Priorities for the development of a European taxonomic information system

Cornelia Löhne & Walter Berendsohn

Botanischer Garten und Botanisches Museum Berlin-Dahlem

Freie Universität Berlin

EPBRS Meeting, Pruhonice 19-22 May 2009

1   Taxonomic information: strategy & methods

1a Inventory and identification

2   Taxonomy as a basis for ecological research and sustainable management of biodiversity

3   Taxonomy, potential users & capacity building

3a Open access to information
(„How to find out about it?")

Could research improve the way in which taxonomic Information is managed, and how it is delivered to its users – especially non-taxonomic users?

Taxonomy: What is it and what do we need it for?

Taxonomy = the science which identifies, describes, classifies and names living beings.

Taxonomy is the most fundamental of life sciences and
is becoming crucial to biodiversity management, public health,
agriculture, and many other aspects of life and society.

[*from EDIT homepage*]

Division: Angiospermae
Class: Dicotyledonae
Subclass: Asteridae Takht.
Order: Gentianales Lindley
Family: Gentianaceae Juss.
Genus: *Gentianella* Moench
Species: ***Gentianella bohemica*** Skalicky
          [reference: Preslia, 41: 144 (1969)]

[Photo: C. Angerer, www.floraweb.de]

Taxonomy: What is it and what do we need it for?

Taxonomy = language for communication.

Imagine if people and things didn't have a name,
or **if many names were wrong**…

[*Johan Liljablad at EPBRS meeting, 20 May 2009*]

Division: Angiospermae
Class: Dicotyledonae
Subclass: Asteridae Takht.
Order: Gentianales Lindley
Family: Gentianaceae Juss.
Genus: *Gentianella* Moench
Species: ***Gentianella bohemica*** Skalicky
          [reference: Preslia, 41: 144 (1969)]

          [Photo: C. Angerer, www.floraweb.de]

Taxonomy: What is it and what do we need it for?

Taxonomy = language for communication.

Imagine if people and things didn't have a name,
or **if many names were wrong**…

[*Johan Liljablad at EPBRS meeting, 20 May 2009*]

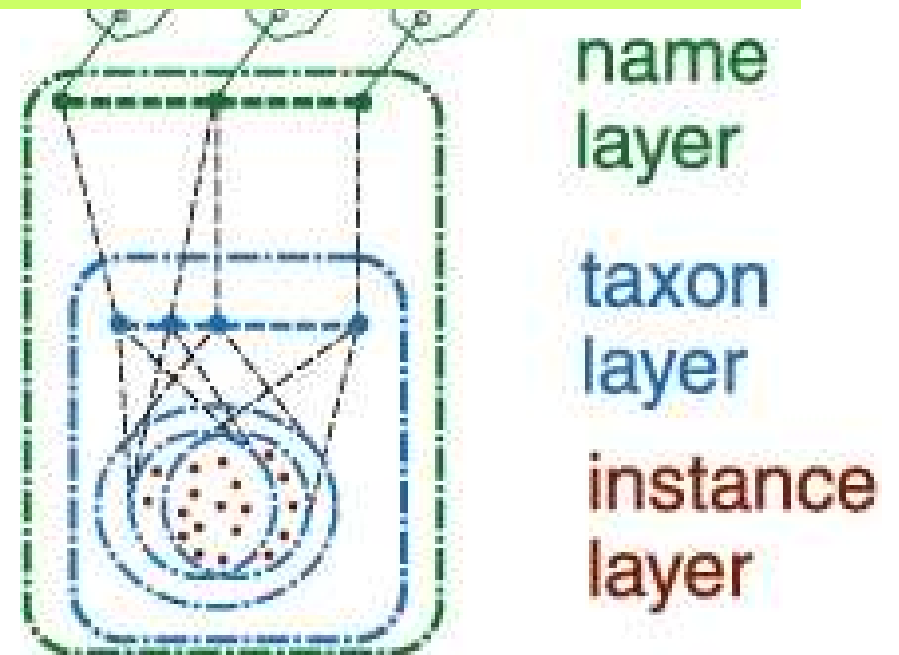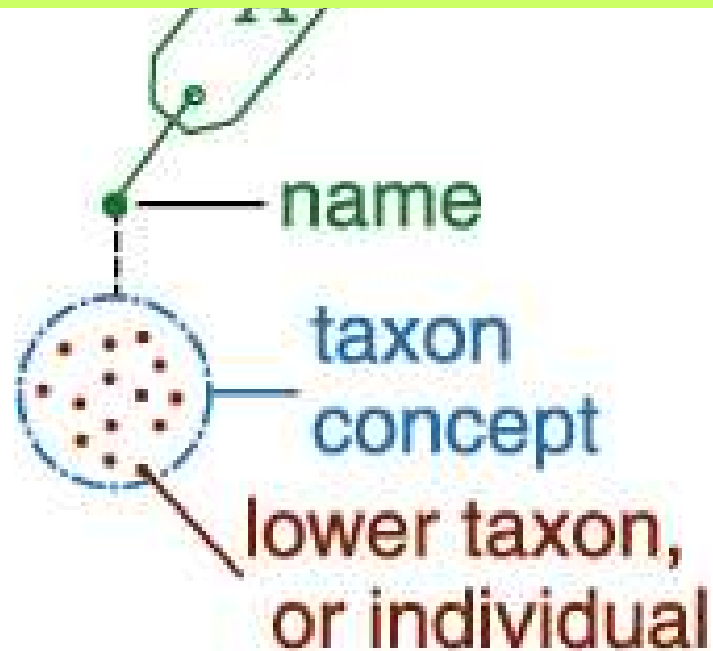## Are names reliable keys for information?

[Photo: C. Angerer, www.floraweb.de]

Taxonomy: Names vs. Concepts

The simple case:
One concept, one name

The complex case (REALITY!):
Different concepts, different names

taxonomic concept = name + explicit usage (reference)

name

taxon concept

lower taxon, or individual

name layer

taxon layer

instance layer

[grafics: www.nomencurator.org]

Taxonomy: What is it and what do we need it for?

Taxonomy = language for communication.

Imagine if people and things didn't have a name,
or **if many names were wrong**...

