The Future of BioDiversity Informatics

What's Possible, What's Not

Robert J. Robbins
Fred Hutchinson Cancer Research Center
rrobbins@fhcrc.org

Abstract

Much biomedical research is carried out collaboratively across institutional boundaries. The computer infrastructure to support this research must often implement security in a manner that cannot (and should not) depend upon the enterprise security of any one institution.

In our experience, requirements analysis for a totally federated access control system shows the need for some interesting properties, such as the need to

(1) implement both groups (aggregations of human beings) and roles (aggregations of permitted actions),

(2) define formal authorization as “explicitly allowing members of group A to act in role B on resource C", where the individuals, groups, roles, and resources may all be located in different enterprises and be operated wholly independently,

(3) support the idea of formal deauthorization, defined as “explicitly prohibiting members of group A to act in role B on resource C", and

(4) support "clarity of roles", defined as ensuring that, at any one time, a user be permitted to act in one and only one role on a resource.

We believe that a properly implemented federated access-control system should, wherever possible, also provide capabilities to support federated usage auditing. We also argue that an enterprise-centric security model can be easily derived from a federation-centric model, while the converse is very difficult. Therefore, we suggest that developers of security systems would benefit from attending to the security needs of federations.

Abstract

Much biomedical research is carried out collaboratively across institutional boundaries. The computer infrastructure to support this research must often implement security in a manner that cannot (and should not) depend upon the enterprise security of any one institution.

In our [illegible] need for sor[illegible]

(1) [illegible]tted

(2) [illegible]ferent

(3) [illegible]roup

(4) [illegible] act in

We believe [illegible]possible, also provide capabilities to support federated usage auditing. We also argue that an enterprise-centric security model can be easily derived from a federation-centric model, while the converse is very difficult. Therefore, we suggest that developers of security systems would benefit from attending to the security needs of federations.

NOTE: This presentation will deal only with a LOGICAL specification of what services and functionality could be useful in meeting access-control needs in a totally federated environment. We will not consider any technical details of how such logical processes could be implemented or what it would take to ensure that such logical processes would operate in a genuinely secure manner.

Issues to be Covered

- Background
- Vision and Goal for the Future
- Challenges and Limits
  - As-yet Unsolved Technical Issues
  - Rapidly Changing Technology
  - The Dynamism of Science
- All Components, All the Time
- Making it work
  - Logical Simplicity
  - Social Scalability
- Conclusions

Background

Documenting the Problem
&
Articulating a Possible Solution

The challenge:
A biodiversity information infrastructure

•Species 2000 Workshop on the Catalogue of Life

•Reading, UK, 5 June 2001

• *by Hannu Saarenmaa* *hannu.saarenmaa@eea.eu.int*

•*European Environment Agency / EIONET / EC CHM*

•*http://www.eionet.eu.int/ and http://biodiversity-chm.eea.eu.int/*

The challenge: A biodiversity information infrastructure

Species 2000 Workshop on the Catalogue of Life
Reading, UK, 5 June 2001

by Hannu Saarenmaa *hannu.saarenmaa@eea.eu.int*
European Environment Agency / EIONET / EC CHM
http://www.eionet.eu.int/ and http://biodiversity-chm.eea.eu.int/

The challenge

Problems

- The Y3K problem...
  - Fast rate of extinction 10^4/y: Extrapolated to Y3K, there will not be a single species left
- Still only 2*10^6 species known
  - Slow and unchanged rate of discovery 10^4/y
  - Poor use of knowledge and difficulty in information access
  - Taxonomy often seen by others as endless sink of resources because results are not shared efficiently → No funding.
- Meanwhile, biodiversity information is mushrooming on Internet
  - Species homepages everywhere
  - ”Pinus sylvestris” returns 6251 hits on Altavista today; no coordination or standards whatsoever
  - Global Species Databases are not found by searches

Plenty of organisational response

- An incomplete inventory of biodiversity information networks and services: ABREN, All Species, BCIS, BIN21, BIODI, BIOSIS, CBD, CHM, CONABIO, DIVERSITAS, ECNC, ETC/NC, EWGRB, GBIF, IABIN, IBIN, ILDIS, INBio, IOPI, IPNI, ITIS, IUCN, MAB/BIS, NBII, Species 2000, TDWG, Tree of Life, WCMC.
- (Individual projects with a begin and an end have not been listed.)

Would the GenBank of biodiversity please stand up!

Plenty of organisational response

- An incomplete inventory of biodiversity information networks and services: ABREN, All Species, BCIS, BIN21, BIODI, BIOSIS, CBD, CHM, CONABIO, DIVERSITAS, ECNC, ETC/NC, EWGRB, GBIF, IABIN, IBIN, ILDIS, INBio, IOPI, IPNI, ITIS, IUCN, MAB/BIS, NBII, Species 2000, TDWG, Tree of Life, WCMC.
- (Individual projects with a begin and an end have not been listed.)

Would the GenBank of biodiversity please stand up!

We do not need more networks, but one information infrastructure

- Provides building blocks
- Defines standards for interchange
- Invites everyone to contribute to a common pool of information
- Enables sharing of data, information, and knowledge efficiently

What, exactly, is infrastructure?

- Building on each other's work using standardised interfaces.
- "[information] infrastructure is viewed as everything that supports the flow and processing of information." – whatis.com
- Building infrastructure does not necessarily require a mega-projects, and new organisations, although they are often viewed as such.
- Applications vs. infrastructure
  - If you want others to build on it -- then you are building infrastructure:
    - Characterised by usage of open standardized interfaces
    - Examples: IP, Z39.50, LDAP, IMAP, CORBA, Posix
  - If you just want to get one job done -- then you can build an application
    - Examples: Microsoft Windows, Microsoft Access, DCOM
  - Mature, successful, shareable applications often migrate to infrastructure

What, exactly, is infrastructure?

- Building on each other's work using standardised interfaces.
- "[information] infrastructure is viewed as everything that supports the flow and processing of information." – whatis.com
- Building infrastructure does not necessarily require a mega-projects, and new organisations, although they are often viewed as such.
- Applications vs. infrastructure
  - If you want others to build on it -- then you are building infrastructure:
    - Characterised by usage of open standardized interfaces
    - Examples: IP, Z39.50, LDAP, IMAP, CORBA, Posix
  - If you just want to get one job done -- then you can build an application
    - Examples: Microsoft Windows, Microsoft Access, DCOM
  - Mature, successful, shareable applications often migrate to infrastructure

Representation of biodiversity content

eXtended Markup Language

- XML is changing the way information is exchanged.
- Concepts
  - HTML: presentation on the web, but no knowledge of content (see example)
  - XML: knowledge of content (see example)
  - XSL: mapping content to presentation
  - CSS: cascading style sheets
  - XMLT: XML Transformation
  - XML Schemas: standardisation of content for domains
    - New approach to metadata
    - Need for standard schemas for biodiversity so that remote XML pages can be queried.

Registering and locating biodiversity content

SpeciesBank requirements

- Link to any accessible existing database that holds information about species.
- Facilitate searching of Internet resources by non-specialists.
- Assist taxonomists to avoid re-naming already-described species.
- Enable rapid dissemination of information on newly discovered species.
- Speed up repatriation of information about species native to developing world.
- Increase the rate at which new species are described.
- Enhance accessibility of species information to users.

SpeciesBank technical design

- Central registry
  - Links to distributed content like Napster
- Central repository
  - Clearing-house database similar to GenBank? Maybe not.
  - Caching, archiving and backup. Certainly!
- How do remote content providers actually form and register their taxon home pages and other services at SpeciesBank?
  - Email notification enough for standardised XML content
  - SpeciesBank protocol and client application?
- Portal services automatically from harvested meta-information
- What is the link between CNKO and SpeciesBank? DNS??

Plethora of available e-business standards

- Infrastructure that motivates users to contribute their pieces of knowledge to a common pool and set up and advertise their services.
- New standards and protocols available from e-business world:
  - ebXML (e-business XML)
  - UDDI (Universal Description and Discovery of Information)
  - SOAP (Simple Object Access Protocol)
  - RSS (Rich Site Summary) → Channels for portals
  - RDF (Resource Description Framework)
  - WSDL (Web Services Description Language)
  - ADS (Advertisement and Discovery of Services)
  - Napster, Freenet, W3C Semantic Net

Back to my talk

An Aside

- Many other scientific communities have offered similar visions (e.g., caBIG), but

An Aside

- Many other scientific communities have offered similar visions (e.g., caBIG), but
- No one has yet implemented a working, production-quality solution.

An Aside

- Many other scientific communities have offered similar visions (e.g., caBIG), but
- No one has yet implemented a working, production-quality solution.
- The problems are…

An Aside

- Technical (unresolved problems remain)
  - Inter-database referential integrity
  - Global, resolvable identifiers
  - Distributed joins
  - Tri-state logic
  - Support for inductive reasoning
  - Inadequate data models
  - Identity and identification

An Aside

- Social
  - Scientific beliefs are always in debate
  - There is no one true science
  - Context affects meaning

An Aside

- Practical
  - Resource limitations
  - Resource limitations
  - Resource limitations

The PROBLEM

- Thousands of different, uncoordinated sources of biodiversity data.
- No common formats, no standards.

The NEED

- A single, coherent information infrastructure for delivering biodiversity data.

The PLAN

- Analyze the problem.
- Articulate a vision for the future.
- Perform feasibility analyses.
- Produce a design.
- Implement SpeciesBANK using modern information-infrastructure tools.

The GOAL

In parallel to the molecular database GenBank (but operating on completely different principles), GBIF envisions a future in which all sorts of information about any species (gene sequences, occurrence in ecosystems, specific locality data, ecological relationships, physiological requirements and so on) would be compiled on demand from many, disparate, continuously updated databases.

SpeciesBANK would effectively be an encyclopedia of species that is continuously filling in missing or supplanting outdated information.

The Future

Very Promising

Technological Advances

- Ubiquitous, high-speed networks.
- Global data-exchange standards.
- Multi-media tools readily available.

Technological Advances

NETWATCH

edited by Mitch Leslie

IMAGES

A Bird in Hand

For a nifty take on how museum collections can benefit from cyberspace, check out this digital specimen case from the Zoological Museum Amsterdam in the Netherlands. The site supplies three-dimensional (3-D) images of 151 avian type specimens from around the world—the original examples taxonomists used to describe the species. You can rotate or tilt animals ranging from crows and owls to this black-capped lory (*Lorius lory viridicrissalis*, right) from Indonesia. The pages also describe where and when the birds were collected, provide their measurements, and compare them to other specimens. The museum plans to post similar 3-D images of its cache of shells and skulls.

www.science.uva.nl/ZMA/3dpics

EDUCATION

Fire Up the Virtual Bunsen Burner

Demonstrating chemical reactions in class is a great way to spark students' interest—assuming the procedures work, everyone can see the lab bench, and nobody gets hurt. An alternative that eliminates these potential problems is this library of some 200 experiments for undergraduate labs from the Swiss Federal

SVAVA HELENAND LINUS PAULING PAPERS

DATABASE

Cytochrome Central

People who inherit a particular version of the gene *CYP2D6* don't get help from standard doses of the pain reliever codeine and can suffer side effects from many other medications. The problem is a sluggish drug-detoxifying enzyme from the cytochrome P450 family. This database from molecular biologist David Nelson of the University of Tennessee, Memphis, can help researchers get a handle on this sprawling group of enzymes, which take part in everything from breaking down Prozac and caffeine to synthesizing cholesterol.

The site lists more than 4000 versions of cytochrome P450 enzymes gleaned from published genomes of humans, honeybees, slime molds, bacteria, and other creatures. The sequences come in standard format, so you can plug them directly into genome analysis software or compare your sequences to those already on the site. For more information about cytochrome P450s, check out transcripts of Nelson's lectures or take a guided tour of some P450 molecules (above, *CYP2C5*).

drnelson.utmem.edu/CytochromeP450.html

ZMA - Bird type specimens online - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address http://ip30.eti.uva.nl/zma3d/home.html Go Links

Google type specimen amsterdam Search Web 0 blocked AutoFill Options type

Bananaquit (*Coereba f. bonairensis*)
ZMA 11091 (paratype)

Click once on the bird to stop the rotation, double-click to restart

Zoological Museum Amsterdam
Bird type specimens online

This website provides direct access to 3D images of 151 type specimens in the bird collection of the **Zoological Museum Amsterdam**. If you see a rotating bird specimen at the left, you are ready to start exploring. If you see a "broken" logo instead, please visit the **Help** page to get started.

All 3D images can be accessed from the **bird type index**, which also lists the most recently updated specimens. The families listed are arranged in the Wetmore-order (Wetmore 1940); within families, taxa are arranged in sequence of description. Taxa belonging to the same species are grouped together for convenience. Information about the specimens has been taken from Roselaar & Prins (2000). If necessary, a list of **museum acronyms** and relevant **literature** can be consulted.

The history of the bird collection is described under **ZMA bird collection**. The collection manager, Tineke G. Prins, can be contacted at **tprins@science.uva.nl**. The complete bird collection, as well as other ZMA collections, can be searched at the Data Portal of the **GBIF** website.

If you have visited the **Help** page and still experience technical problems, please contact the **webmaster**.

Internet

Access to images of type specimens…

Eclectus cornelia - Eclectus Parrot type in the ZMA - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address http://ip30.eti.uva.nl/zma3d/psittacidae.html Go Links

Google type specimens amsterdam Search Web Search Site Options type spec

Home | Bird type index | ZMA bird collection | Search specimens | Links | Help

Psittacidae

Eclectus cornelia
Bonaparte, 1849: 142

Current name:
Eclectus roratus cornelia

Eclectus Parrot
1 holotype?

Holotype(?): ZMA 261, adult female, undated, ex coll. Amsterdam Zoo; wing 260, tail 150, bill 39.4, tarsus 27.5.

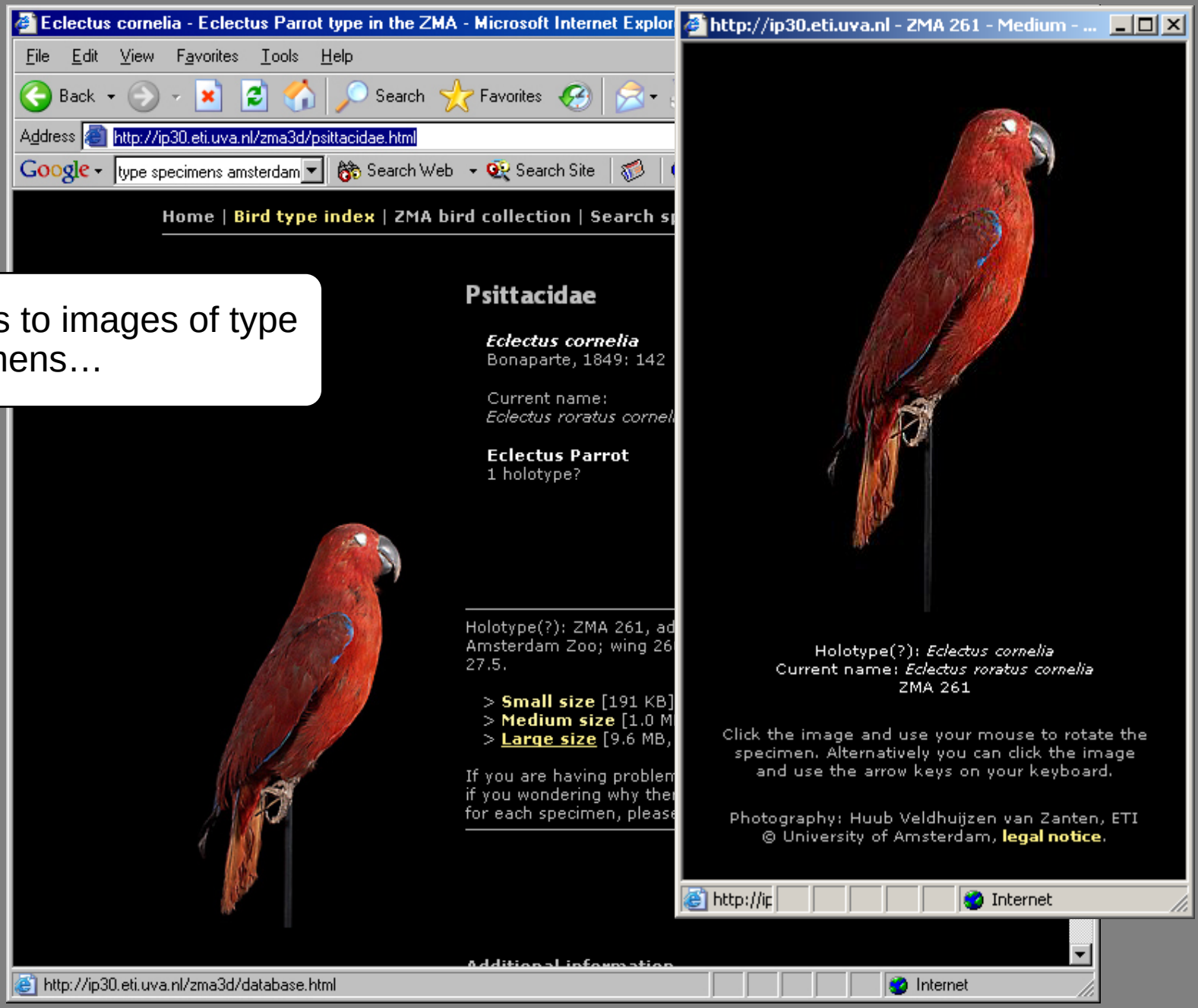
> **Small size** [191 KB]
> **Medium size** [1.0 MB]
> **Large size** [9.6 MB, large download!]

