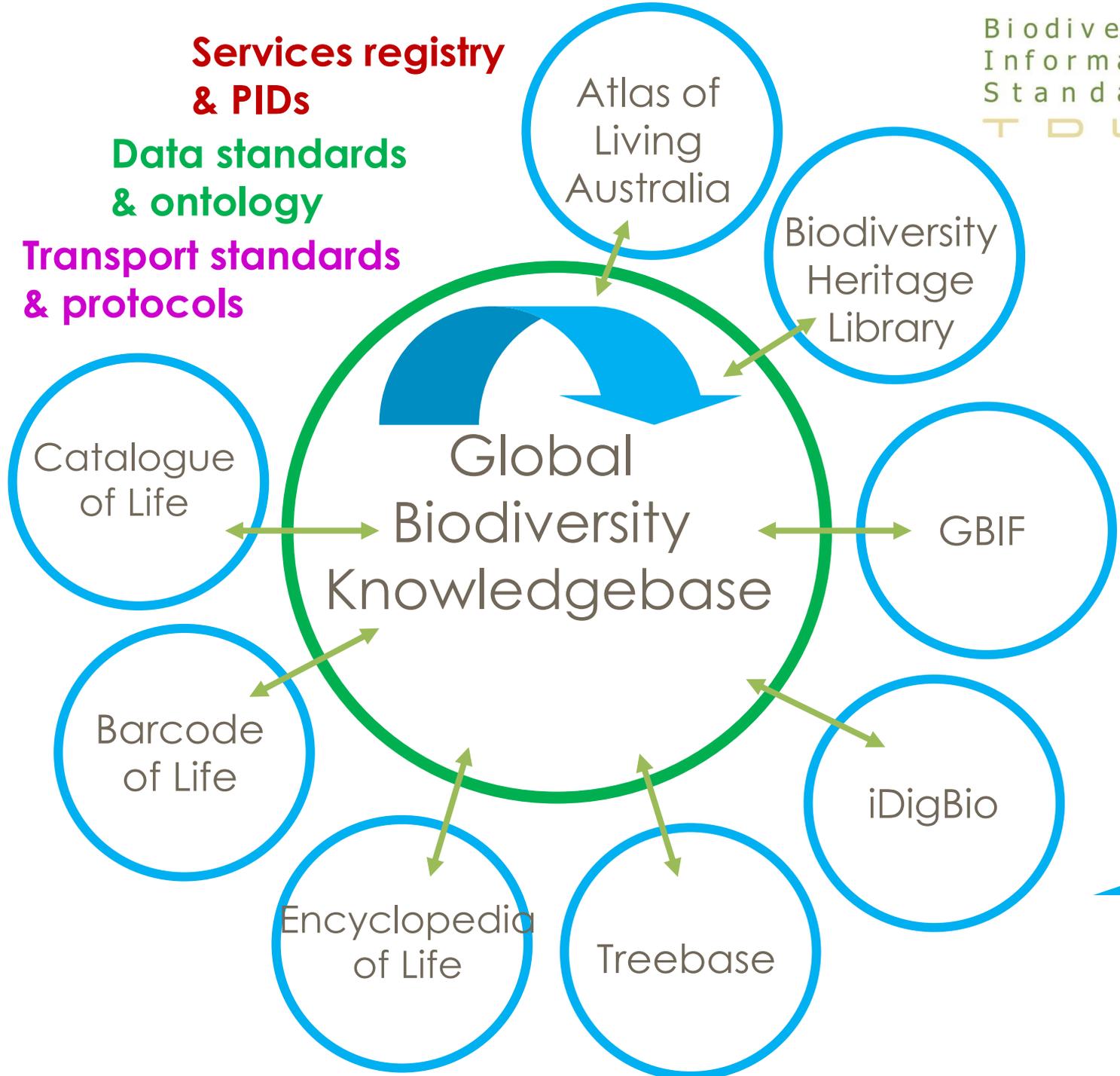


\* a graph with linked data objects that allows heterogeneous data from multiple sources to be linked and used by different clients and services