[J. Liljablad at EPBRS meeting, 20 May 2009]

Are names reliable keys for information?
Not always !
Is this relevant?

[Photo: C. Angerer, www.floraweb.de]

Relevance of the Concept Problem: Example 1

- Wisskirchen, Haeupler & al. 1998: **German Standard List of Vascular Plants**

- 4709 accepted taxa are listed (3811 species), with indication of congruent taxon concepts in 6 (7) recent floristic works

- Fully databased (Floraweb.de)

- Funded by the German Federal Agency for Nature Conservation (BfN)

ANAGALLIS L. (Primulaceae) – Sp. Pl.: 148 (1.5.1753) – Typus: Anagallis arvensis L. – NCU – <u>Gauchheil</u>

  CENTUNCULUS L. – Sp. Pl.: 116 (1.5.1753) – Typus: Centunculus minimus L. – NCU

A **Anagallis arvensis L.** – Sp. Pl.: 148 (1753)* – Typus: Herb. Linn. No. 208.1, LINN (lecto- DYER in DYER et al. 1963: Fl. Southern Africa <u>26</u>: 14)   E F h M O R S
  <u>Acker-Gauchheil</u>

    Anagallis caerulea L. – Amoen. Acad. <u>4</u>: 479 (1759)
    Anagallis phoenicea SCOP. – Fl. Carniol., ed. 2, <u>1</u>: 139 (1771)*, nom. illeg. (nom. superfl.)
    Anagallis carnea SCHRANK – Baier. Fl. <u>1</u>: 461 (1789)
    Anagallis arvensis subsp. phoenicea VOLLM. – Ber. Bayer. Bot. Ges. <u>9</u>: 44 (1904), nom. inval.
    Anagallis arvensis fo. carnea (SCHRANK) LÜDI   O
    Anagallis arvensis var. phoenicea GOUAN
    Anagallis arvensis var. caerulea (L.) GOUAN – Fl. Monsp.: 30 (1765)
    Anagallis arvensis fo. azurea HYL. – Uppsala Univ. Årsskr. <u>7</u>: 256 (1945)   O

  Im Gebiet nur die subsp. *arvensis*

  Anagallis foemina MILL. – Gard. Dict., ed. 8: no. 2 (1768)   E F M O R S
    <u>Blauer Gauchheil</u>

    Anagallis caerulea SCHREB. – Spic. Fl. Lips.: 5 (1771) non L. 1759, nom. illeg.

[Wisskirchen & Haeupler 1998]

Relevance of the Concept Problem: Example 1

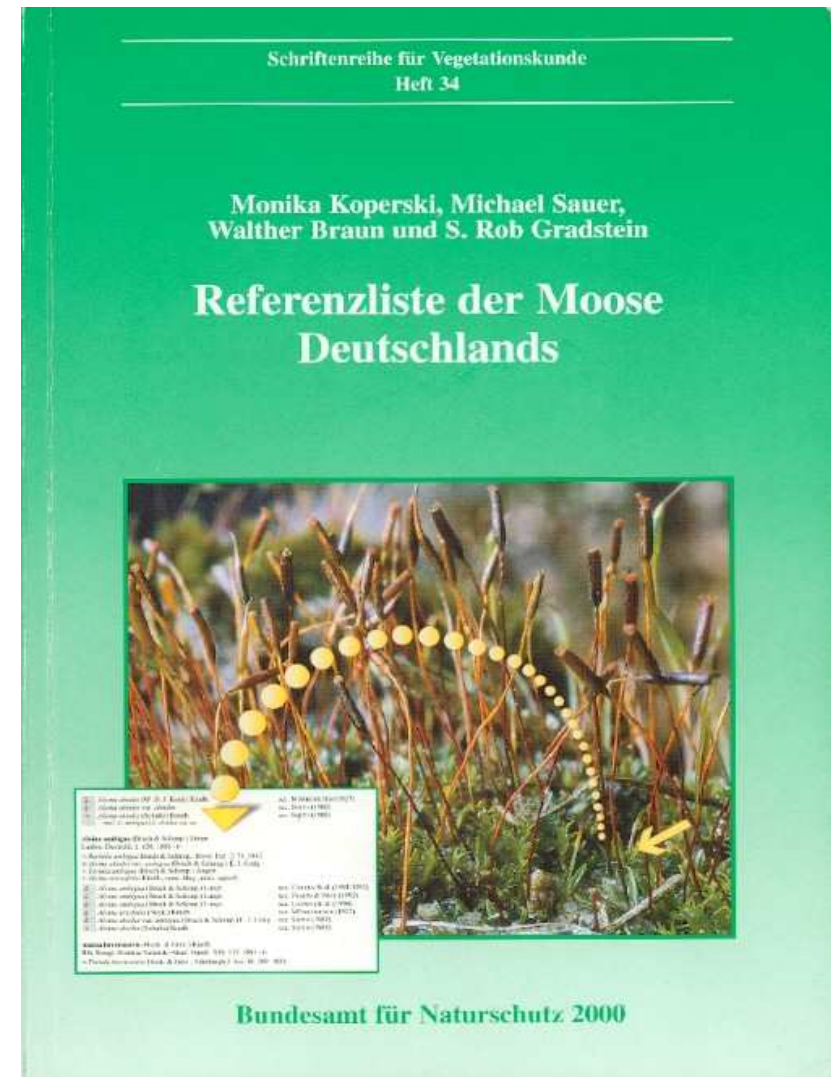- Wisskirchen, Haeupler & al. 1998: **German Standard List of Vascular Plants**

Result:

- **Roughly half of German vascular plant species are stable as to name and concept throughout works in current use**