If you are having problems opening the 3D images, or if you wondering why there are three types of movies for each specimen, please visit the **Help** page.

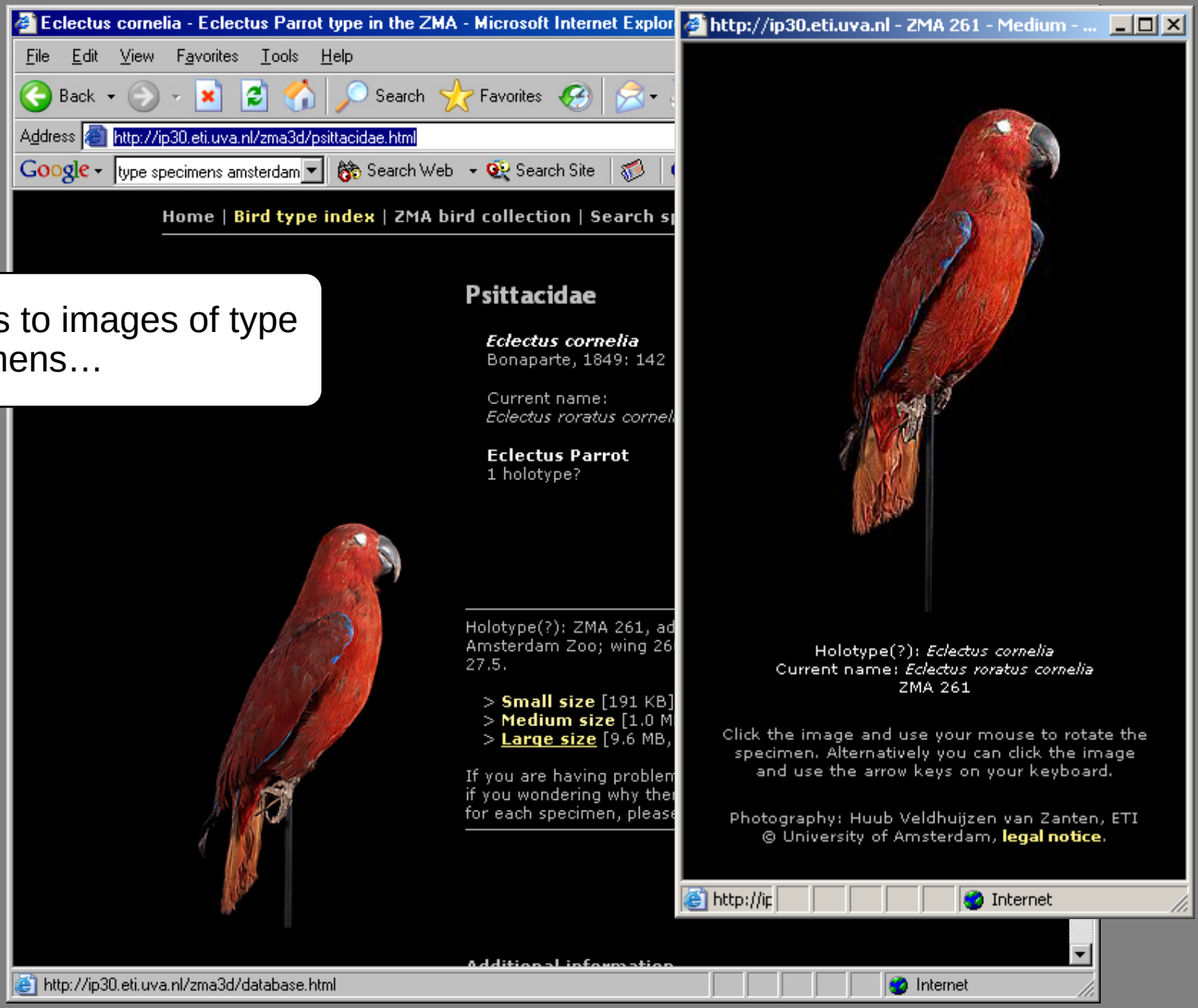
Additional information

http://ip30.eti.uva.nl/zma3d/database.html Internet

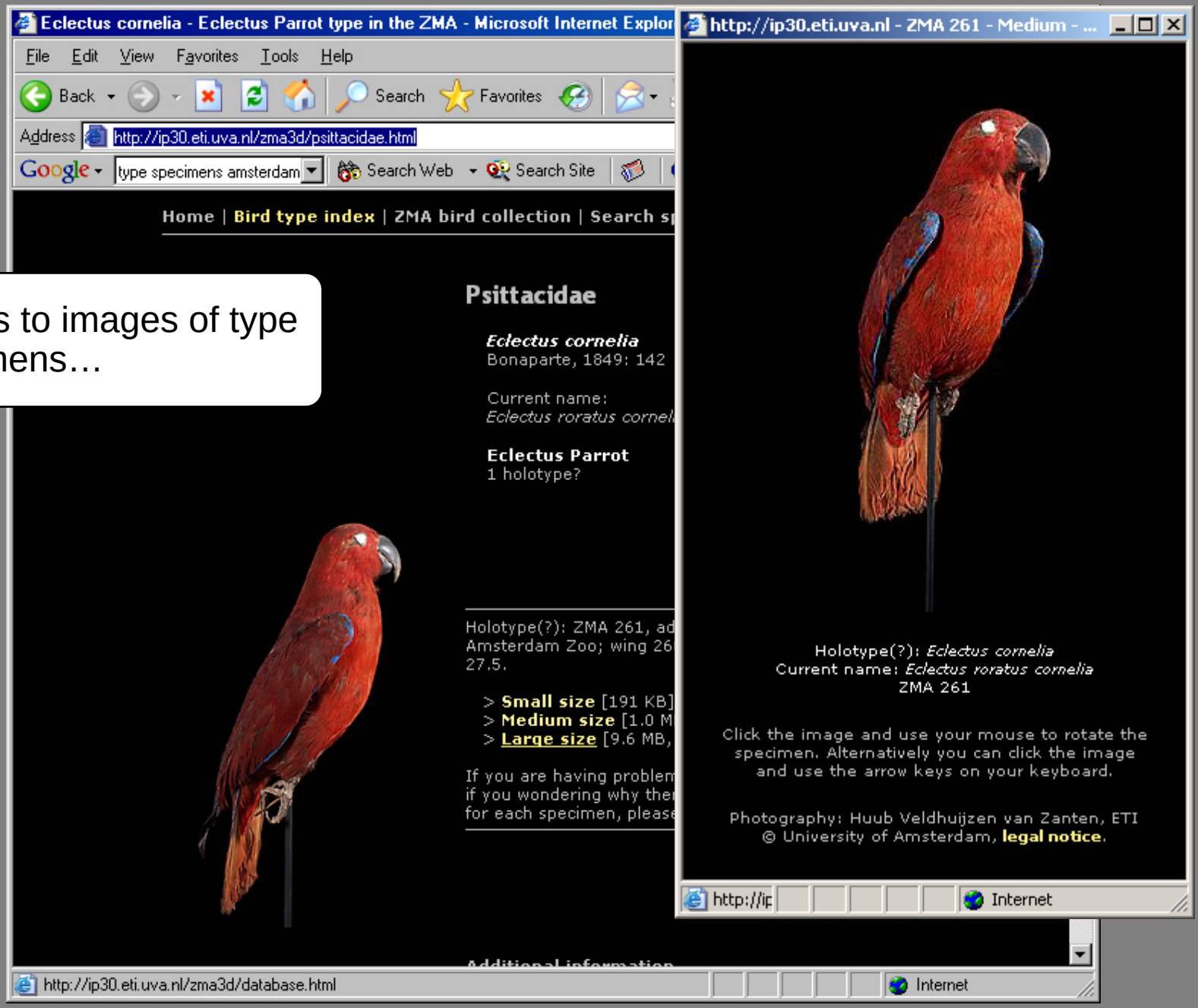
Access to images of type specimens…



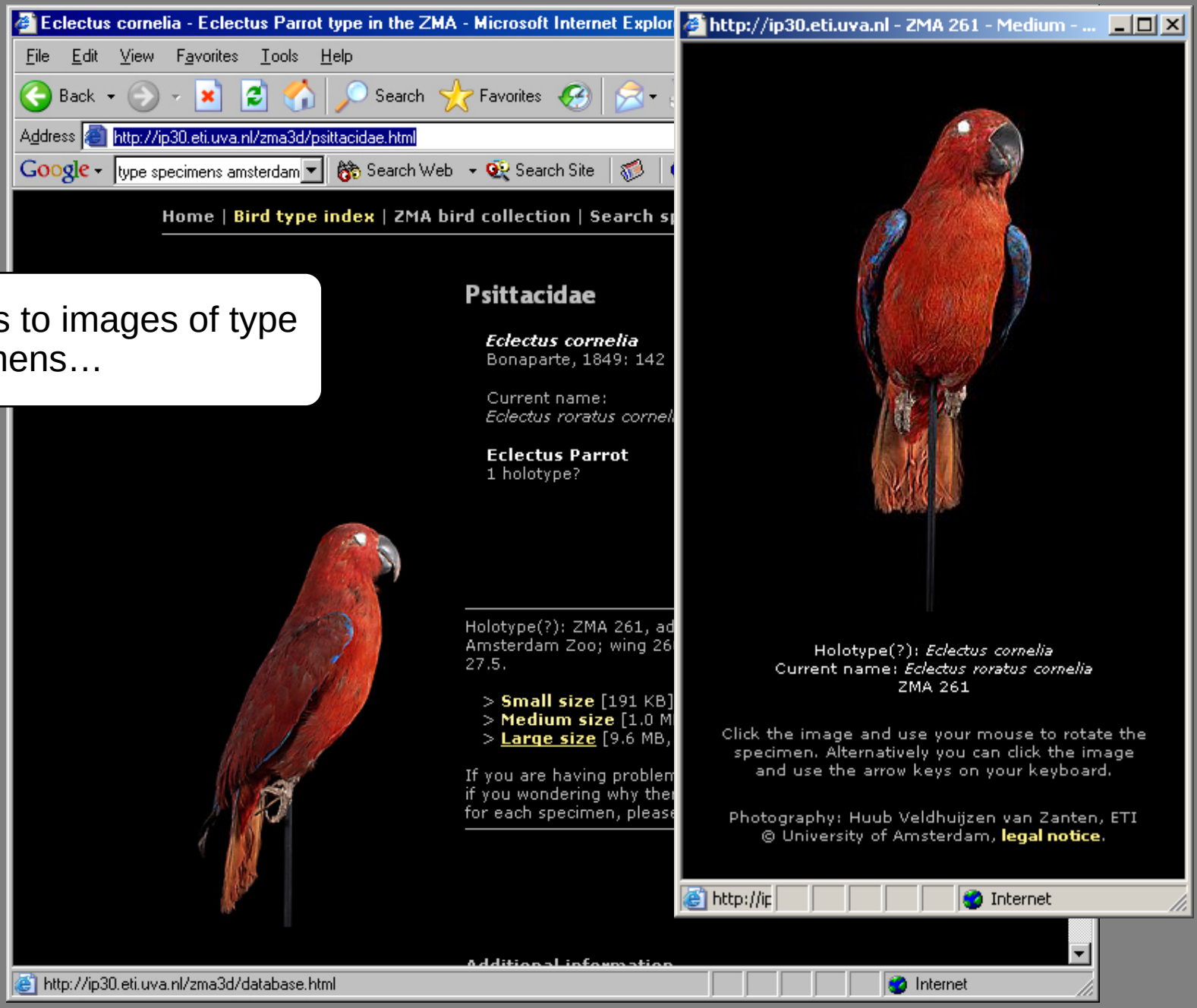
Access to images of type specimens…



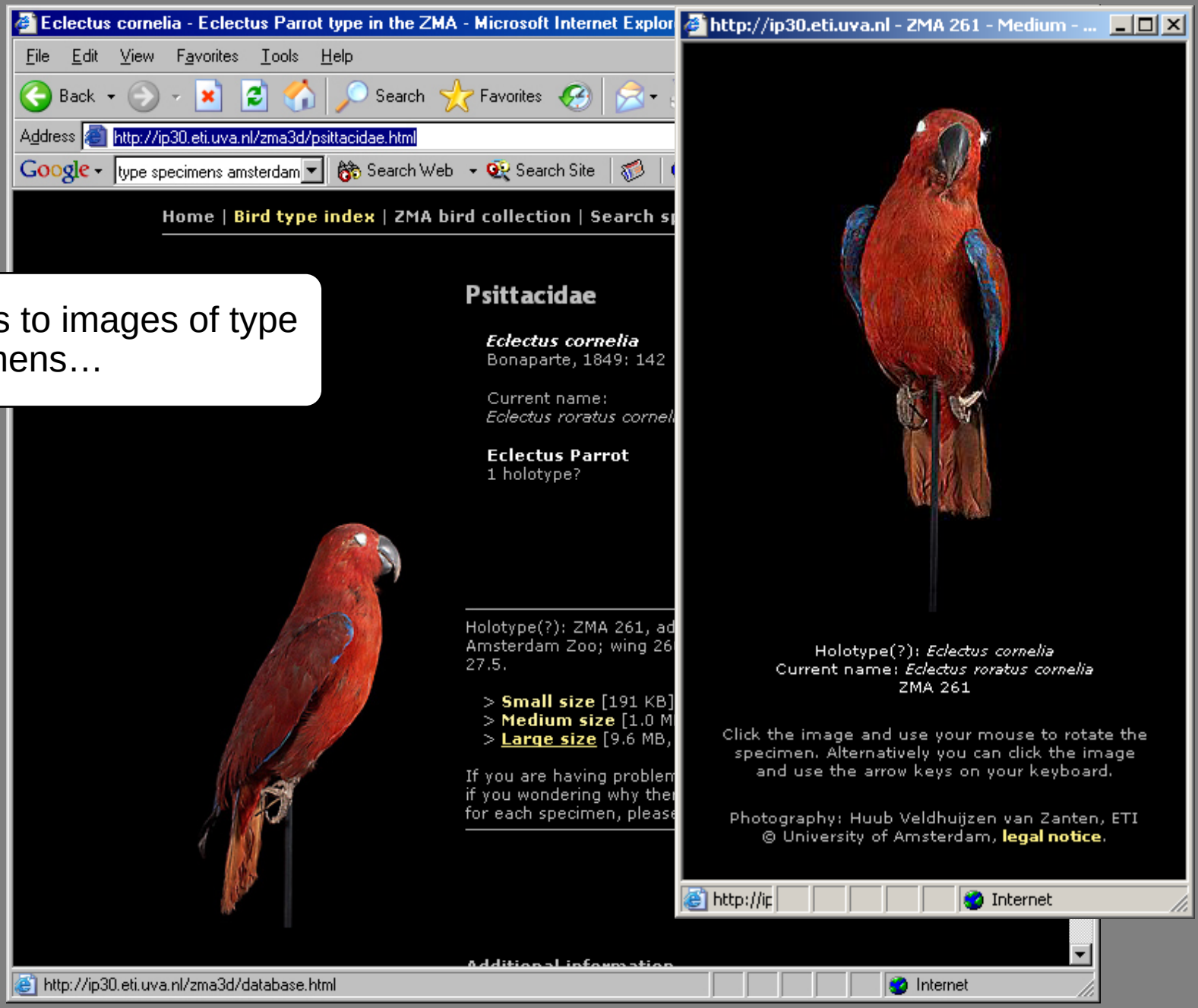
Access to images of type specimens…



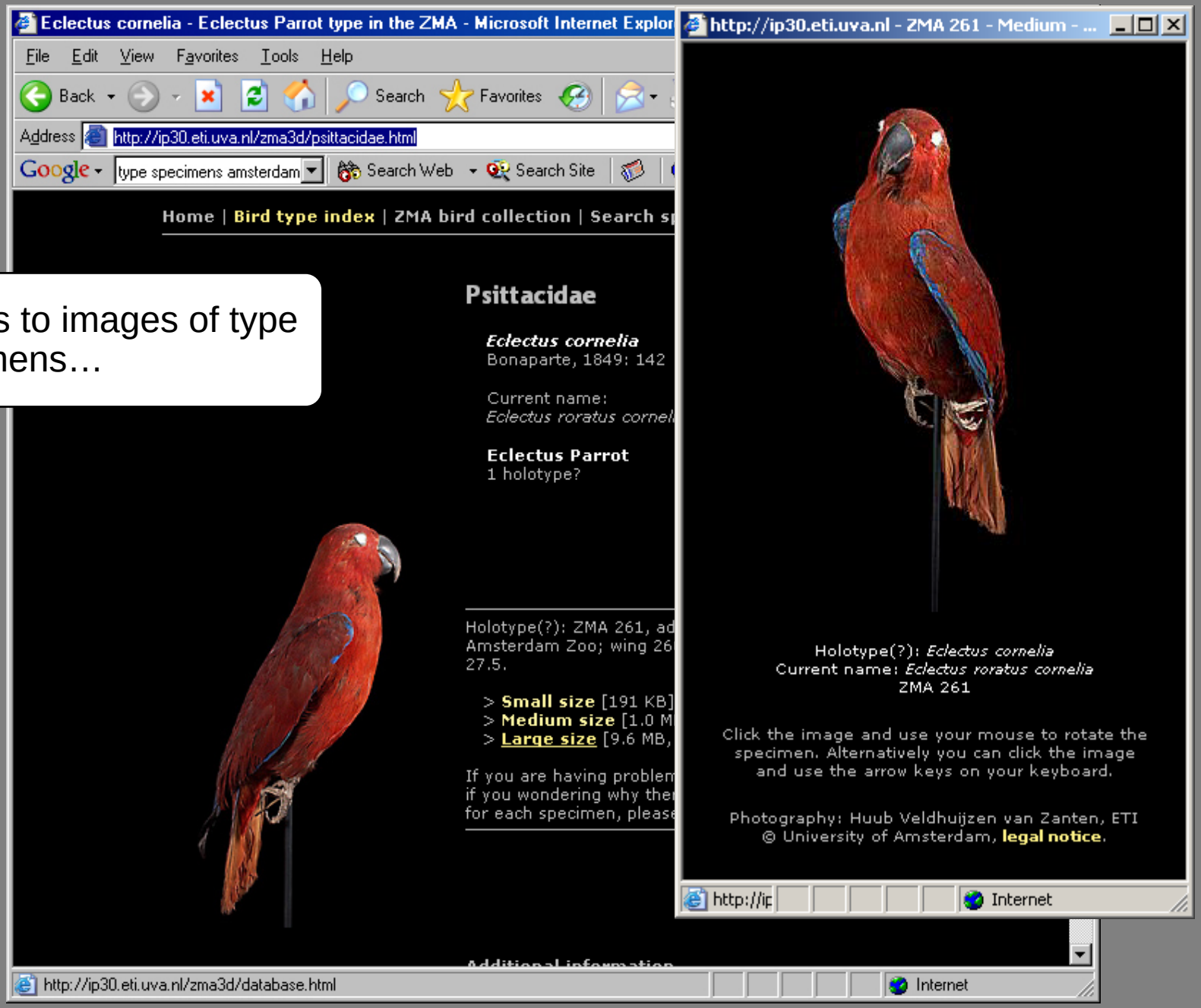