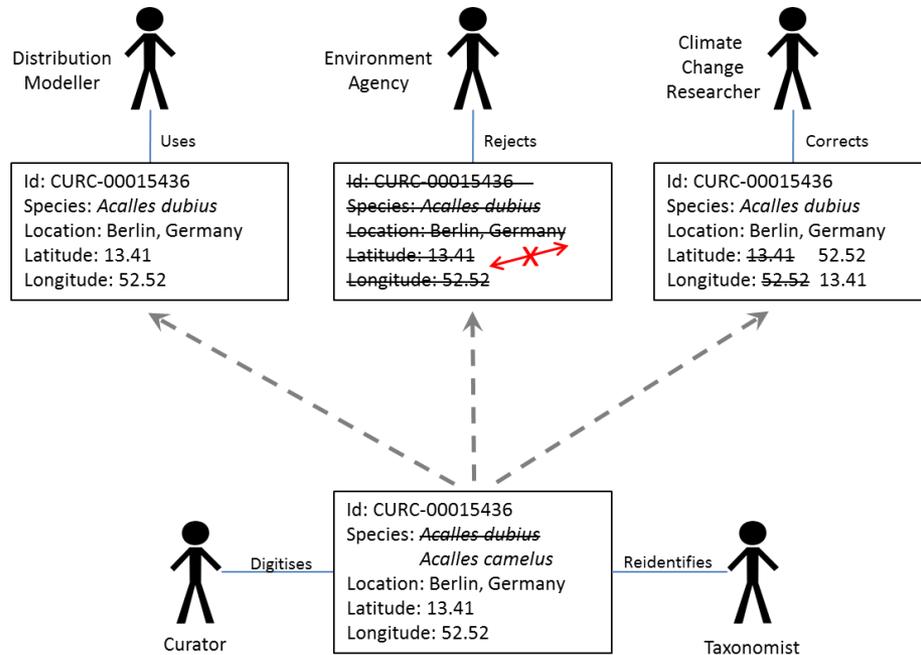
**Services registry  
& PIDs**

**Data standards  
& ontology**

**Transport standards  
& protocols**

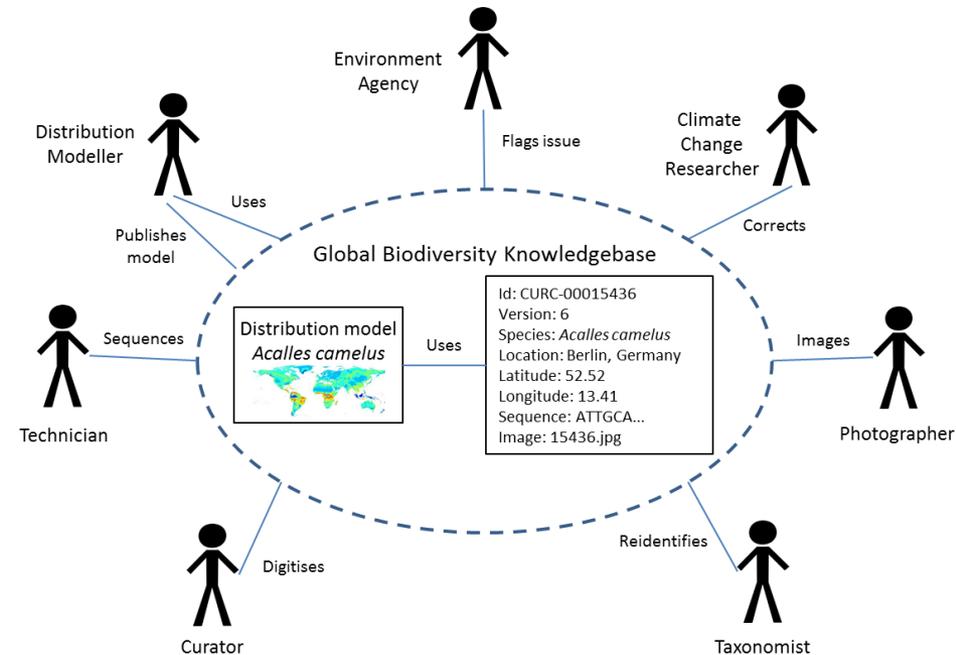


# Towards a Global Biodiversity Knowledgebase



## Today

- Disconnected publishing and use
- Limited feedback
- Poor ability to reference data



## Future?

- Stable, connected data from all sources
- Collaborative management
- Continuous step-wise improvement



# Annual Conferences

ONCE UPON A TIME,

1985

When Courier the Old Ruled, and Most Spake xBase,  
A Dozen Birds of a Feather Flocked Together...