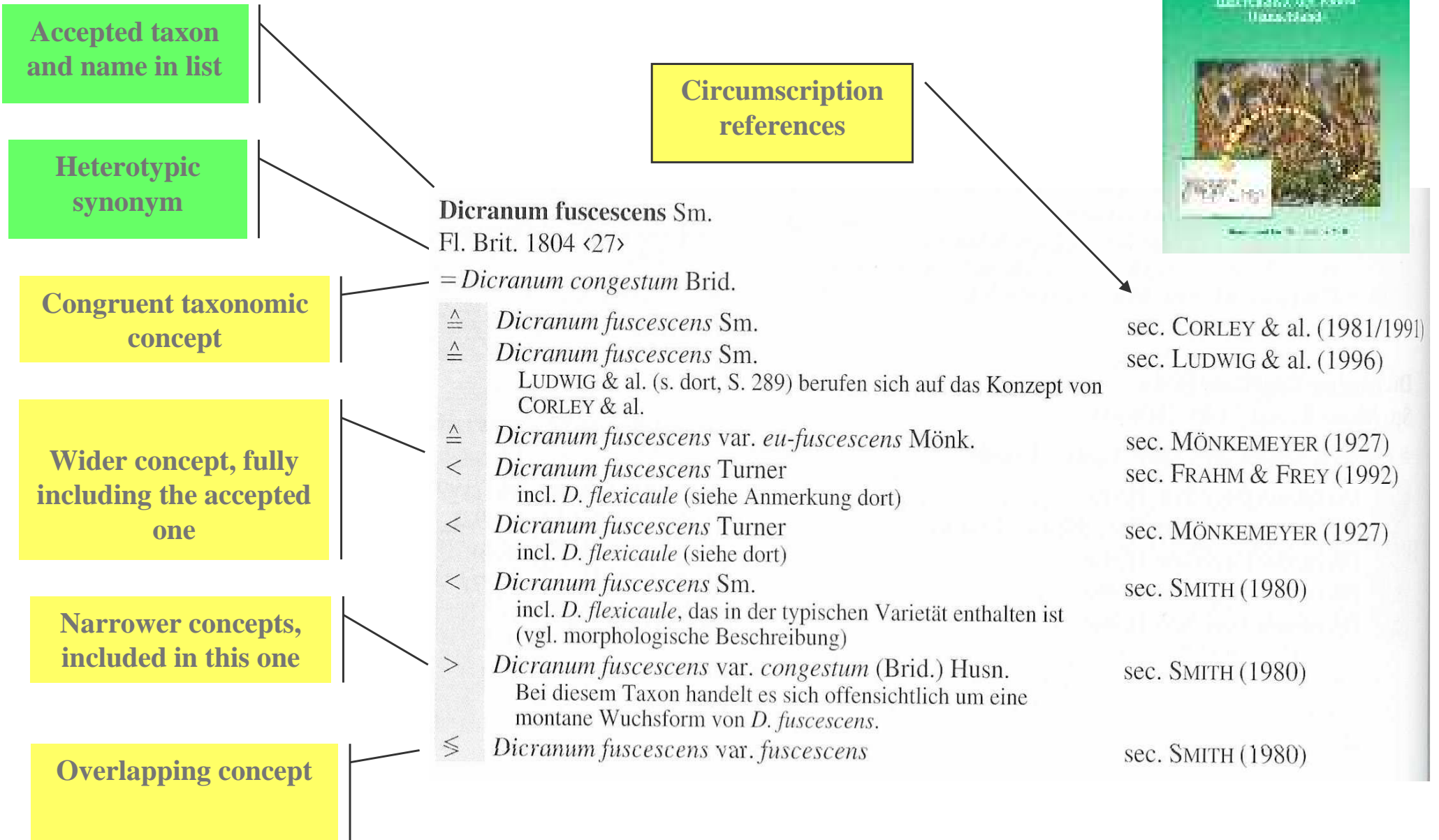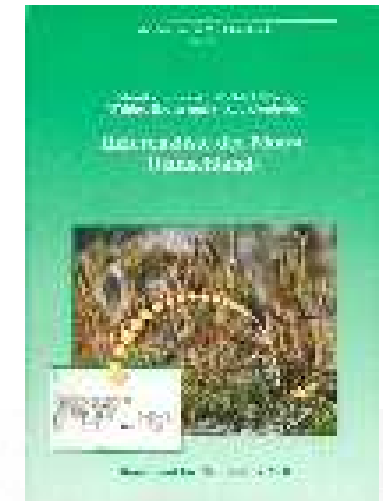- **→ the other half is not!**

Relevance of the Concept Problem: Example 2

- **Mosses of Germany**
  (Koperski, Sauer, Braun
  & Gradstein 2000)

- 1548 accepted taxa,
  analysis of 11 floristic or
  taxonomic treatments
  (mostly very recent, 1 from 1927)

- Funded by a user:
  German Federal Agency for
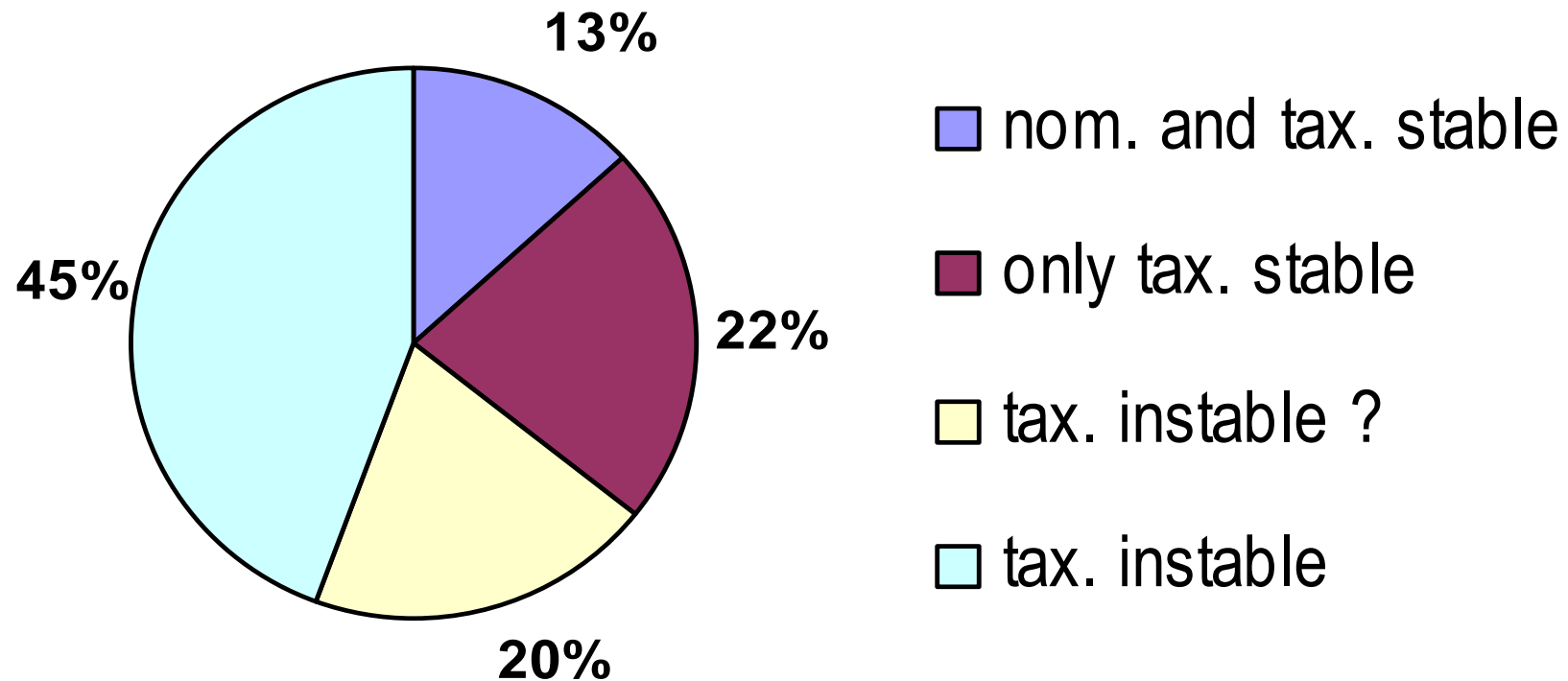 Nature Conservation (BfN)