Access to images of type specimens…



Access to images of type specimens…



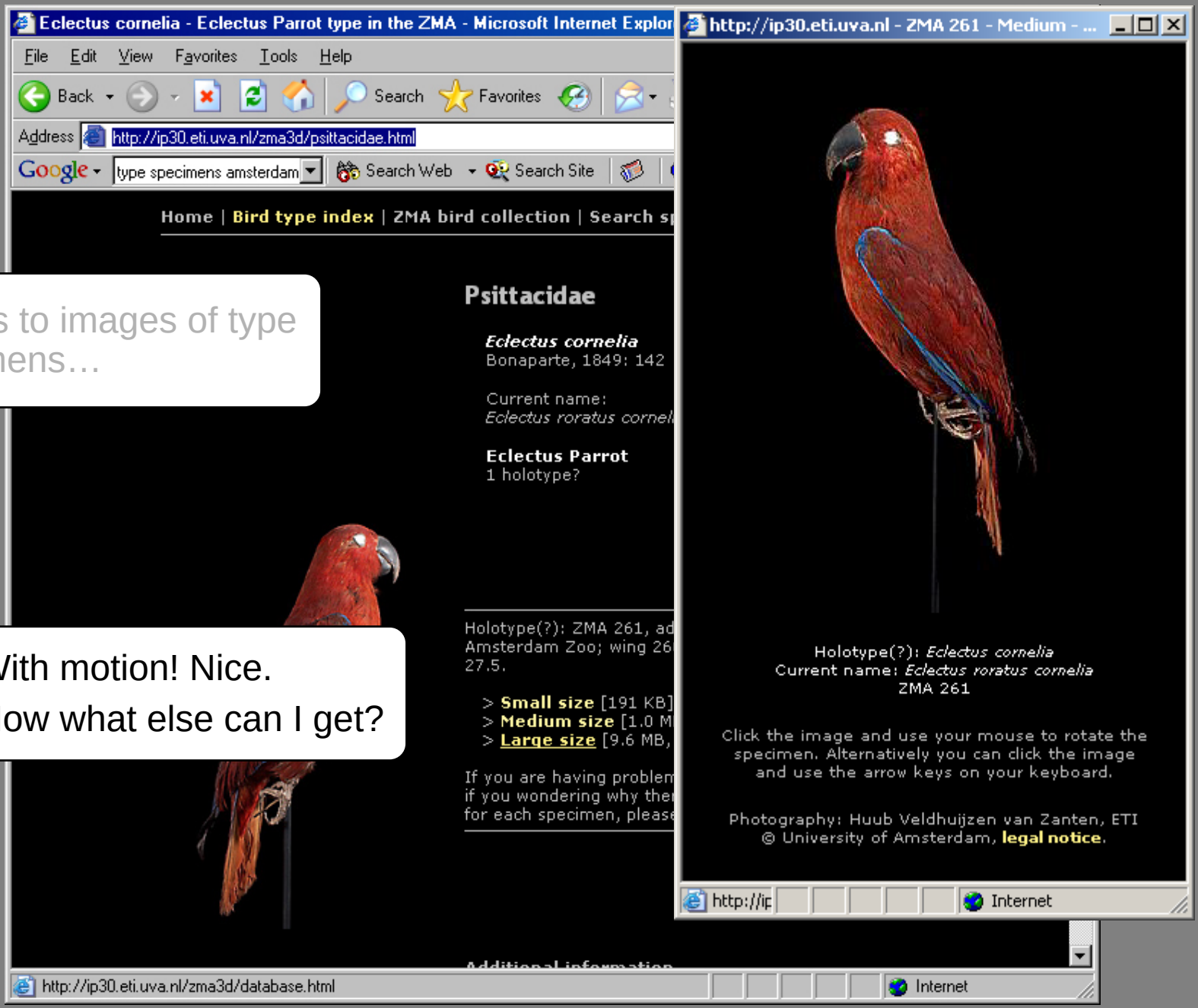
Access to images of type specimens…



Access to images of type specimens…

With motion! Nice.

Now what else can I get?

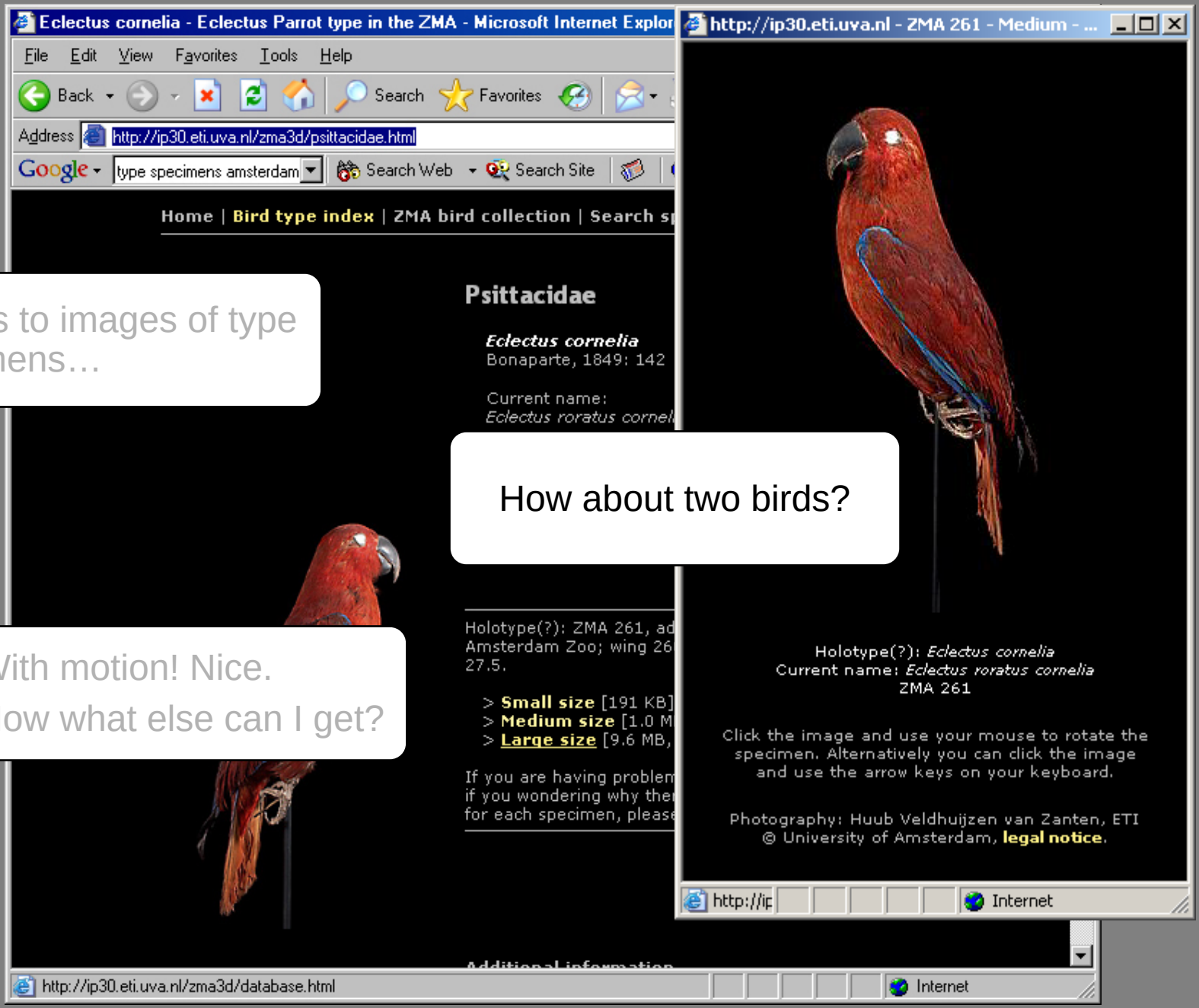


Access to images of type specimens…

With motion! Nice.

Now what else can I get?

How about two birds?



X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

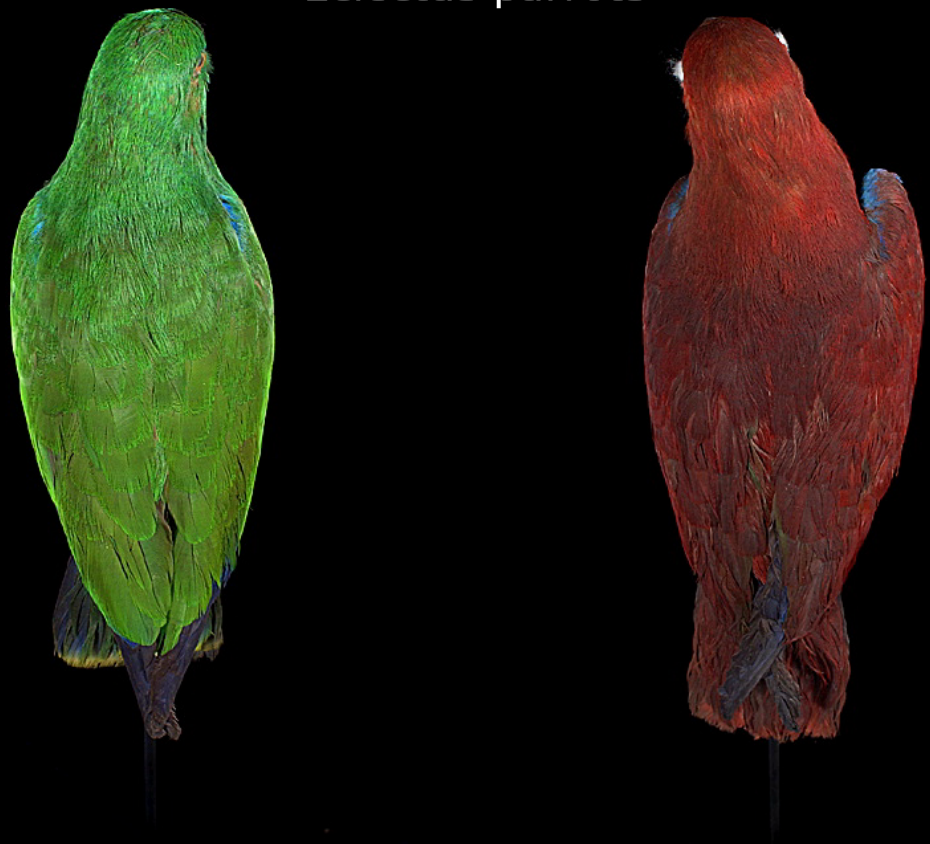
File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots



Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

Done Local intranet

X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html Go Links

Google Search Web 235 blocked AutoFill Options

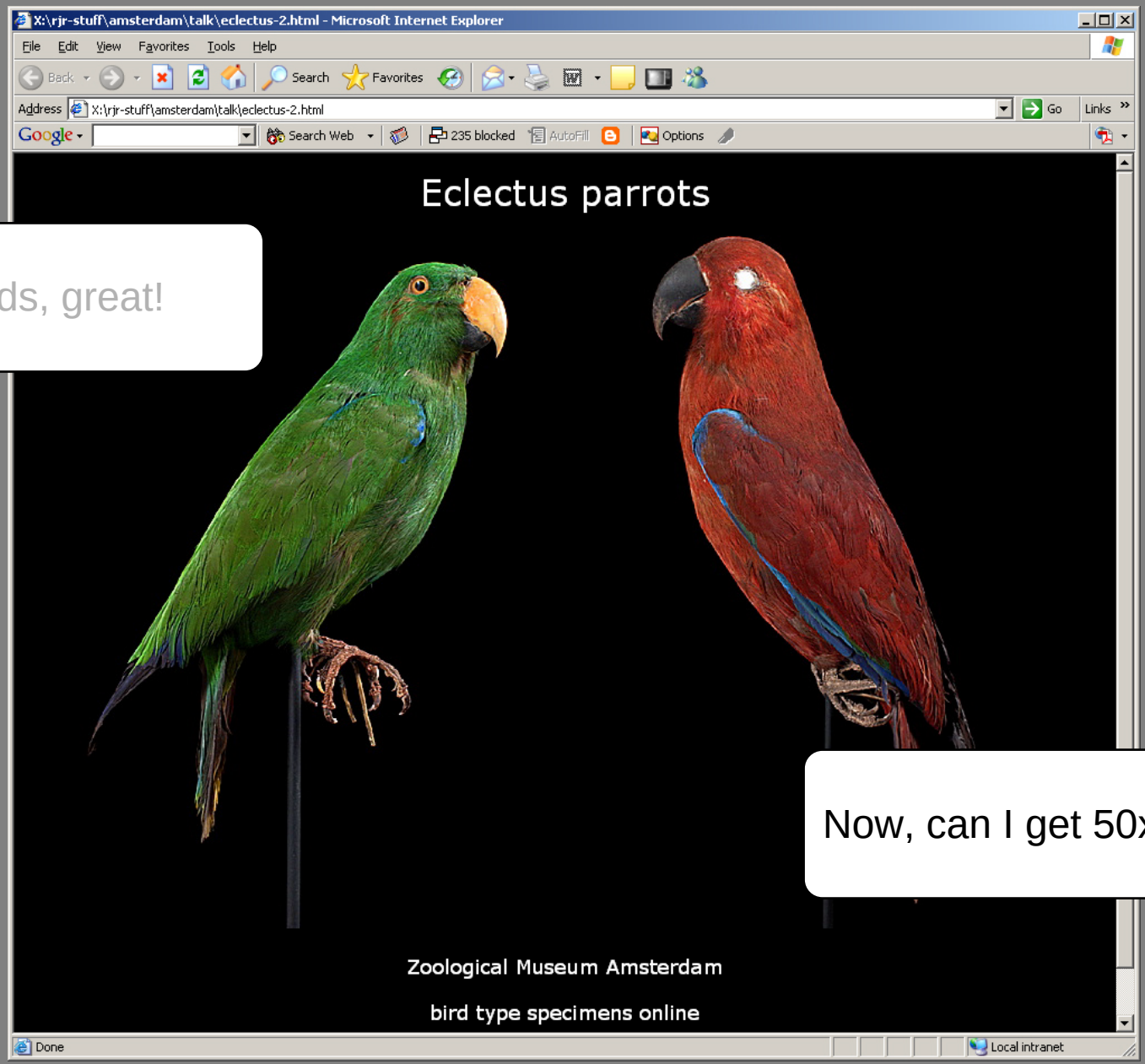
Eclectus parrots

Zoological Museum Amsterdam

bird type specimens online

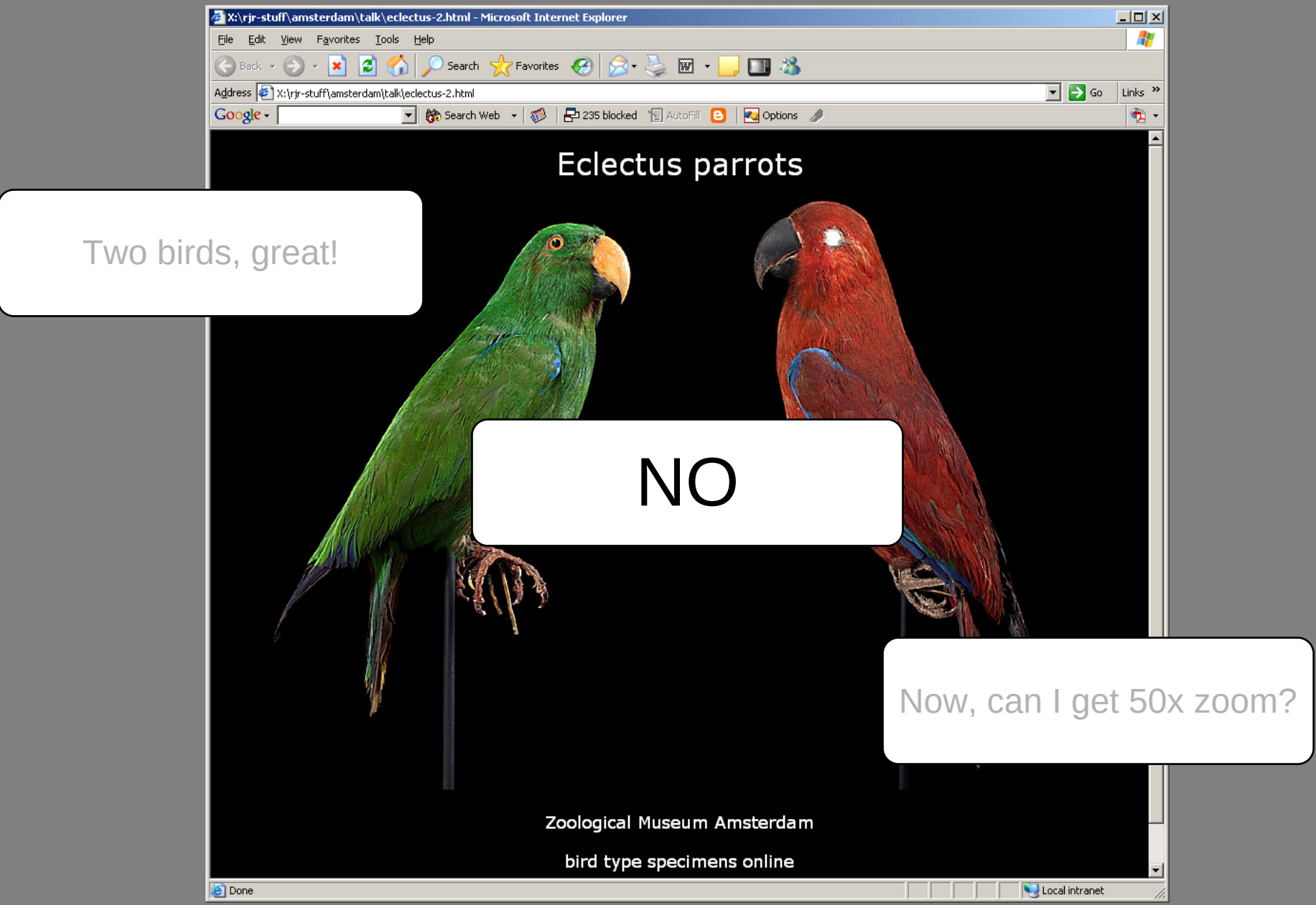
Done Local intranet

Two birds, great!



Two birds, great!

Now, can I get 50x zoom?


X:\rjr-stuff\amsterdam\talk\eclectus-2.html - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address X:\rjr-stuff\amsterdam\talk\eclectus-2.html
Go Links
Google Search Web 235 blocked AutoFill Options
Eclectus parrots
Two birds, great!
NO
Now, can I get 50x zoom?
Zoological Museum Amsterdam
bird type specimens online
Done
Local intranet

Managing Scope

At some point you have to stop accepting new requirements.

At some point you have to ask

What's possible?

What's not?

At some point you have to compare your goals and your resources (including time) and decide on the SCOPE of the project:

What will be IN SCOPE and what will be OUT OF SCOPE.

You must set boundaries and manage expectations.

Managing Scope

In building production systems, deciding what

WILL NOT BE DONE

is just as important as deciding what

WILL BE DONE

You must set boundaries and manage expectations.

SCOPE Problem

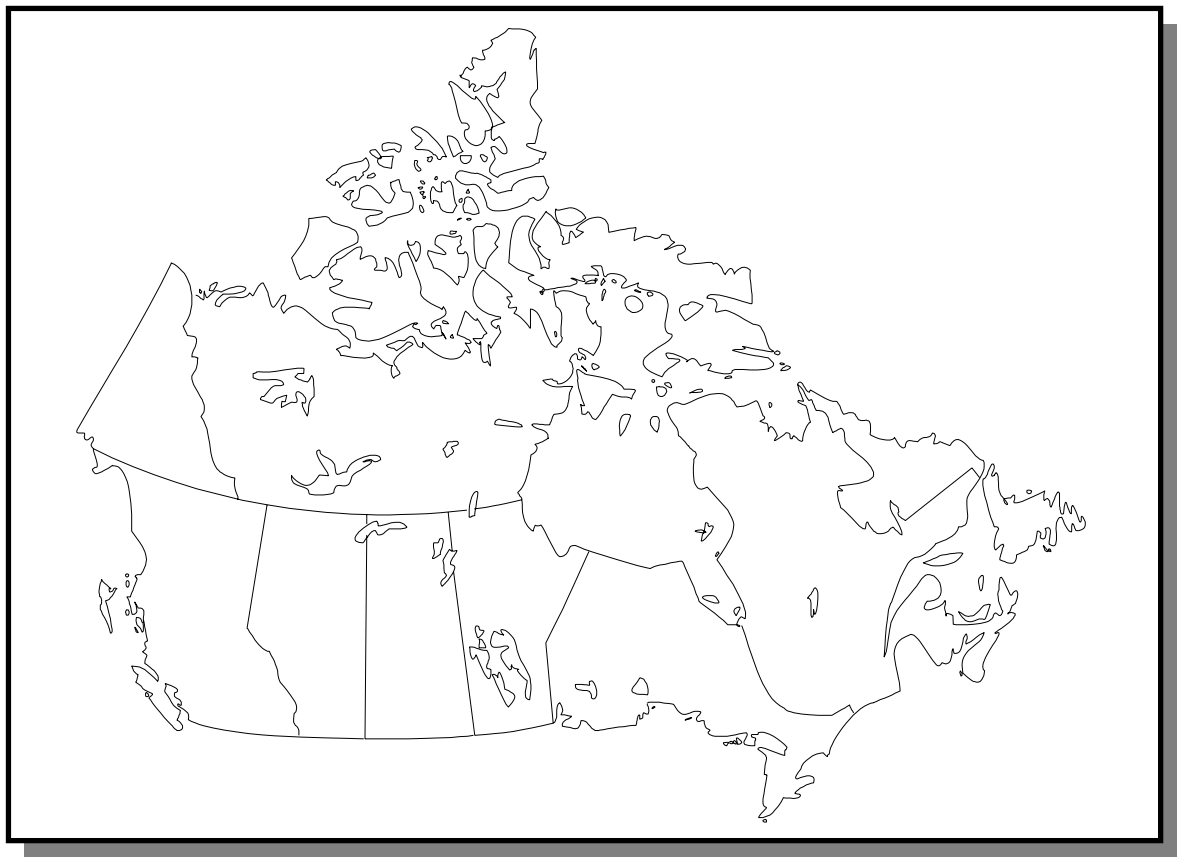
Documenting Global Biodiversity

Documenting Biodiversity

- Documenting global biodiversity will require access to global data sets on:
  - species diversity
  - species distribution and density
  - environmental parameters
  - times series records of biological and environmental data
  - genetic diversity within species
  - individual differences in gene expression
  - more…

Documenting Biodiversity

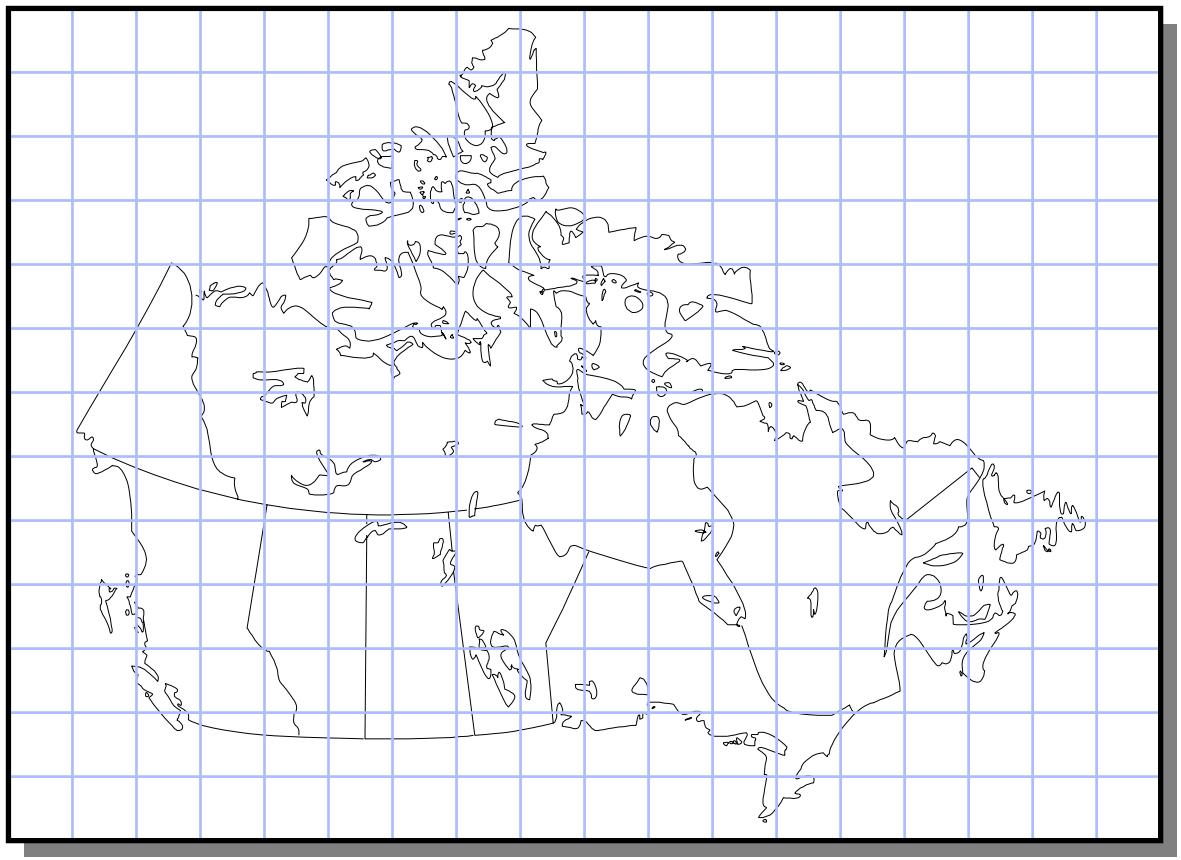
At an elementary level, documenting biodiversity involves tracking species presence/absence per unit of area.



Documenting Biodiversity

At an elementary level, documenting biodiversity involves tracking species presence/absence per unit of area.

But at what resolution?

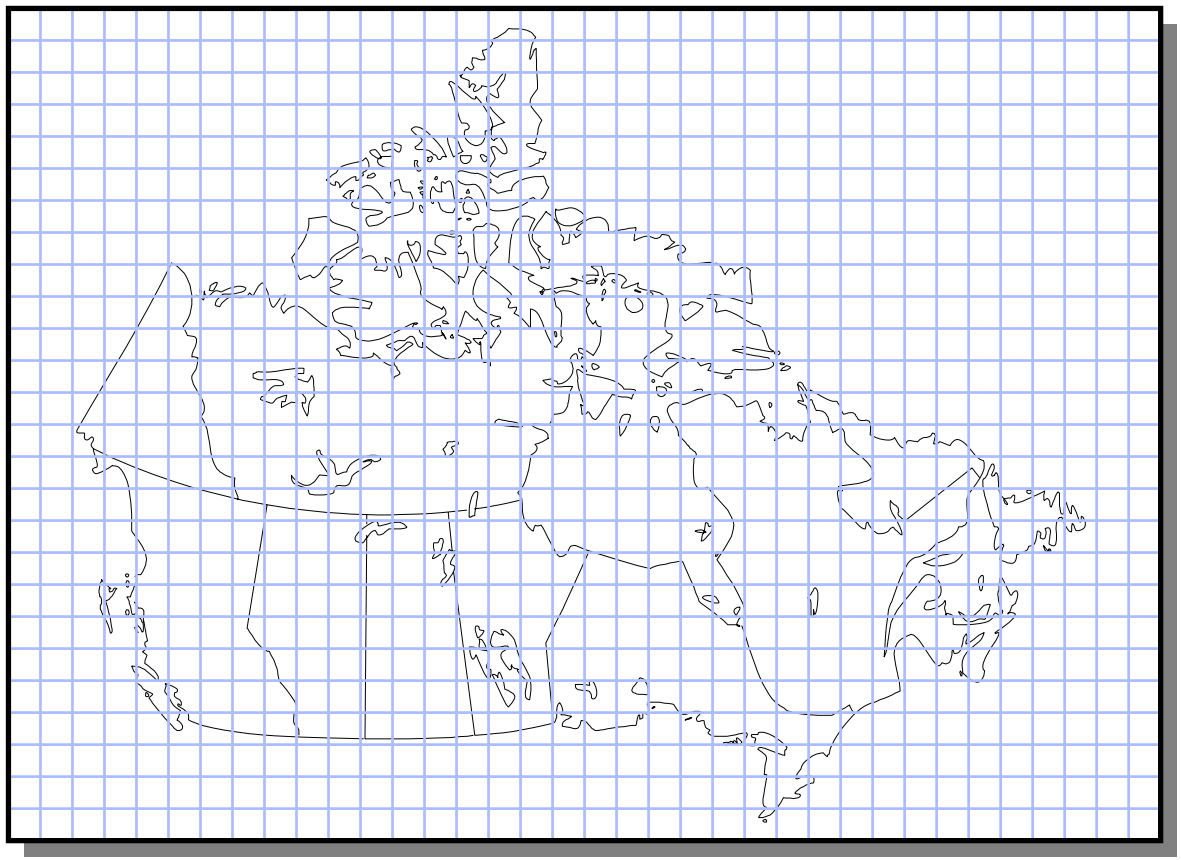


Documenting Biodiversity

At an elementary level, documenting biodiversity involves tracking species presence/absence per unit of area.

But at what resolution?

With every increase in resolution, the data set grows exponentially...