D. Clayton  
D. Brummit  
D.L. Hawksworth  
M.N. Dadd  
H. Synge  
D.C. Mackinder  
E. Forero



F. Bisby  
G.T. Bocket  
G.T. Prance  
J.F.M. Cannon  
J.M. Mascherpa  
M. Crosby  
V.H. Heywood

# What are TDWG conferences like now ...

Typically a mixture of forums:

- **Symposia:** Sessions with talks dedicated to a topic
- **Workshops:** Discussion-based, typically related to standards development
- **Computer demonstrations:** Implementation of standards in software
- **Interest Group meetings:** Members discuss tasks that are underway

# Conference Evolution

1985: 9 attendees talking about harmonizing taxonomical databases



2017: large conference about anything around biodiversity data

2017



# Conferences Overview

Prof. Heywood about the first TDWG meeting held in 1985 in Geneva: *"It provided a stimulating discussion amongst a dozen or so participants, punctuated by the elegant hospitality of Prof. Bocquet and colleagues"*

1985	Geneva	Switzerland	2002	Campinas	Brazil
1986	Pittsburgh	USA	2003	Lisbon	Portugal
1987	Edinburgh	Scotland	2004	Christchurch	New Zealand
1988	St Louis	USA	2005	St Petersburg	Russia
1989	Las Palmas	Gran Canaria	2006	St Louis	USA
1990	Delphi	Greece	2007	Bratislava	Slovakia
1991	Canberra	Australia	2008	Fremantle	Australia
1992	Xalapa	Mexico	2009	Montpellier	France
1993	Washington	USA	2010	Woods Hole	USA
1994	Paris	France	2011	New Orleans	USA
1995	Madrid	Spain	2012	Beijing	China
1996	Toronto	Canada	2013	Firenze	Italy
1997	Taipei	Taiwan	2014	Jönköping	Sweden
1998	Reading	England	2015	Nairobi	Kenya
1999	Harvard	USA	2016	Santa Clara de San Carlos	Costa Rica
2000	Frankfurt	Germany	2017	Ottawa	Canada

# TDWG 2018

## Dunedin, New Zealand

Biodiversity  
Information  
Standards  
TDWG



**Society for the Preservation of Natural History Collections &  
Taxonomic Database Working Group Joint Conference 2018**  
[www.spnhc-tdwg2018.nz](http://www.spnhc-tdwg2018.nz)