Mosses of Germany
(Koperski, Sauer, Braun & Gradstein 2000)

Accepted taxon and name in list

Heterotypic synonym

Circumscription references

Congruent taxonomic concept

Wider concept, fully including the accepted one

Narrower concepts, included in this one

Overlapping concept

Dicranum fuscescens Sm.
Fl. Brit. 1804 ‹27›

= *Dicranum congestum* Brid.

≜ *Dicranum fuscescens* Sm. · sec. CORLEY & al. (1981/1991)
≜ *Dicranum fuscescens* Sm. · sec. LUDWIG & al. (1996)
　 LUDWIG & al. (s. dort, S. 289) berufen sich auf das Konzept von
　 CORLEY & al.

≜ *Dicranum fuscescens* var. *eu-fuscescens* Mönk. · sec. MÖNKEMEYER (1927)
< *Dicranum fuscescens* Turner · sec. FRAHM & FREY (1992)
　 incl. *D. flexicaule* (siehe Anmerkung dort)

< *Dicranum fuscescens* Turner · sec. MÖNKEMEYER (1927)
　 incl. *D. flexicaule* (siehe dort)

< *Dicranum fuscescens* Sm. · sec. SMITH (1980)
　 incl. *D. flexicaule*, das in der typischen Varietät enthalten ist
　 (vgl. morphologische Beschreibung)

> *Dicranum fuscescens* var. *congestum* (Brid.) Husn. · sec. SMITH (1980)
　 Bei diesem Taxon handelt es sich offensichtlich um eine
　 montane Wuchsform von *D. fuscescens*.

≶ *Dicranum fuscescens* var. *fuscescens* · sec. SMITH (1980)

Relevance of the Concept Problem: Example 2

Mosses of Germany (Koperski, Sauer, Braun & Gradstein 2000)

- 1548 accepted taxa, 11 floristic or taxonomic treatments (mostly very recent, one from 1927)

Pie chart legend:
- nom. and tax. stable — 13%
- only tax. stable — 22%
- tax. instable ? — 20%
- tax. instable — 45%

Names vs. Concepts: Relevance of the problem

Who are the USERS of taxonomic information?

- **Conservation**: Red Lists, priority setting in conservation, environmental monitoring, management of invasive species,…

- **Use of biodiversity**: Agriculture (e.g. crop wild relatives, pollinators), biological pest control, medicine & pharmaceutics, …

- **Research**: Ecology, evolution, physiology, biotechnology, genetics, pharmaceutical research, …

- e.g., human well-being, landscape planning …
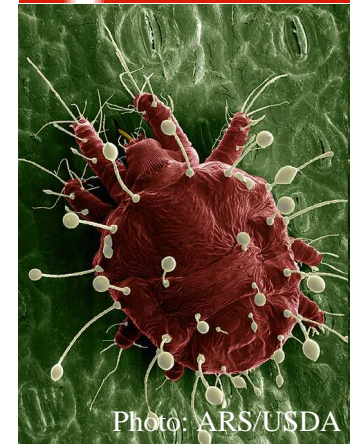
Photo: ARS/USDA

Relevance of the Concept Problem: more examples

BIODIVERSITY AND CONSERVATION

Taxonomic monographs in relation to global Red Lists

Jan Kirschner & Ždenek Kaplan

Department of Botany, Czech Academy of Sciences, CS-252 43 Pruhonice u Prahy, Czech Republic. kirsvhner@ibot.cas.cz (author for correspondence); kaplan@ibot.cas.cz.

Juncaceae & Potamogenotaceae:
50 – 75% of names in the IUCN red list either apply to widespread, non-threatened taxa or are doubtful

"…It may be concluded that the overall accuracy of the IUCN list is rather low."