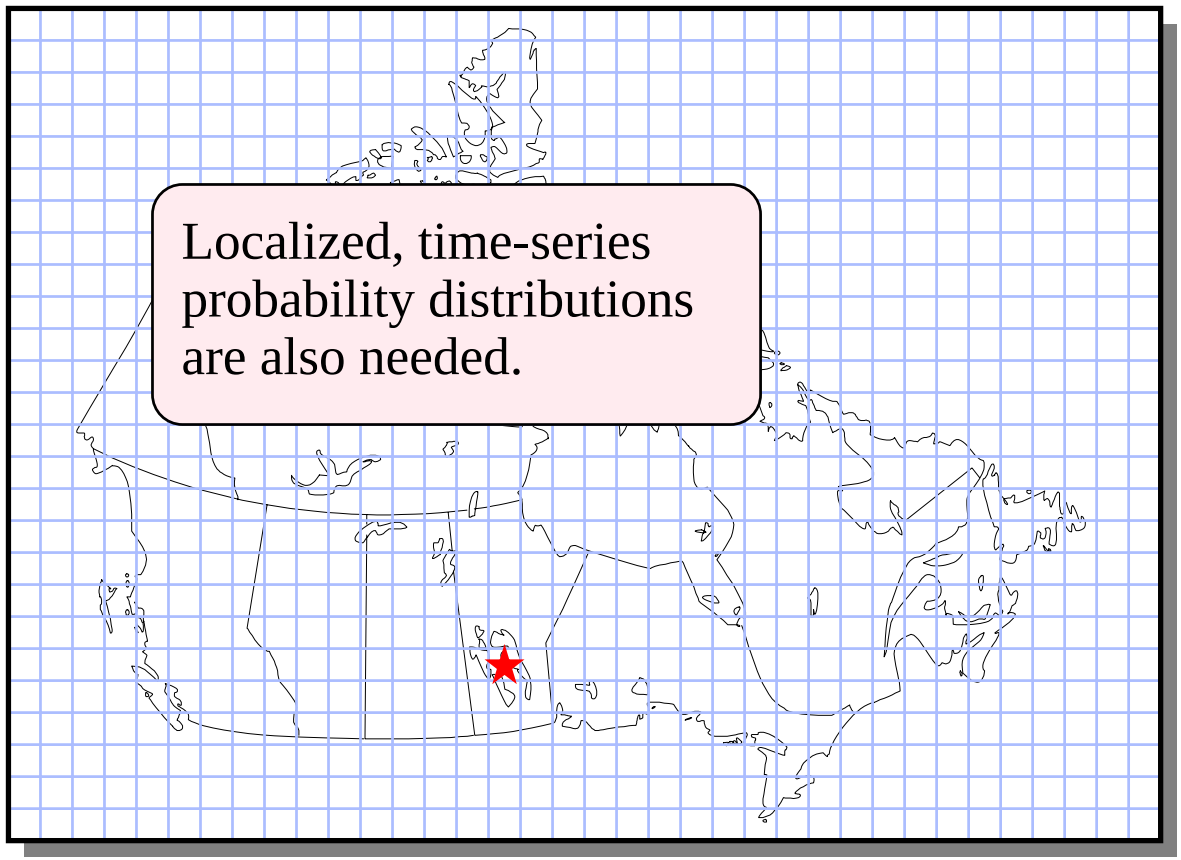


Documenting Biodiversity

At an elementary level, documenting biodiversity involves tracking species presence/absence per unit of area.

But at what resolution?

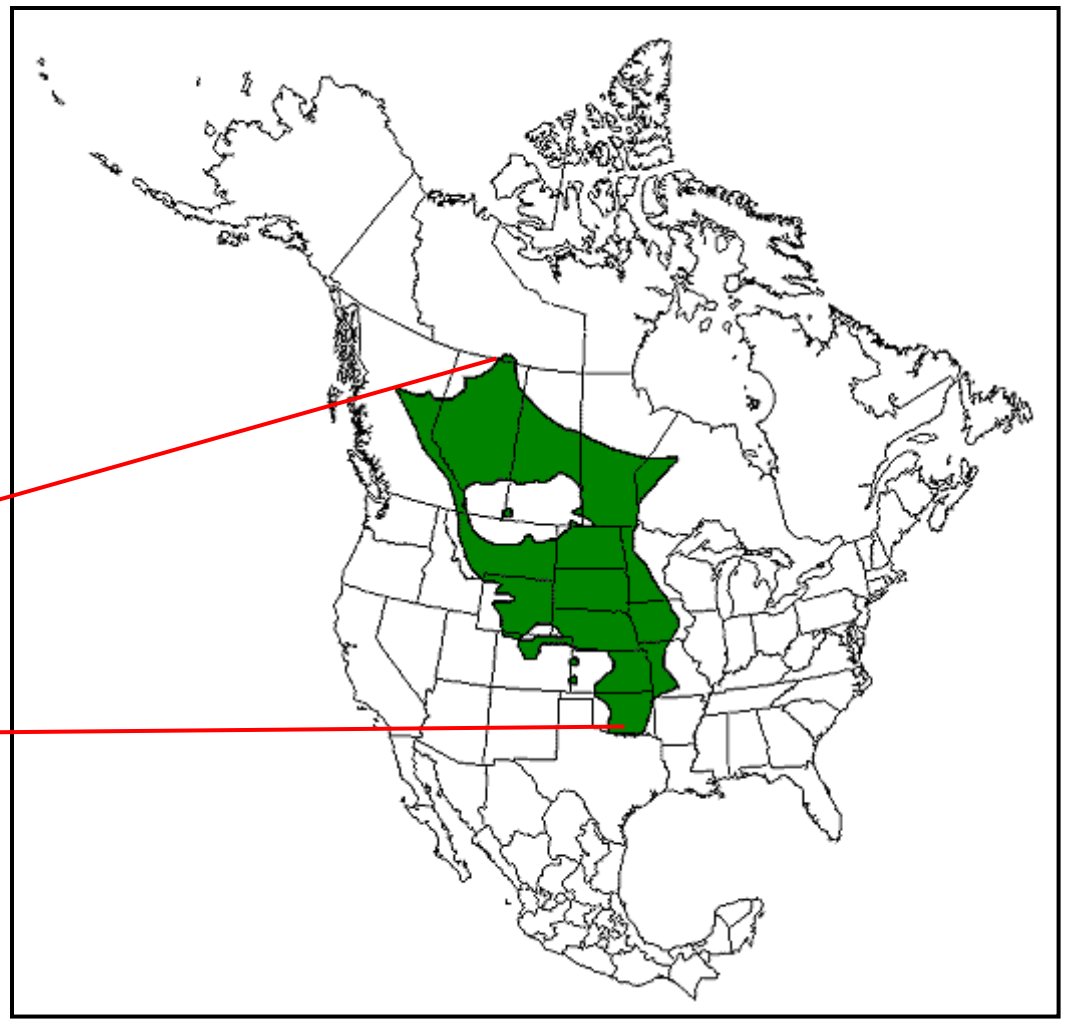
With every increase in resolution, the data set grows exponentially...



SpeciesBANK

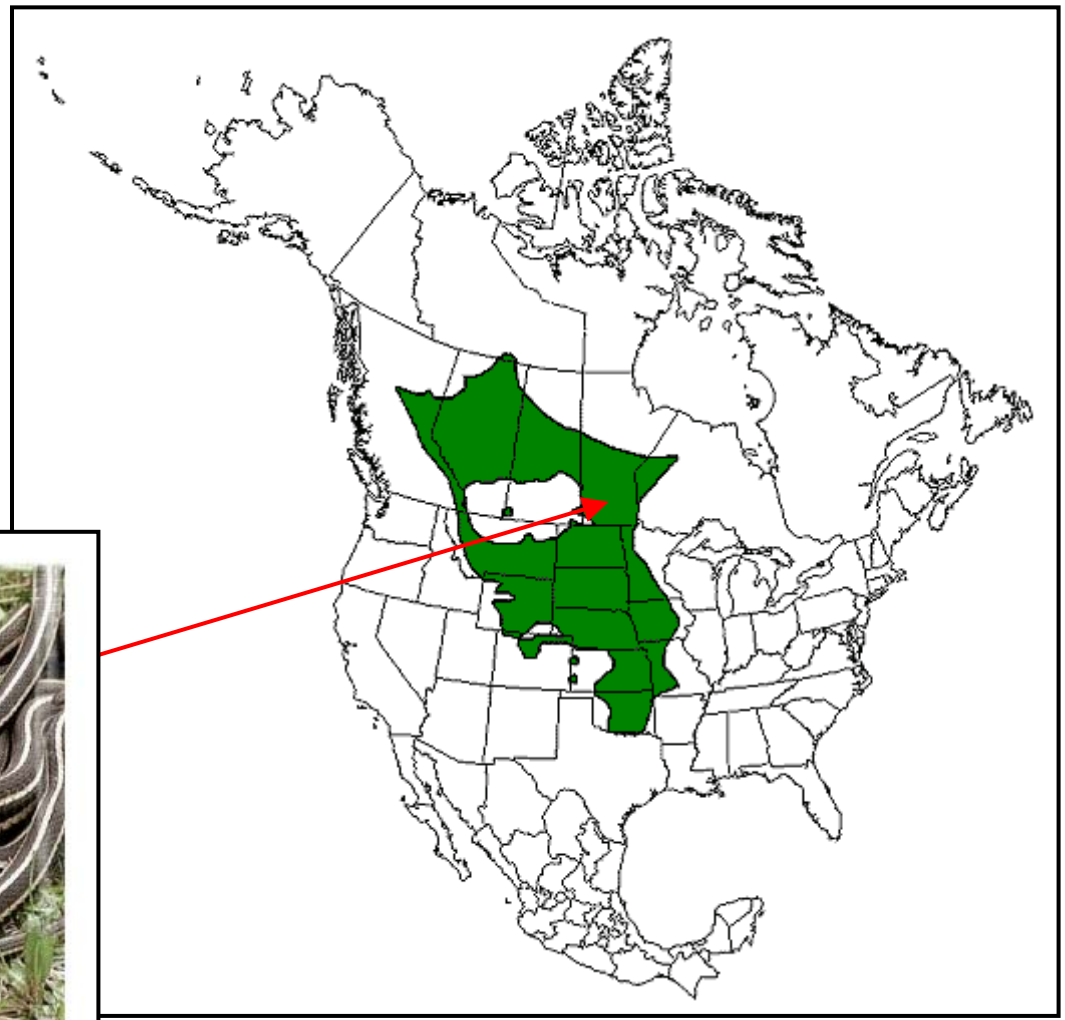
It's one thing to say that:

The red-sided garter snake occurs throughout central North America.



It's another to note:

Every fall and spring, more than 65,000 red-sided garter snakes congregate at local over-wintering dens in southern Manitoba, resulting in the most locally dense concentration of snakes in the world.



Documenting Biodiversity

- Surface of the Earth = 10^9 km^2.
- Representing the distribution of just one million species would require a two-dimensional distribution grid with 10^{15} cells.
- Adding the third dimension, at one-km scale, to document the diversity in a bio-film that is, say, ten kilometers thick, would require a three-dimensional grid with 10^{16} cells.
- Storing 1 byte of information for each of 10,000 species per cell requires 10^{20} bytes of storage.

Documenting Biodiversity

- Surface of the Earth = 10^{15} m^2.
- Representing the distribution of just one million species would require a two-dimensional distribution grid with 10^{21} cells.
- Adding the third dimension, at one-meter scale, to document the diversity in a bio-film that is, say, ten kilometers thick, would require a three-dimensional grid with 10^{26} cells.
- Storing 1 byte of information for each of 10,000 species per cell requires 10^{30} bytes of storage.

Documenting Biodiversity

- Surface of the Earth = 10^{15} m^2.

Let's compute:

10^{30} bytes = 10^{18} terabytes

= $4x10^{18}$ 250 gB drives

= $5.5x10^{16}$ cubic feet of disks

increase the complexity substantially.

Documenting Biodiversity

- Surface of the Earth = 10^{15} m^2.

Let's compute:

10^{30} bytes = 10^{18} terabytes

= 4×10^{18} 250 gB drives

= 5.5×10^{16} cubic feet of disks

That would cover all Europe to a depth of 160 meters.

increase the complexity substantially.

Documenting Biodiversity

The simple fact is that we will never get the resources necessary to produce PERFECT KNOWLEDGE of the world's species. So, what can we do?

What's possible, what's not?

Figuring that out is our real challenge.

No More Promises

Stand and Deliver

Promises, Promises

I dream of a world with seamless access…

Stand and Deliver

It’s time to deliver a production system.

Challenges/Limits

Challenges/Limits

- Science is constantly changing
- Scientific “facts” are never globally consistent
- Scientific databases are never perfect
- Resources are always limiting
- Needs are constantly changing
- Technology keeps evolving

Challenges/Limits

THE REAL CHALLENGE:

Doing something genuinely useful anyway.

Challenges/Limits

Data Inconsistency

Logic 101

- If premise “A” is false, then the statement “IF A then B” is always true, regardless of the truth value of “B”.

Logic 101

- If premise “A” is false, then the statement “IF A then B” is always true, regardless of the truth value of “B”.
- That is, with a false antecedent you can prove anything.

Logic 101

- If premise “A” is false, then the statement “IF A then B” is always true, regardless of the truth value of “B”.
- That is, with a false antecedent you can prove anything.
- “A and not A” is always false.

Logic 101

- If premise “A” is false, then the statement “IF A then B” is always true, regardless of the truth value of “B”.
- That is, with a false antecedent you can prove anything.
- “A and not A” is always false.
- Feeding inconsistent premises into a logical calculator yields nonsense.

Logic 101

- If premise “A” is false, then the statement “IF A then B” is always true, regardless of
- A and not A” is always false.
- Feeding inconsistent premises into a logical calculator yields nonsense.

Seamless access to inconsistent data is a bad idea.

Challenges/Limits

The GenBank Model

The GenBank Model

- In some ways, GenBank provides a good model for other bioinformatics efforts…

The GenBank Model

- In some ways, GenBank provides a good model for other bioinformatics efforts…

Track record of success

Single source for critical data

Integrated query tools

Integration with other relevant data sets

…

The GenBank Model

- In many other ways, GenBank provides a very bad model for other efforts…

The GenBank Model

- In many other ways, GenBank provides a very bad model for other efforts…

Single, trivial data type

Monolithic, data-warehouse mechanism

Highly constrained update mechanism

Huge (and growing) budget

…

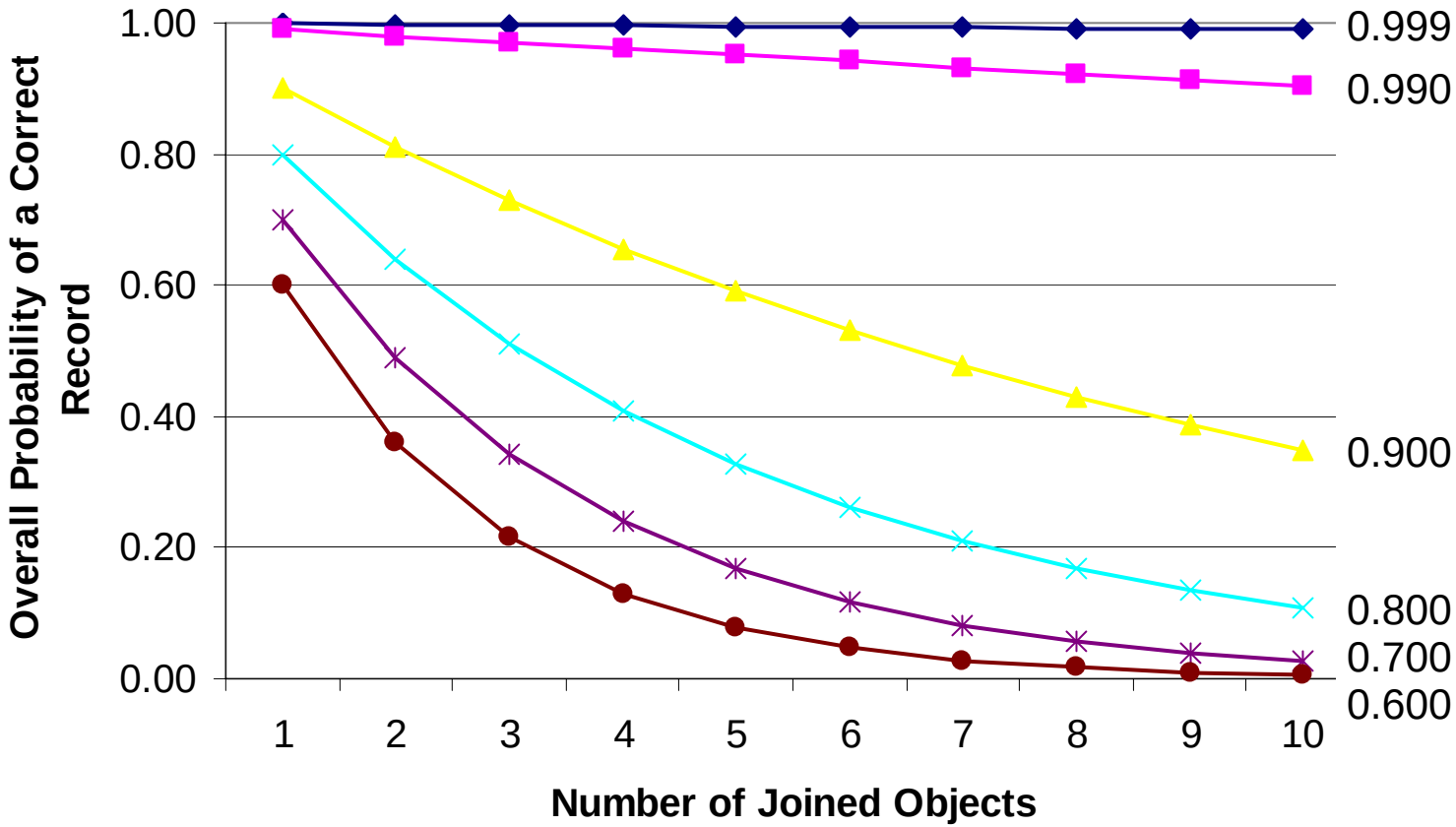
Challenges/Limits

Errors Accrete

Declining Overall Probabilities

- If a “record” in SpeciesBANK is assembled (joined) from data components maintained independently, and
- If the component data collections are not perfect (e.g., the probability of correct = p),
- Then the proportion of completely correct SpeciesBANK records in a query will be given by p^n, where n is the number of elements joined in the query.