# TDWG

# Peer-Reviewed Journal

Biodiversity  
Information  
Standards  
TDWG

**BISS** Biodiversity  
Information  
Science and  
Standards



Full Text Author Title

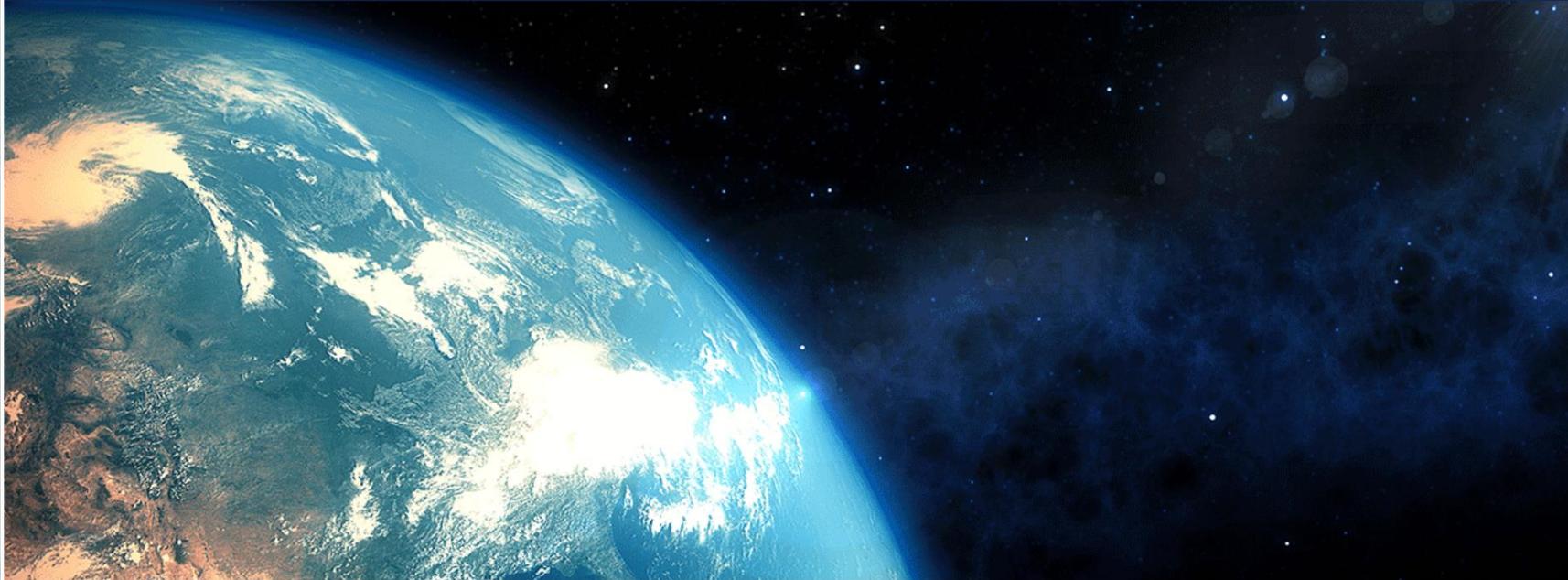
Start a manuscript

Articles

About

Contacts

Editorial team



► BISS launched in October 2017...



# TDWG

## Peer-Reviewed Journal

# BISS

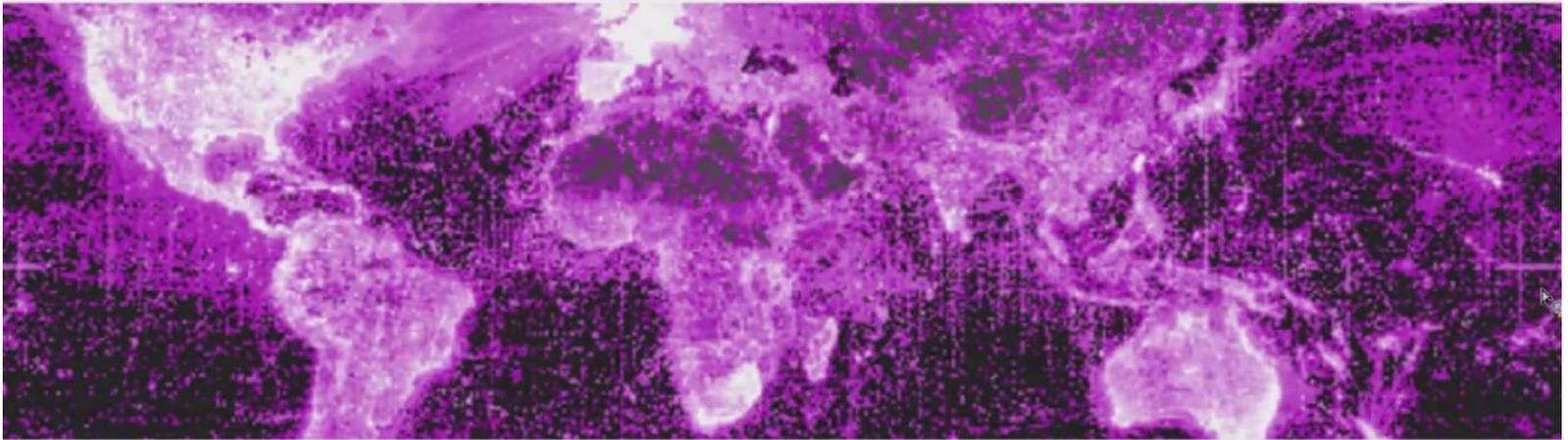
Biodiversity  
Information  
Science and  
Standards