Relevance of the Concept Problem: more examples

Examples mentioned by J. Liljeblad:

- *Scotinophara coarctata* (Malayan rice black bug)
- *Phenacoccus manihoti* (Cassava mealybug)
- Malaria vectors in Europe

Open Access to Information: Data bases

All historic and contemporary information about a species is tied to a scientific name […] This is especially problematic because **names of species are neither stable nor unique** *[…] These factors all significantly impact on the means to find, access, and effectively synthesize biodiversity data."*

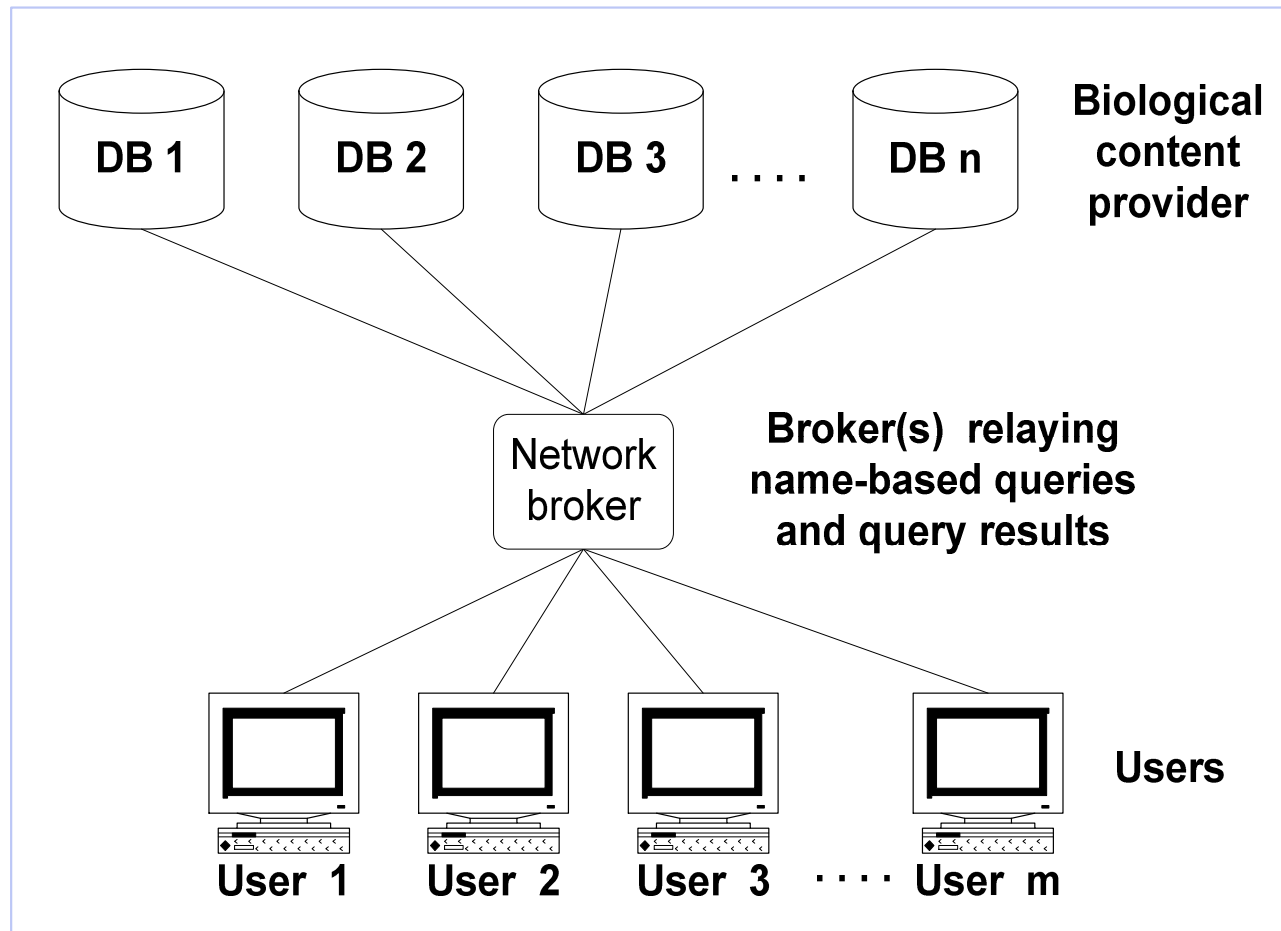D. Remsen, EPBRS e-conference, May 2009

Content linked to taxon names, e.g.:

- Uses (mostly human) and threats (to species itself, to hosts, to health, to environment, etc.)

- Ecology (pollination, symbiosis, parasitism, indicator value, edaphic and climatic requirements, etc.)

- Molecular data (natural substances, genes, sequences, physiology, etc.)