Declining Overall Probabilities



As p goes down, p^n goes down a lot faster.

Declining Overall Probabilities

What kinds of error rates (or inconsistency rates) occur in real data sets?

A recent study of human genome data (cytogenetic location of genes), in two large, curated databases, showed an average error rate of 0.1, giving p = 0.9.

What about some species data?

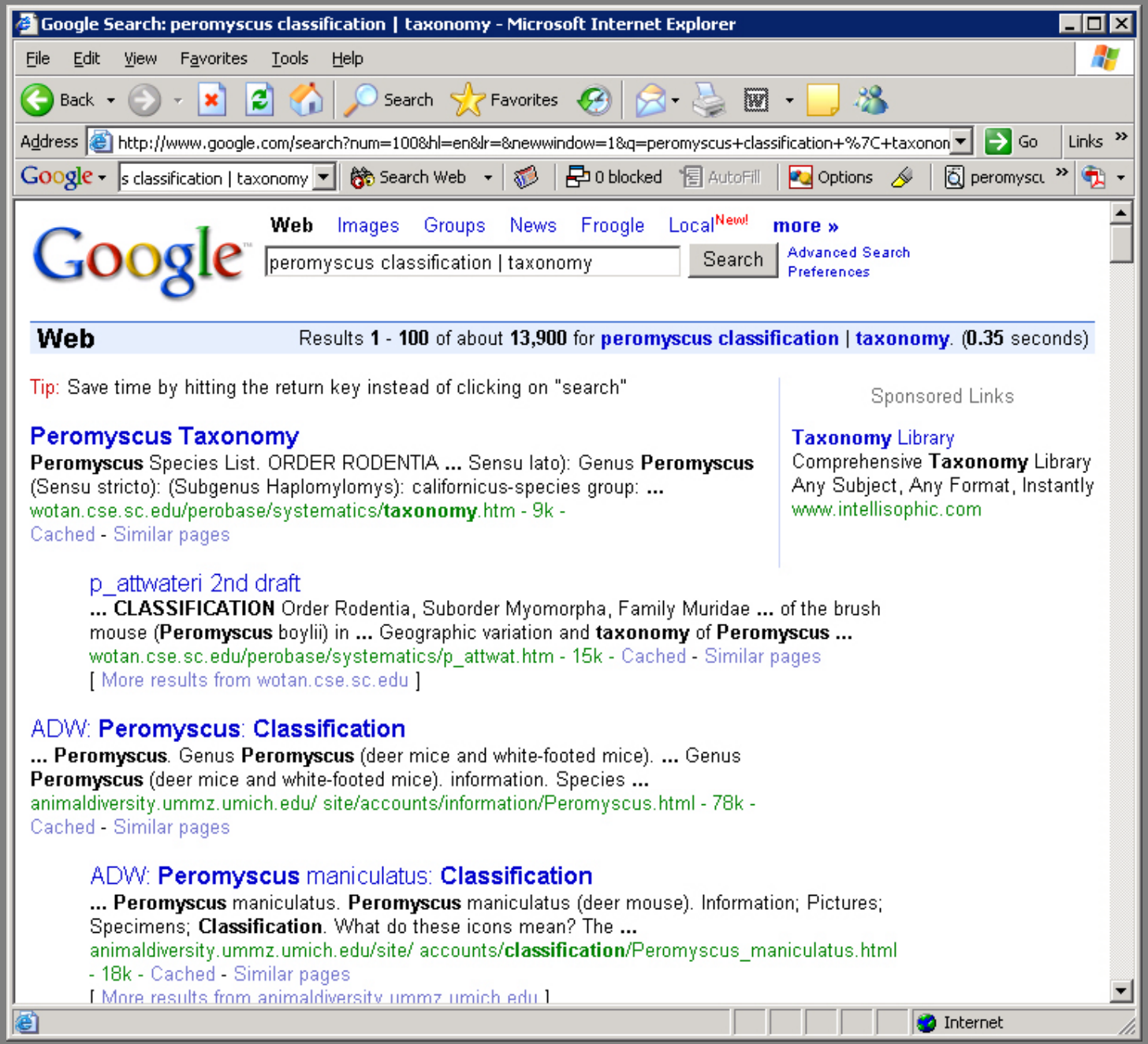
Challenges/Limits

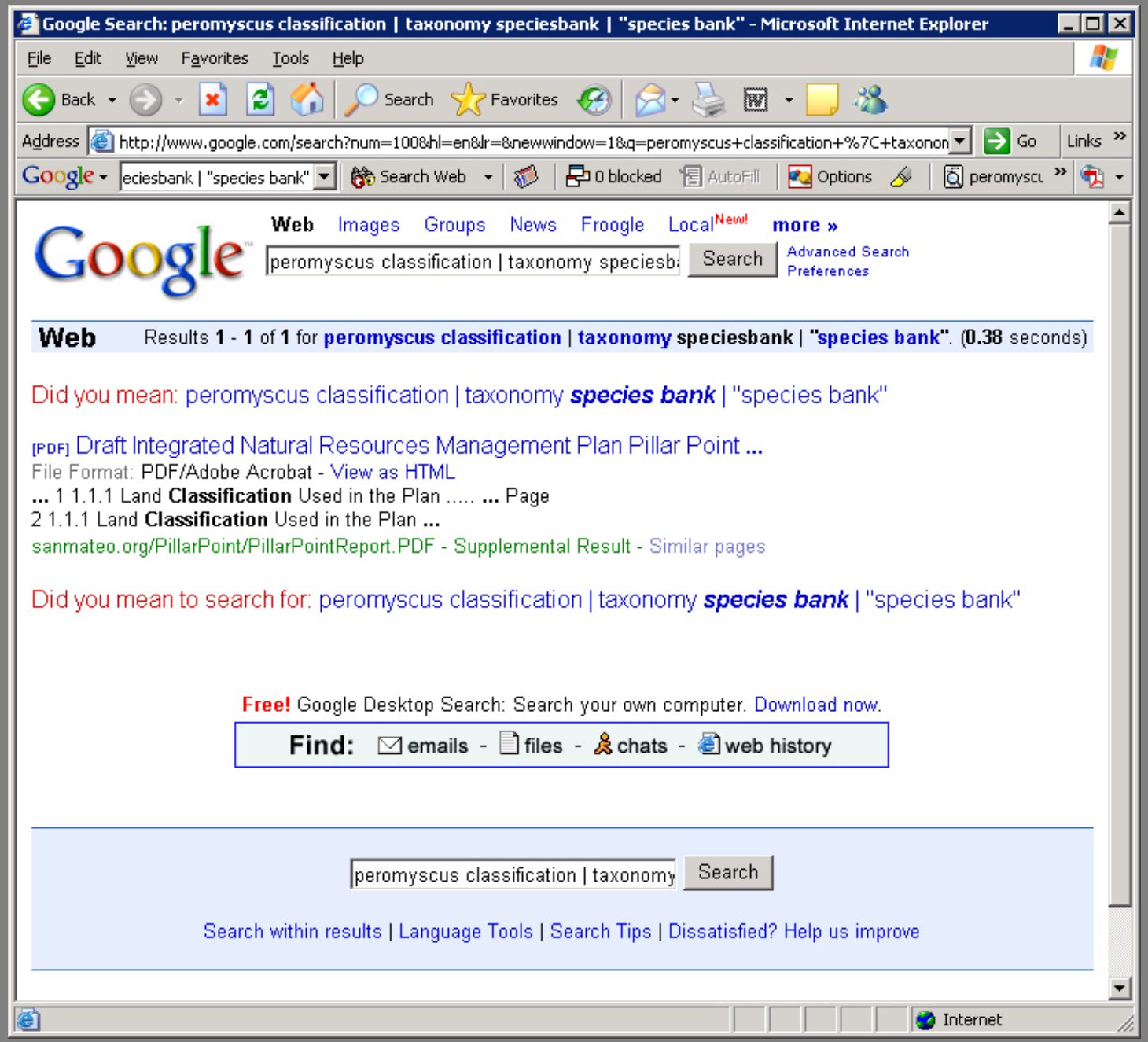
An Example

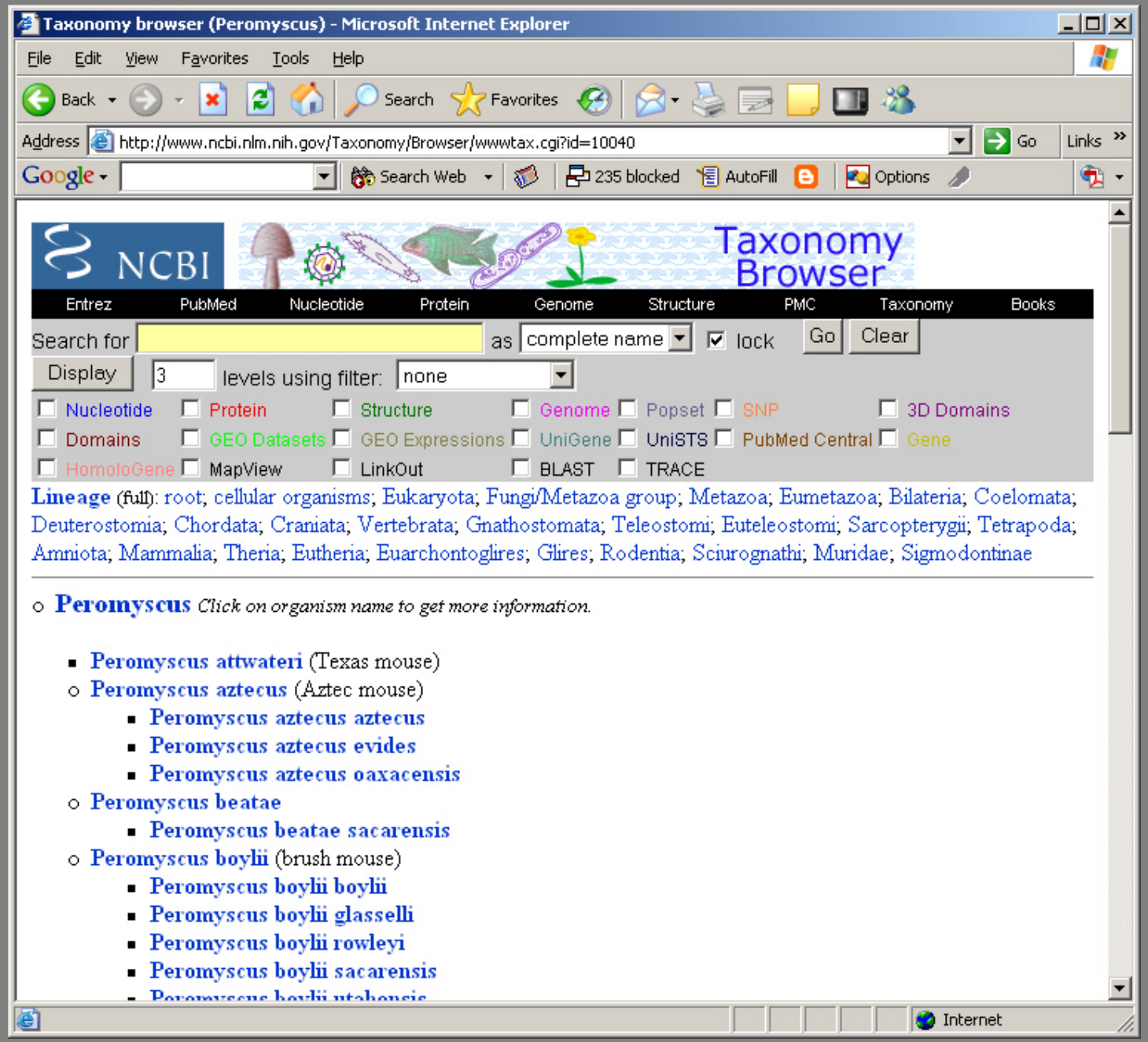
Peromyscus: example

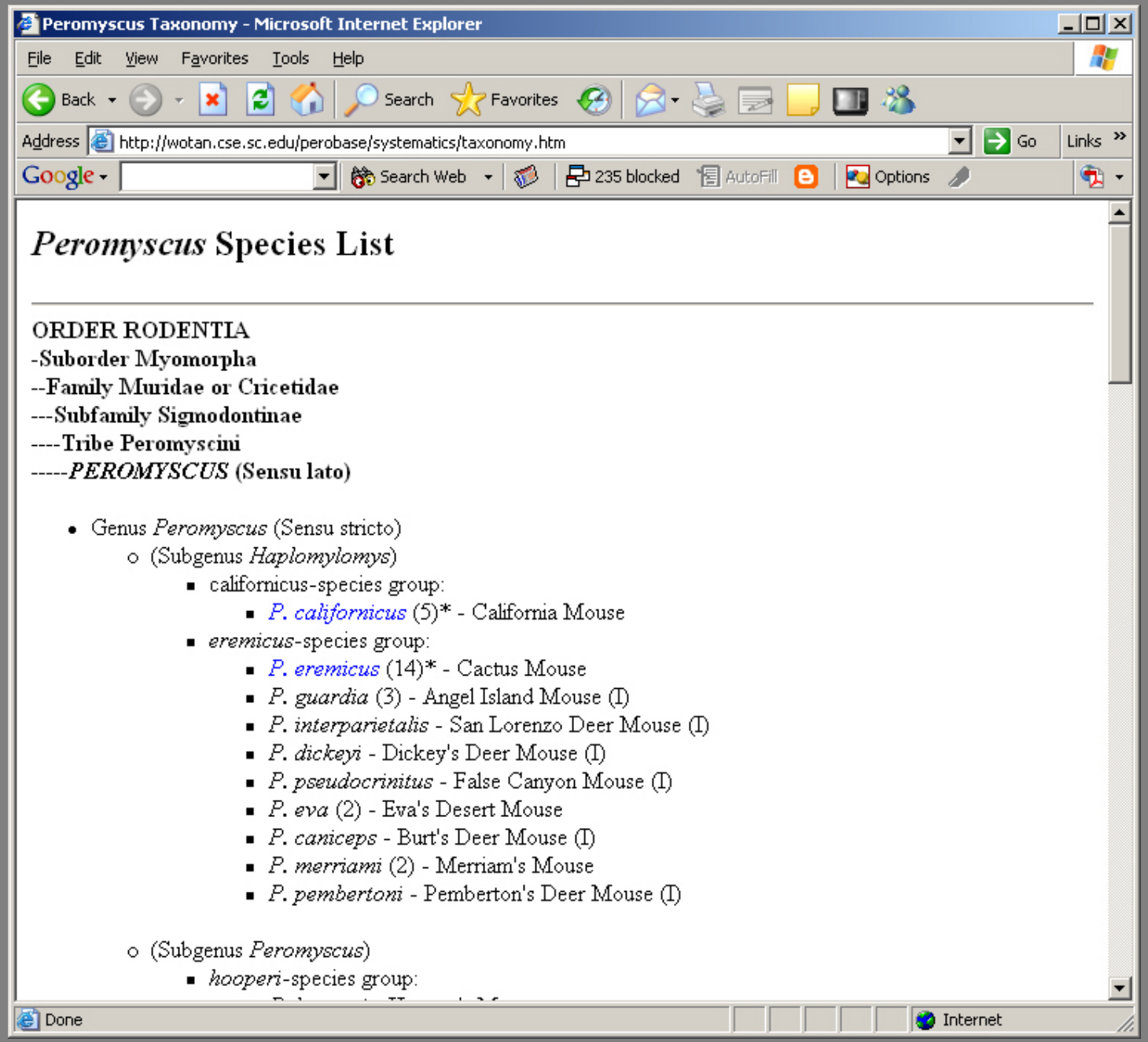
Source: http://cedarcreek.umn.edu/mammals/cricetidae.html