Biodiversity  
Information  
Standards  
T D W G

Biodiversity Information Science and Standards (BISS) is an innovative open access journal publishing abstracts related to biodiversity standards, methods, guidelines, models and applications in biodiversity informatics submitted to Biodiversity Information Standards (TDWG) for presentation at annual meetings.

The journal also publishes conventional research articles. They will only be considered if they illustrate the development or application of new methods and approaches in biodiversity informatics in the form of case studies.

# TDWG Achievements



- ▶ **>1.000.000.000 occurrence records** exchanged with TDWG standards
- ▶ Used by over **1000 data publishers** worldwide
- ▶ TDWG standards **enabled the exchange of natural history and biodiversity data** and are the basis of the current global (GBIF) and regional data infrastructures
- ▶ **25 standards** produced by our volunteers

# TDWG Achievements

- ▶ TDWG standards are **globally adopted** in the natural sciences community and also used to exchange data from natural sciences with other domains, like cultural heritage (Europeana)
- ▶ **Recommended and supported by influential global and regional organisations** and initiatives in the natural sciences and biodiversity domain, e.g. GBIF, iDigBio, ALA, BioCAsE.
- ▶ More than **30 global annual conferences** have been organised, attended by thousands of participants from taxonomists to data scientists.
- ▶ A true **community effort** in which hundreds of volunteers worked together on TDWG publications that have been **cited by thousands** of authors

# Priorities for the future

- ▶ **High-level meta-model**
  - *Essential data classes for biodiversity informatics*
  - *Key relationships between class instances*
- ▶ **Mandatory and recommended properties**
  - *Revision of existing TDWG vocabularies*
  - *Some generic terms independent of data class*
  - *Others tied to a specific data class*
- ▶ **Serialisation models**
  - *Guidance on sharing data in different formats*
  - *Best practice for denormalised data exchange*
- ▶ **GUID schemes and HTTP URIs**
  - *Best practice to support and manage linking data*

# Our challenges

- ▶ **Developing effective standards is not easy**
  - ▶ Involves people with different interests and perspectives – also need to deal with conservatism, rivalries, egos...
  - ▶ Involves complex areas of knowledge – trivial issues are not trivial...
  - ▶ Involves multiple disciplines – experts from different areas need to understand each other
  - ▶ Information Technology evolves very quickly – standards can soon become obsolete!
- ▶ **TDWG needs more resources to facilitate rapid development of effective standards**

# How to Participate

- ▶ Take a look at the TDWG website (tdwg.org) ... see anything interesting? Dig in deeper – registering on the site is free



The screenshot shows the TDWG website homepage. The header features the TDWG logo and a navigation bar with 'Login' and 'SEARCH' buttons. A left sidebar contains a menu with items like 'Home', 'About TDWG', 'Getting Started', 'Standards', 'Activities', 'Membership', 'Past Meetings', 'Acknowledgements', '2017 Conference', 'Biodiversity Information Projects of the World', 'Biodiversity Information Networks Database', 'Proceedings', 'Mailing Lists', and 'Glossary'. The main content area is divided into three columns: 'Basic Standards Recommendations', 'TDWG 2017 Annual Conference: Oct 1-6 in Ottawa, Canada', and 'Latest News'. The 'Latest News' section includes a 'Call for Tender: Website Development' announcement from June 2017 and a 'TDWG Community Support Fund - Call for Proposals 2017' announcement from February 2017.