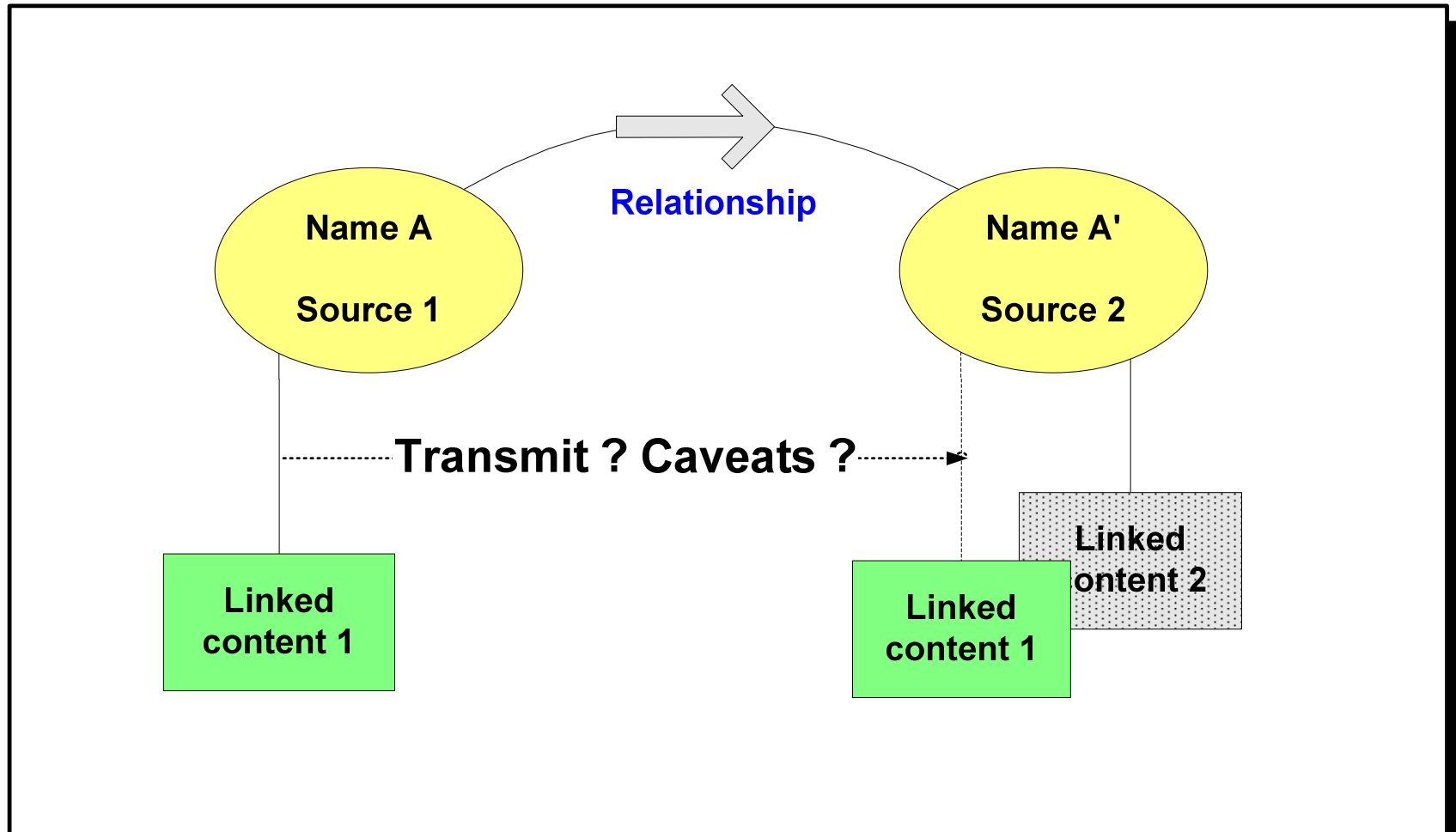- Geographical range or occurrence

- Descriptive data

Users and providers want to link all these data.

Users want a web-based "Unitary Taxonomy" to get reliable access to species information.

Open Access to Information: Data bases architecture (1)

- **Works only if all providers agree on a common standard taxonomy**

Diagram labels:
- DB 1, DB 2, DB 3 . . . . DB n
- Biological content provider
- Network broker
- Broker(s) relaying name-based queries and query results
- User 1, User 2, User 3 . . . . User m
- Users

Names vs. Concepts: Relevance of the Problem

Relationship

Name A

Source 1

Name A'

Source 2

Transmit ? Caveats ?

Linked content 1

Linked content 1

Linked content 2

Biological content provider

DB 1 DB 2 DB 3 DB n

Concept DB

Taxonomic Editor

Network broker

Broker(s) using stored concept relationships to handle name-based queries and query results

Users

User 1   User 2   User 3 · · · · User m

e.g., the "Berlin Model"

Common concept-oriented database core

How to find reliable keys to biodiversity information?

Various data models and exchange standards have been developed to deal with differing names and taxonomic concepts, e.g.:

- **ABCD – Access to Biological Collection Data** (http://www.bgbm.org/tdwg/codata/schema/),

- **Berlin Taxonomic Information Model** (http://www.bgbm.org/biodivinf/Docs/BGBM-Model/default.htm),

- **EDIT Common Data Model** (http://dev.e-taxonomy.eu/trac/wiki/CommonDataModel),

- **Global Names Architecture** (http://gnapartnership.org/gna/wiki)

- **Nomencurator**: a nomenclatural history model to handle multiple taxonomic views (www.nomencurator.org),

e.g., EDIT Common Data Model (CDM)
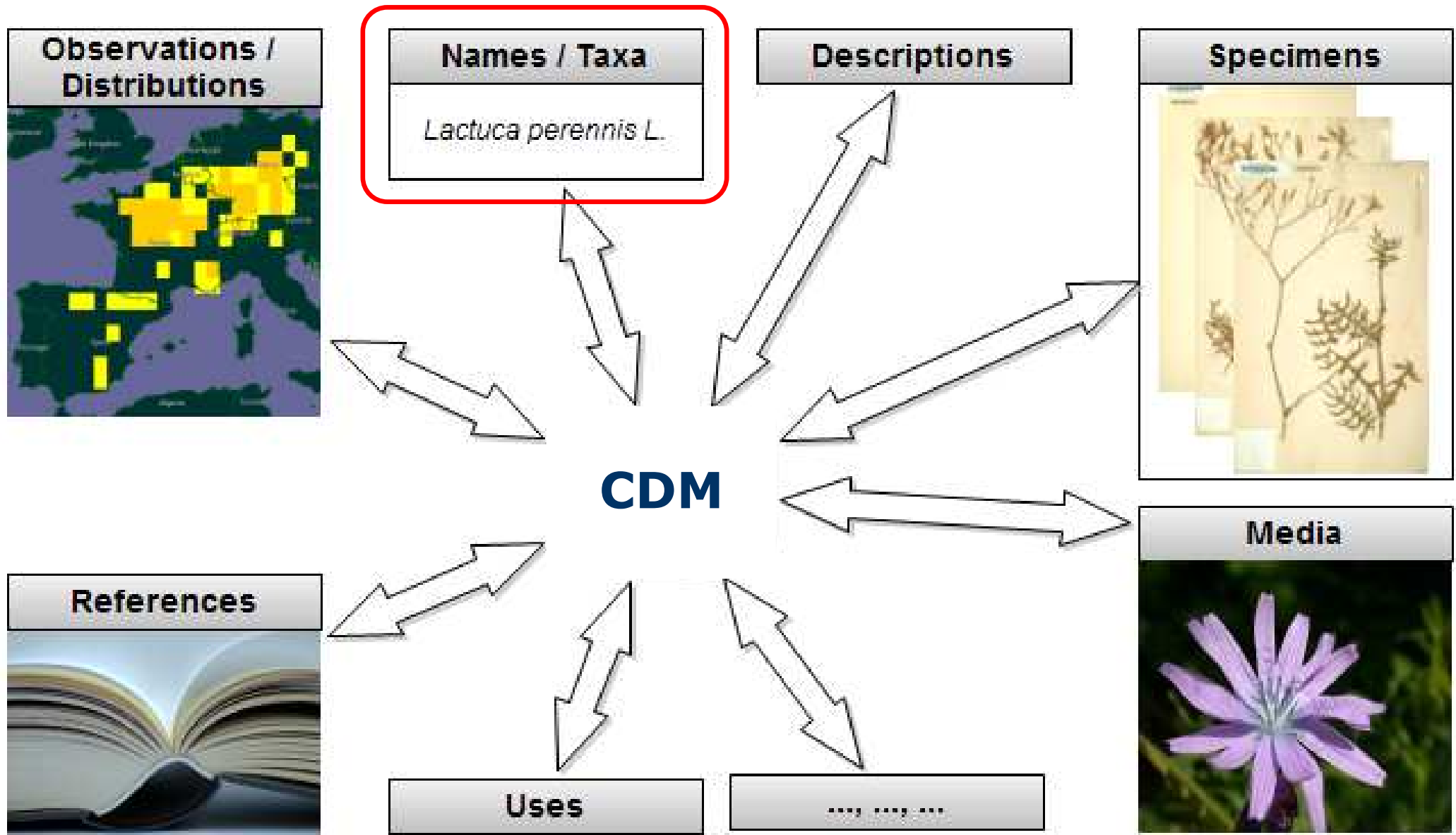
e.g., EDIT Common Data Model (CDM)