Peromyscus maniculatus


Google Search: peromyscus classification | taxonomy - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.google.com/search?num=100&hl=en&lr=&newwindow=1&q=peromyscus+classification+%7C+taxonon
Go Links
Google s classification | taxonomy Search Web 0 blocked AutoFill Options peromyscu
Google
Web Images Groups News Froogle Local New! more »
peromyscus classification | taxonomy
Search
Advanced Search
Preferences
Web
Results 1 - 100 of about 13,900 for peromyscus classification | taxonomy. (0.35 seconds)
Tip: Save time by hitting the return key instead of clicking on "search"
Peromyscus Taxonomy
Peromyscus Species List. ORDER RODENTIA ... Sensu lato): Genus Peromyscus (Sensu stricto): (Subgenus Haplomylomys): californicus-species group: ...
wotan.cse.sc.edu/perobase/systematics/taxonomy.htm - 9k -
Cached - Similar pages
Sponsored Links
Taxonomy Library
Comprehensive Taxonomy Library
Any Subject, Any Format, Instantly
www.intellisophic.com
p_attwateri 2nd draft
... CLASSIFICATION Order Rodentia, Suborder Myomorpha, Family Muridae ... of the brush mouse (Peromyscus boylii) in ... Geographic variation and taxonomy of Peromyscus ...
wotan.cse.sc.edu/perobase/systematics/p_attwat.htm - 15k - Cached - Similar pages
[More results from wotan.cse.sc.edu]
ADW: Peromyscus: Classification
... Peromyscus. Genus Peromyscus (deer mice and white-footed mice). ... Genus Peromyscus (deer mice and white-footed mice). information. Species ...
animaldiversity.ummz.umich.edu/ site/accounts/information/Peromyscus.html - 78k -
Cached - Similar pages
ADW: Peromyscus maniculatus: Classification
... Peromyscus maniculatus. Peromyscus maniculatus (deer mouse). Information; Pictures; Specimens; Classification. What do these icons mean? The ...
animaldiversity.ummz.umich.edu/site/ accounts/classification/Peromyscus_maniculatus.html
- 18k - Cached - Similar pages
[More results from animaldiversity.ummz.umich.edu]
Internet


Google Search: peromyscus classification | taxonomy speciesbank | "species bank" - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.google.com/search?num=100&hl=en&lr=&newwindow=1&q=peromyscus+classification+%7C+taxonon
Go Links
Google eciesbank | "species bank" Search Web 0 blocked AutoFill Options peromyscu
Google
Web Images Groups News Froogle Local New! more »
peromyscus classification | taxonomy speciesb
Search
Advanced Search
Preferences
Web Results 1 - 1 of 1 for peromyscus classification | taxonomy speciesbank | "species bank". (0.38 seconds)
Did you mean: peromyscus classification | taxonomy species bank | "species bank"
[PDF] Draft Integrated Natural Resources Management Plan Pillar Point ...
File Format: PDF/Adobe Acrobat - View as HTML
... 1 1.1.1 Land Classification Used in the Plan .... Page
2 1.1.1 Land Classification Used in the Plan ...
sanmateo.org/PillarPoint/PillarPointReport.PDF - Supplemental Result - Similar pages
Did you mean to search for: peromyscus classification | taxonomy species bank | "species bank"
Free! Google Desktop Search: Search your own computer. Download now.
Find: emails - files - chats - web history
peromyscus classification | taxonomy
Search
Search within results | Language Tools | Search Tips | Dissatisfied? Help us improve
Internet

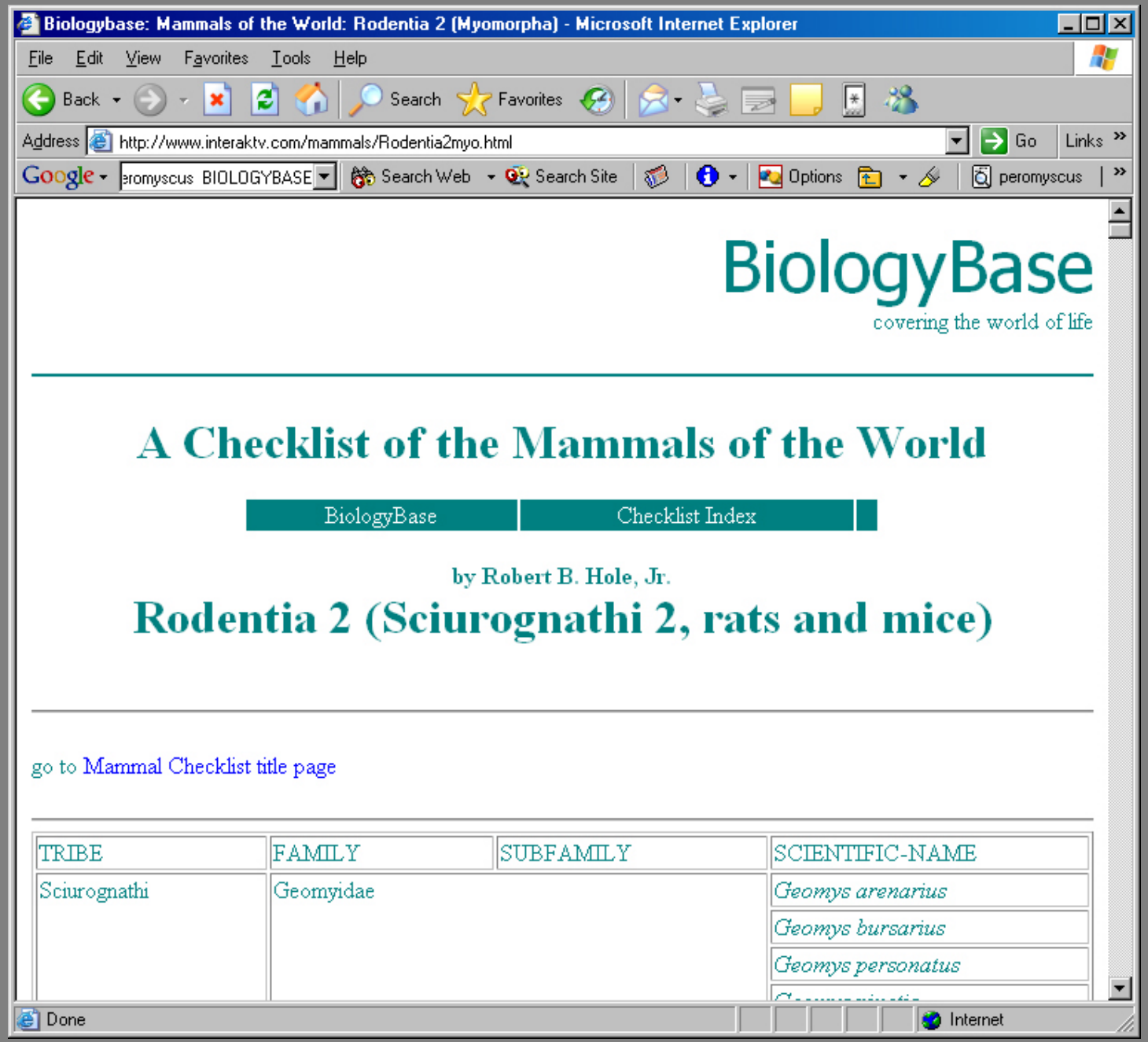

Taxonomy browser (Peromyscus) - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=10040
Go Links
Google Search Web 235 blocked AutoFill Options
NCBI
Taxonomy Browser
Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Books
Search for as complete name lock Go Clear
Display 3 levels using filter: none
Nucleotide Protein Structure Genome Popset SNP 3D Domains
Domains GEO Datasets GEO Expressions UniGene UniSTS PubMed Central Gene
HomoloGene MapView LinkOut BLAST TRACE
Lineage (full): root; cellular organisms; Eukaryota; Fungi/Metazoa group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Sigmodontinae
Peromyscus Click on organism name to get more information.
Peromyscus attwateri (Texas mouse)
Peromyscus aztecus (Aztec mouse)
Peromyscus aztecus aztecus
Peromyscus aztecus evides
Peromyscus aztecus oaxacensis
Peromyscus beatae
Peromyscus beatae sacarensis
Peromyscus boylii (brush mouse)
Peromyscus boylii boylii
Peromyscus boylii glasselli
Peromyscus boylii rowleyi
Peromyscus boylii sacarensis
Internet



Peromyscus Species List

ORDER RODENTIA
-Suborder Myomorpha
--Family Muridae or Cricetidae
---Subfamily Sigmodontinae
----Tribe Peromyscini
-----_PEROMYSCUS_ (Sensu lato)

- Genus *Peromyscus* (Sensu stricto)
  - (Subgenus *Haplomylomys*)
    - californicus-species group:
      - *P. californicus* (5)* - California Mouse
    - *eremicus*-species group:
      - *P. eremicus* (14)* - Cactus Mouse
      - *P. guardia* (3) - Angel Island Mouse (I)
      - *P. interparietalis* - San Lorenzo Deer Mouse (I)
      - *P. dickeyi* - Dickey's Deer Mouse (I)
      - *P. pseudocrinitus* - False Canyon Mouse (I)
      - *P. eva* (2) - Eva's Desert Mouse
      - *P. caniceps* - Burt's Deer Mouse (I)
      - *P. merriami* (2) - Merriam's Mouse
      - *P. pembertoni* - Pemberton's Deer Mouse (I)
  - (Subgenus *Peromyscus*)
    - *hooperi*-species group:


Biologybase: Mammals of the World: Rodentia 2 (Myomorpha) - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.interaktv.com/mammals/Rodentia2myo.html
Go Links
Google eromyscus BIOLOGYBASE Search Web Search Site Options peromyscus
BiologyBase
covering the world of life
A Checklist of the Mammals of the World
BiologyBase
Checklist Index
by Robert B. Hole, Jr.
Rodentia 2 (Sciurognathi 2, rats and mice)
go to Mammal Checklist title page
TRIBE
FAMILY
SUBFAMILY
SCIENTIFIC-NAME
Sciurognathi
Geomyidae
Geomys arenarius
Geomys bursarius
Geomys personatus
Done
Internet

Peromyscus: classification

NCBI:	Muridae
Perobase:	Muridae or Cricetidae
BiologyBASE:	Muridae

Peromyscus: number of species

NCBI:	42
Perobase:	55 + 15 from other genera
BiologyBASE:	53

Peromyscus: number of species

NCBI:	42
Perobase:	55
BiologyBASE:	53

Peromyscus: number of species

NCBI:	42
Perobase:	55
BiologyBASE:	53
Total:	64

Peromyscus: number of species

NCBI:	42
Perobase:	55
BiologyBASE:	53
Total:	64
In common:	32

Peromyscus: number of species

NCBI:	42
Perobase:	55
BiologyBASE:	53
Total:	64
In common:	32

Hmmm. Fifty percent concordance across only three resources.

Not so hot…

Peromyscus: number of subspecies

NCBI: 30

Perobase: 209

P. maniculatus: number of subspecies

NCBI: 1

Perobase: 64

Challenges/Limits

Constant Revision

GOAL: A Caution

In parallel to the molecular database GenBank (but operating on completely different principles), GBIF envisions a future in which all sorts of information about any species (gene sequences, occurrence in ecosystems, specific locality data, ecological relationships, physiological requirements and so on) would be compiled on demand from many, disparate, continuously updated databases.

SpeciesBANK would effectively be an encyclopedia of species that is continuously filling in missing or supplanting outdated information.

GOAL: A Caution

In parallel to the molecular database GenBank (but operating on completely different principles), GBIF envisions a future in which all sorts of information about any species (gene sequences, occurrence in ecosystems, specific locality data, ecological relationships, physiological requirements and so on) would be compiled on demand from many, disparate, continuously updated databases.

SpeciesBANK would effectively be an encyclopedia of species that is continuously filling in missing or supplanting outdated information.

Primary Literature

- Each contribution to the primary literature is an original contribution. It may be based on prior findings, or it may completely overturn prior findings.
- There is NO REQUIREMENT OF CONSISTENCY between any two documents in the primary literature.

Encyclopedia of Science

Should a biological database be a compilation of scientific ***truths***, or should it be a collection of scientific observations? A compilation of facts is appealing, so that one might consult the database to determine ***the*** amino-acid sequence of human beta-hemoglobin, or ***the*** map location of the beta-hemoglobin gene. But scientific facts have a way of changing with more scientific observations, and the growing burden of constant editing to achieve accuracy and internal consistency would be difficult. Ziman (1978) has made a relevant observation, although not in the context of database publishing.

Encyclopedia of Science

Science continually evolves. Scientific knowledge is under constant revision in the light of new evidence. From a practical point of view, it is not the ultimate truth of the scientific world picture that matters, but the [current] scientific answers to particular questions...

The concept of an archive of reliable scientific knowledge is much too schematic. There is no *Encyclopaedia* where *all* well-established science, and only well-established science, may be consulted. If such an institution existed, it would be in constant agitation, as new information was being added, and old facts and assertions struck out.

Ziman, J. 1978. *Reliable Knowledge: An Exploration of the Grounds for Belief in Science*. London: Cambridge University press.

Encyclopedia of Science

Building a database of scientific truths is equivalent to creating an electronic version of Ziman's Encyclopaedia of all well-established science. Maintaining perfect consistency would require that every existing entry in the database be checked for continuing validity every time any new entry is made. Even with a linear flow of new data, this seems impossible. Also, assertions about the real world may be initially believed, then rejected, then accepted again, albeit in a modified form. Catastrophism in geology is an example. Thus, maintaining a database of scientific truth would be an editorial nightmare, if not an outright impossibility.

Encyclopedia of Science

Building a database of scientific truths is equivalent to creating an electronic version of Ziman's Encyclopaedia of all well-established science. Maintaining perfect consistency would require that every existing entry in the database be checked for continuing validity every time any new entry is made. Even with a linear flow of new data, this seems impossible. Also, assertions about the real world may be initially believed, then rejected, then accepted again, albeit in a modified form. Catastrophism in geology is an example. Thus, maintaining a database of scientific truth would be an editorial nightmare, if not an outright impossibility.

Constant Revision

An Example

The Perils of Constant Revision

St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph taken in 1897

The Perils of Constant Revision

St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph taken in 1897

The Perils of Constant Revision

Idealistic young men, whose efforts ultimately had some very practical consequences.

In the spirit of “one for all and all for one” they worked together, but …

St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph taken in 1897

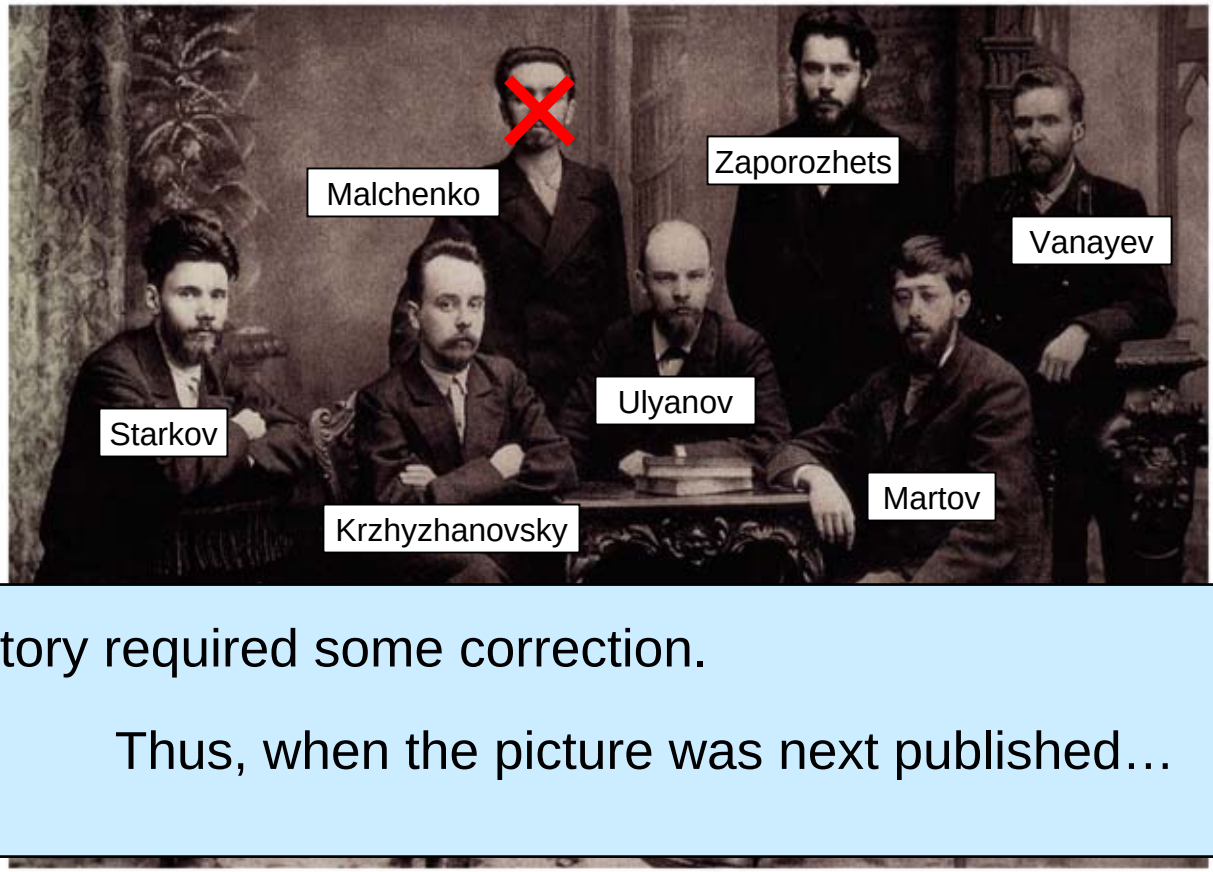
The Perils of Constant Revision

In 1929, Malchenko was arrested and accused of being a "wrecker". He was executed 18 November 1930.