Biodiversity Information Standards TDWG

Home  
About TDWG  
Getting Started  
Standards  
Activities  
Membership  
Past Meetings  
Acknowledgements

2017 Conference

Biodiversity Information Projects of the World  
Biodiversity Information Networks Database

Proceedings  
Mailing Lists  
Glossary

Basic Standards Recommendations

The TDWG community's priority is the development of standards for the exchange of biological/biodiversity data.

The most widely deployed formats for biodiversity occurrence data are [Darwin Core](#) and [ABCD](#).

The TDWG [GUID Applicability Statement](#) provides guidance on the assignment of Globally Unique Identifiers to biodiversity information resources.

TDWG's activities within the biodiversity informatics domain can be found in the [Activities](#) section of this website.

TDWG 2017 Annual Conference: Oct 1-6 in Ottawa, Canada

TDWG's 2017 Annual Conference will hosted by Agriculture and Agri-Food Canada and the Canadian Museum of Nature in Ottawa, Canada. The conference theme is:

"Data Integration in a Big Data Universe: Associating Occurrences with Genes, Phenotypes, and Environments."

Latest News

Call for Tender: Website Development

04-Jun-2017  
TDWG seek a contract with a motivated and experienced web developer to revise the TDWG.org website. TDWG needs a website to communicate and support its outputs and activities. One that 1) provides a clear overview of what TDWG...[\[more\]](#)

[\[more\]](#)

TDWG Community Support Fund - Call for Proposals 2017

06-Feb-2017  
TDWG Interest (IG) and Task (TG) Groups undertake the core business of our organization.

# How to Participate

## ► Individual Membership: US\$75 / Year

Individual membership is open to anyone. TDWG develops international standards that enable the open sharing of the world's biodiversity data. These standards are therefore crucial to the management and preservation of the life of the planet. TDWG needs your expertise and experience to improve the sharing of these data.

- Participate in the primary international forum for solving problems in biodiversity information management.
- Develop international standards and protocols that enable individuals and organizations to exchange and integrate biodiversity information.
- Have full access to the TDWG Online Environment, the annual meetings, and a dynamic learning environment.
- Receive information that will help you track the rapidly developing world of biodiversity informatics.
- Be part of a knowledgeable and welcoming community of enthusiasts dedicated to expanding and sharing biodiversity information.
- Learn about new developments that will make your institution more effective and efficient in accessing, managing, and publishing biodiversity data.
- Comment on new standards, protocols, procedures and vote on constitutional matters.
- Receive discounts on registration fees for the annual meetings.

# How to Participate

## ► Institutional Membership: US\$500 / Year

Institutional membership of TDWG is open to any organization, company or government agency. Institutional members are agencies that use or support TDWG standards.

- The annual meeting draws an extremely broad cross-section of the community together. Bioinformatics specialists, biologists, information scientists, botanists, curators, librarians and a host of others present their work at this meeting. Meeting people from similar institutions to your own can be an enlightening experience.
- Thousands of agencies receiving, managing and distributing biodiversity data use TDWG's standards and therefore should contribute to the development, testing and deployment of those standards.
- Membership provides an opportunity for staff members to maintain an awareness of advances in bioinformatics and bring these advances into their institution.
- Institutions can save hundreds of thousands of dollars by adopting advances from other TDWG institutional members. Don't reinvent the wheel. Read how the Chicago Botanic Garden achieved a better outcome and saved thousands of dollars on the Testimonials page.
- You would be supporting a non-profit, volunteer organization. An organization that is producing international standards that provide the infrastructure that helps to manage the planet's biodiversity.

# Thank You

The logo for the Coalition for Networked Information (cni) features the lowercase letters 'cni' in a bold, blue, sans-serif font. A small orange circle is positioned above the letter 'i'. The logo is enclosed within a thin orange border that forms a partial frame.

Coalition for Networked Information

## Biodiversity Information Standards T D W G

Martin R. Kalfatovic  
TDWG Fundraising and Partnerships  
Subcommittee, Chair

Contributions from:

- Wouter Addink  
TDWG Europe Regional  
Representative
- Dimitris Koureas  
TDWG Chair
- James A Macklin  
TDWG Deputy Chair

12 December 2017  
Washington, DC

---