- Covering all relevant areas (names, agents, references, taxa, descriptions, locations, occurrences)

- Based on existing standards and models

- Exchange format

- Expandable
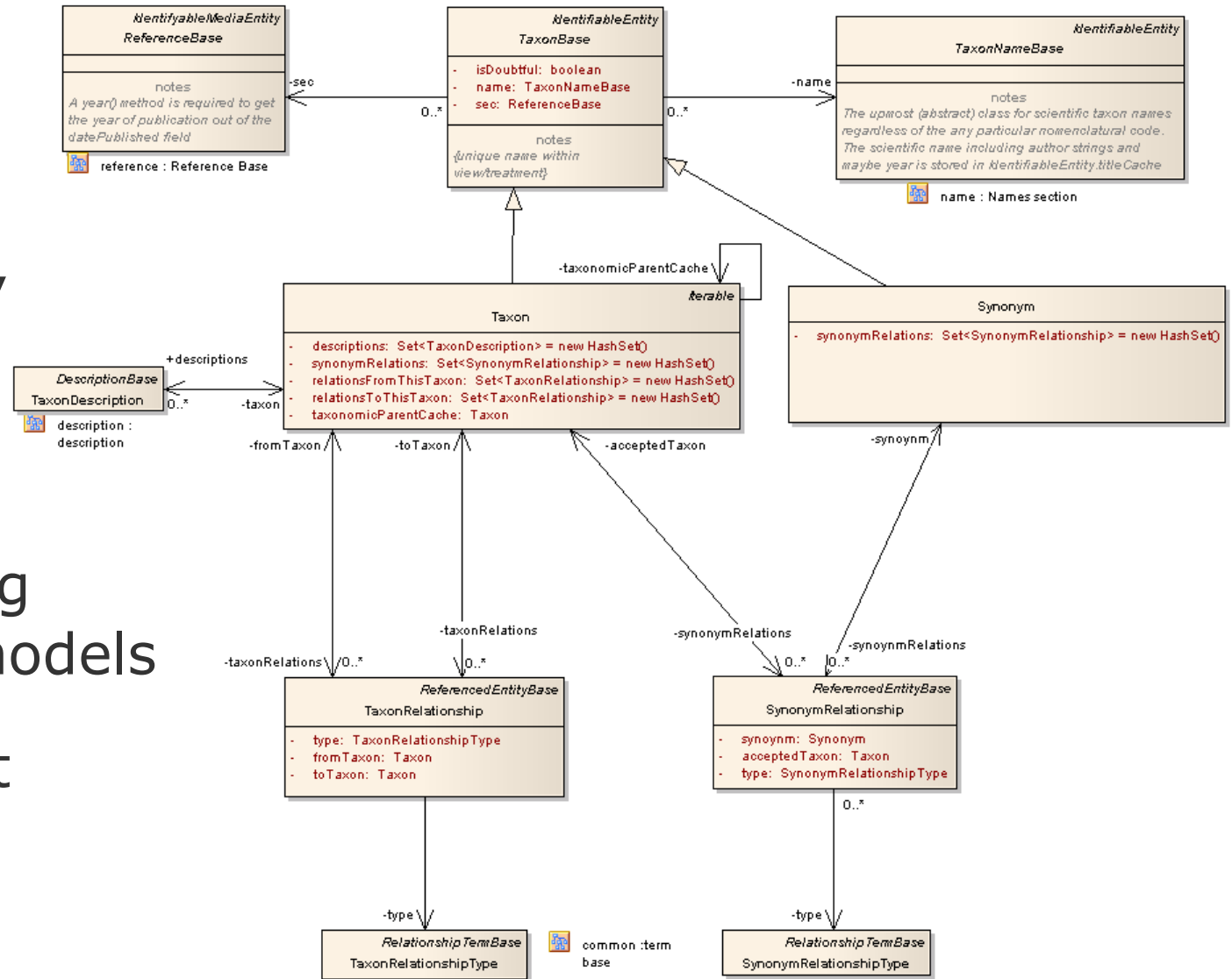
e.g., PESI – implementation at the European level

PESI: Pan-European Species directories Infrastructure
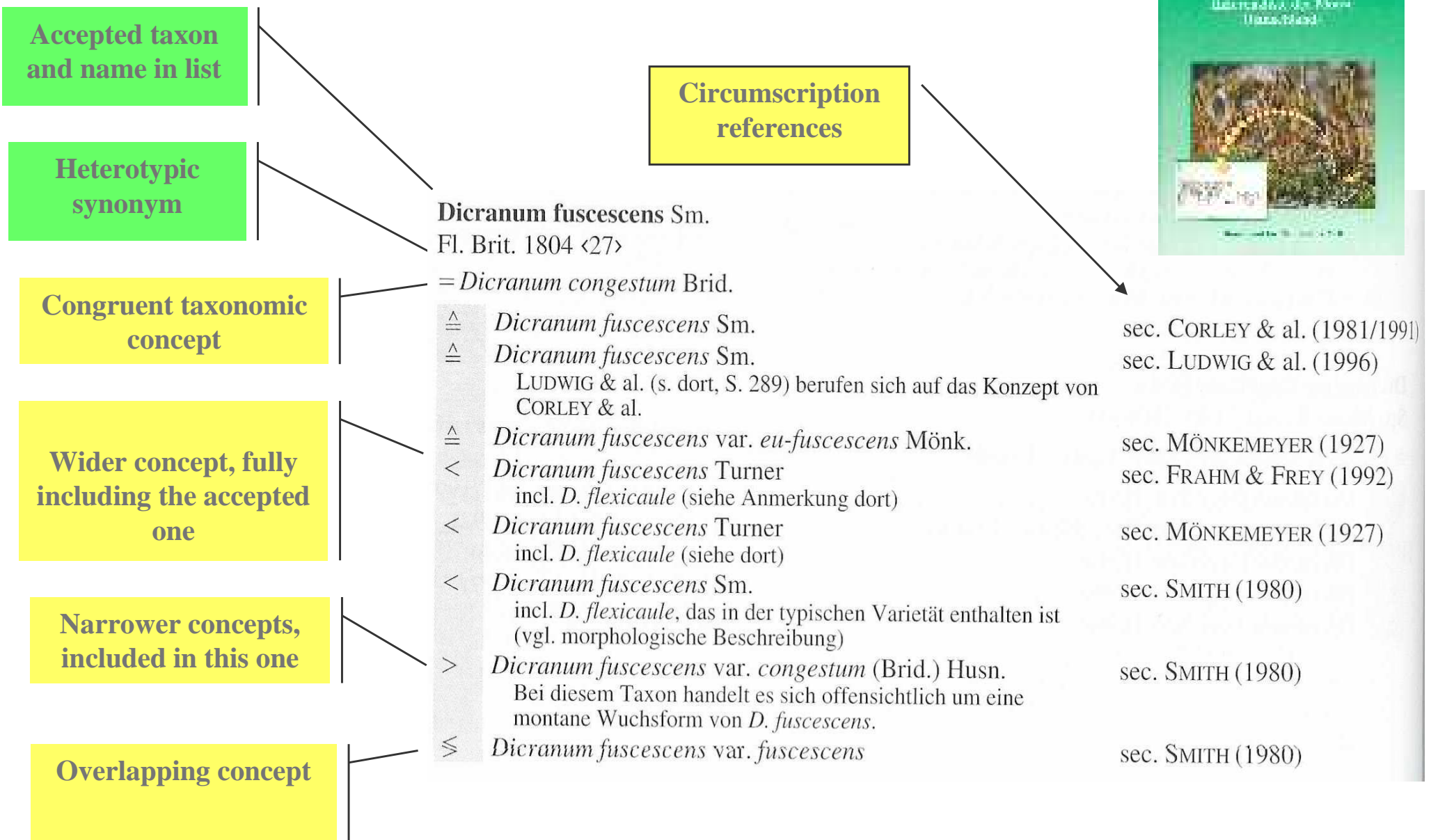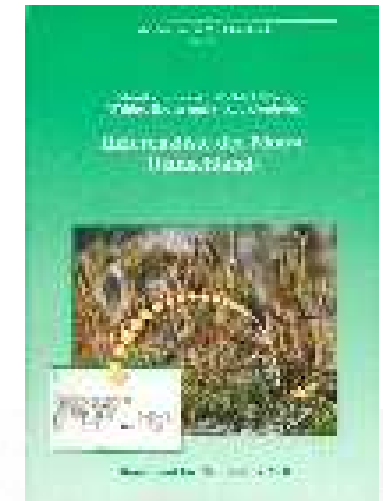
Networks of experts and focal points in order to build up taxonomic backbone

- Coordination of taxonomic metadata standard assessment

- Coordination and integration of information e-infrastructures (establish mirror sites for ERMS, Fauna Europaea, and Euro+Med Plant Base)

- Integrated e-Services for users and dissemination (i.e. an interactive, multilingual web portal)

Mosses of Germany
(Koperski, Sauer, Braun & Gradstein 2000)

Accepted taxon and name in list

Heterotypic synonym

Circumscription references

Congruent taxonomic concept

Wider concept, fully including the accepted one

Narrower concepts, included in this one

Overlapping concept

Dicranum fuscescens Sm.
Fl. Brit. 1804 ‹27›
= *Dicranum congestum* Brid.

≙ *Dicranum fuscescens* Sm. — sec. CORLEY & al. (1981/1991)

≙ *Dicranum fuscescens* Sm. — sec. LUDWIG & al. (1996)
   LUDWIG & al. (s. dort, S. 289) berufen sich auf das Konzept von CORLEY & al.

≙ *Dicranum fuscescens* var. *eu-fuscescens* Mönk. — sec. MÖNKEMEYER (1927)

< *Dicranum fuscescens* Turner — sec. FRAHM & FREY (1992)
   incl. *D. flexicaule* (siehe Anmerkung dort)

< *Dicranum fuscescens* Turner — sec. MÖNKEMEYER (1927)
   incl. *D. flexicaule* (siehe dort)

< *Dicranum fuscescens* Sm. — sec. SMITH (1980)
   incl. *D. flexicaule*, das in der typischen Varietät enthalten ist (vgl. morphologische Beschreibung)

> *Dicranum fuscescens* var. *congestum* (Brid.) Husn. — sec. SMITH (1980)
   Bei diesem Taxon handelt es sich offensichtlich um eine montane Wuchsform von *D. fuscescens*.

≶ *Dicranum fuscescens* var. *fuscescens* — sec. SMITH (1980)

Conclusions & Recommendations

The following approach is suggested:

1. Identify relevant groups of organisms, where a taxonomically sound backbone is needed for decision making and practical applications (*e.g. as part of a European monitoring scheme, e.g. pollinators, e.g. invasive species, e.g. crop wild relatives*)
→ **Communication between taxonomists and users of taxonomic information (user-driven research priorities)**

2. Clarify if there are problems in taxonomic concepts within the identified groups of organisms → **taxonomists**

3. Re-assess the respective groups (revision / standardization) and include information on concept relationships into open access data bases → **taxonomists**

Develop and apply standardized protocols for data collection in the field (e.g. provide reference for identification of the taxa!)