As a counter-revolutionary wrecker of the party, he could hardly have been a participant in its early creation, so…

St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph taken in 1897

The Perils of Constant Revision



History required some correction.

Thus, when the picture was next published…

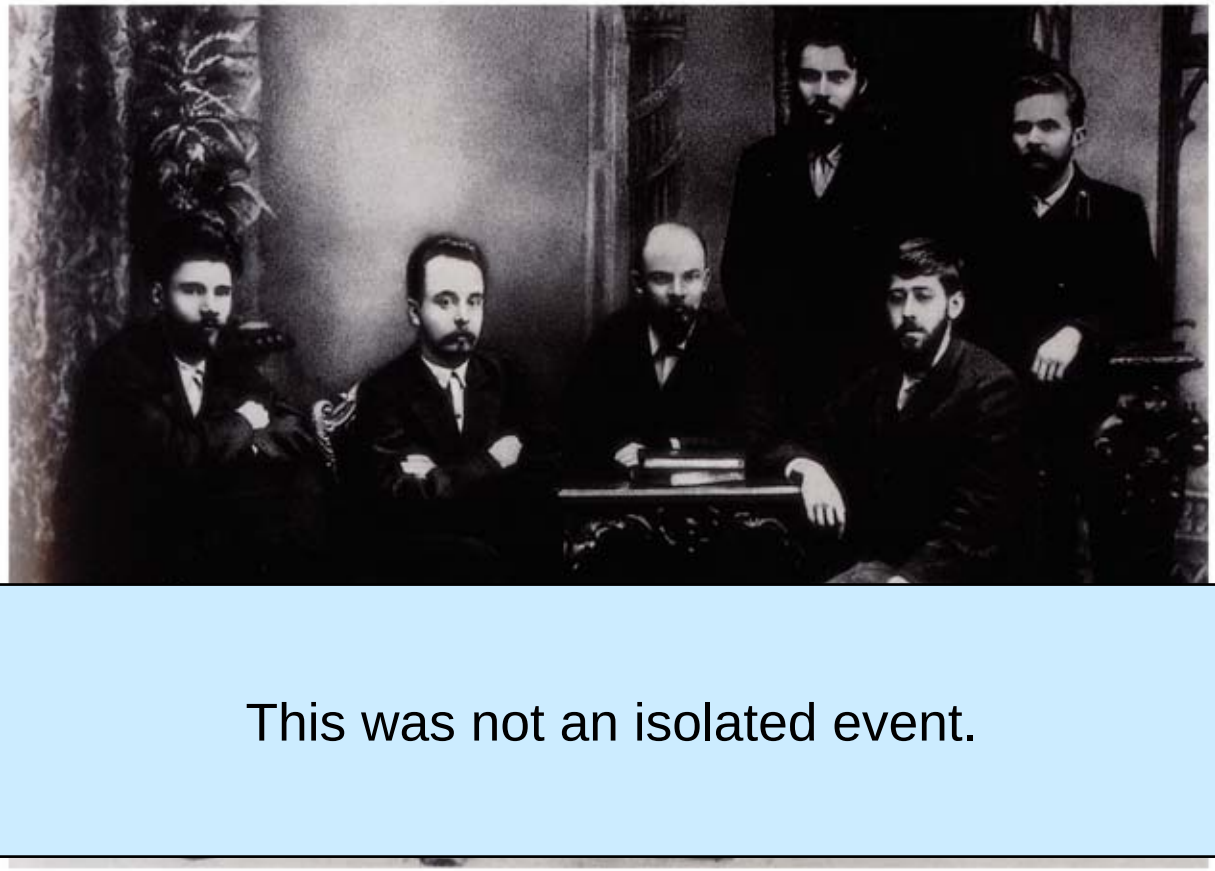
St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph taken in 1897

The Perils of Constant Revision

Malchenko was gone.

St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph published in 1939

The Perils of Constant Revision



This was not an isolated event.

St. Petersburg Union of Struggle for the Liberation of the Working Class
Photograph published in 1939

The Perils of Constant Revision

Stalin, with comrades

The Perils of Constant Revision

Stalin, with fewer comrades

The Perils of Constant Revision

Photograph from 1934 Russian edition of *Ten Years of Uzbekistan*

The Perils of Constant Revision

Photograph from 1935 Uzbek edition of *Ten Years of Uzbekistan*

The Perils of Constant Revision

Ten Comrades at the 14th Party Congress in 1925

The Perils of Constant Revision

In 1939 there were four.

The Perils of Constant Revision

Four

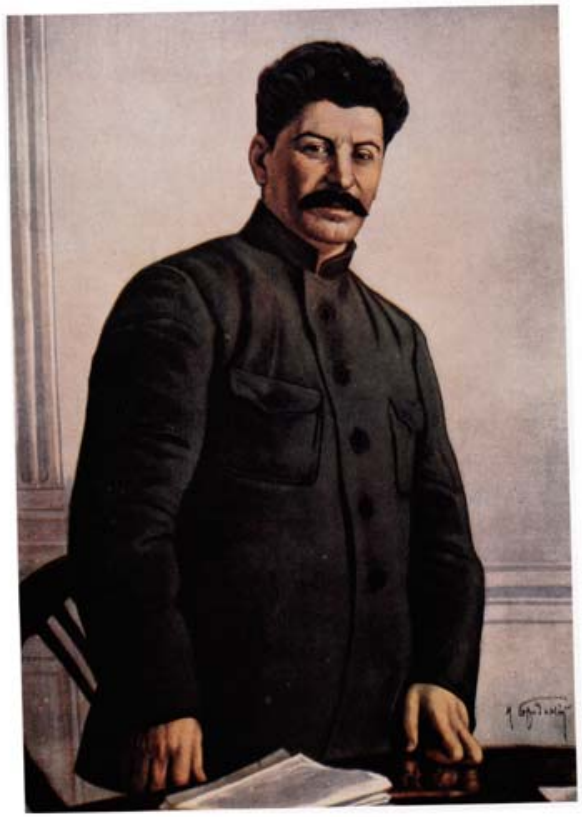
The Perils of Constant Revision

Three

The Perils of Constant Revision

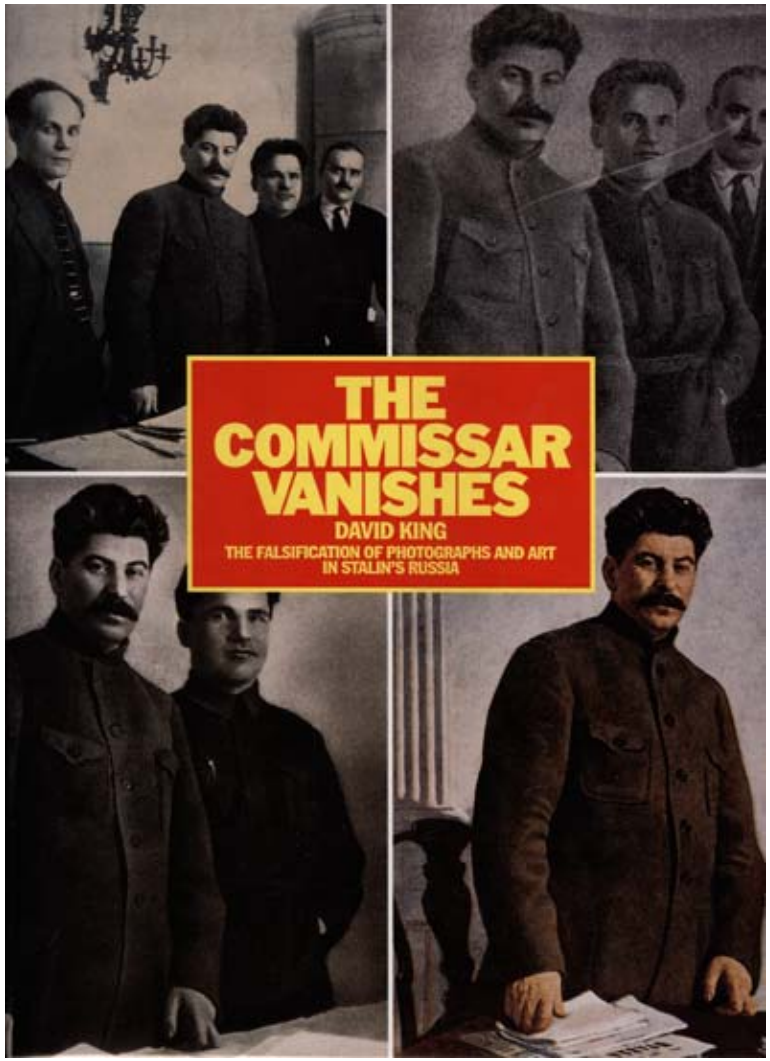
Two

The Perils of Constant Revision



One

SpeciesBANK



Challenges/Limits

Standards

Standards

- Using standards always seems like a good idea, but

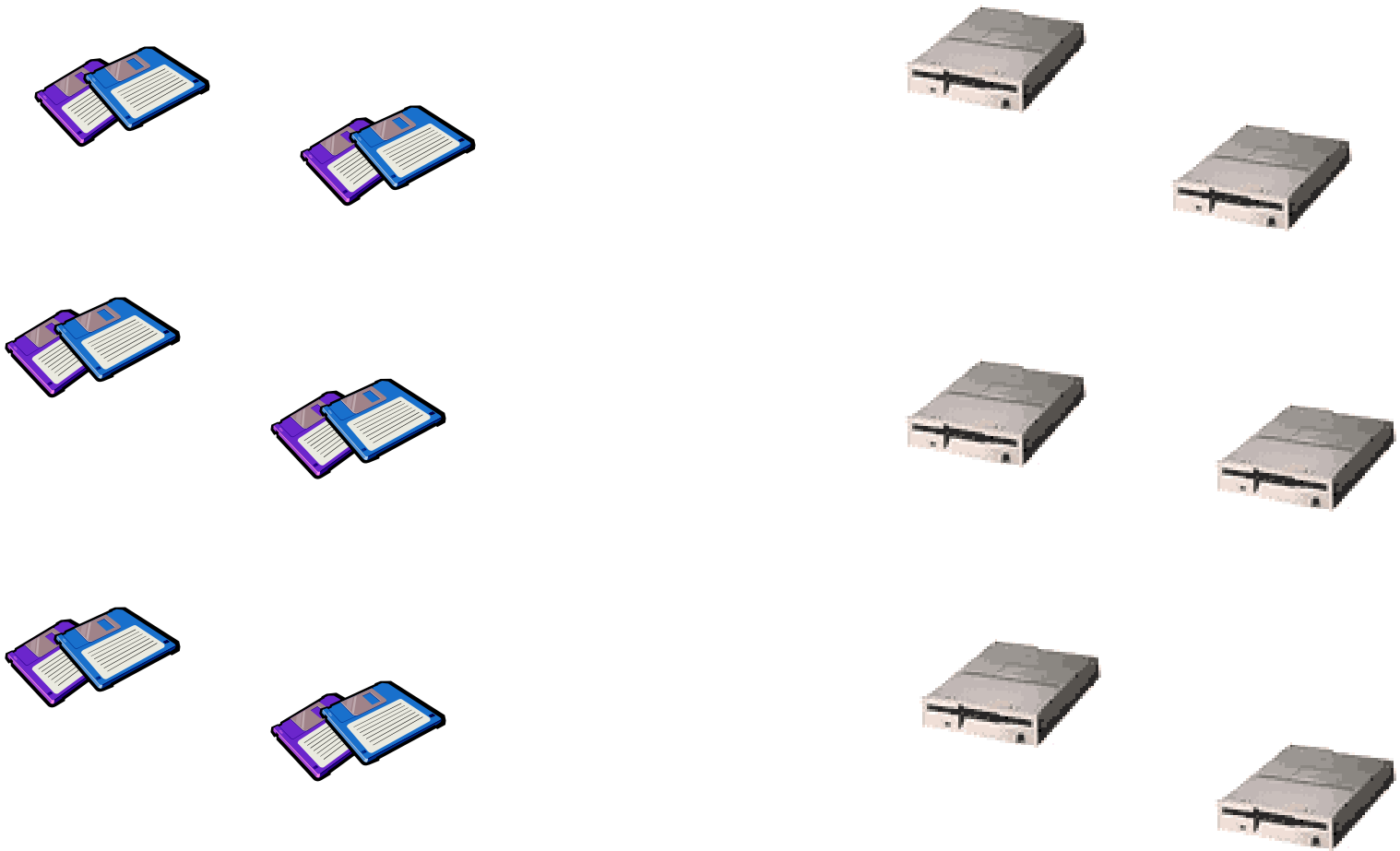
Standards

- Using standards always seems like a good idea, but
- avoiding premature standards is important, and

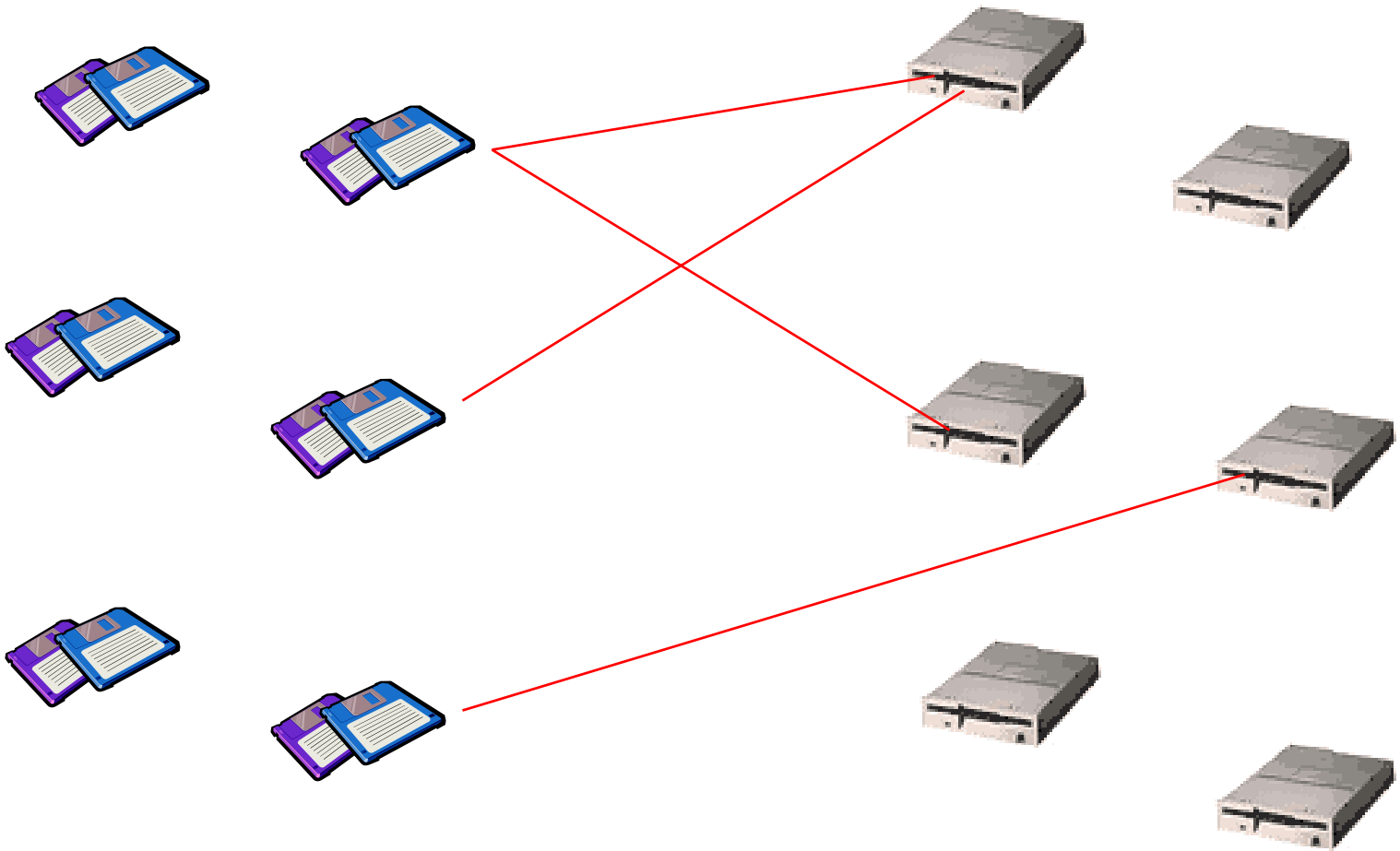
Standards

- Using standards always seems like a good idea, but
- avoiding premature standards is important, and
- adopting bad standards can cripple an IT endeavor, especially one with global ambitions.

Bad Data-exchange Standard

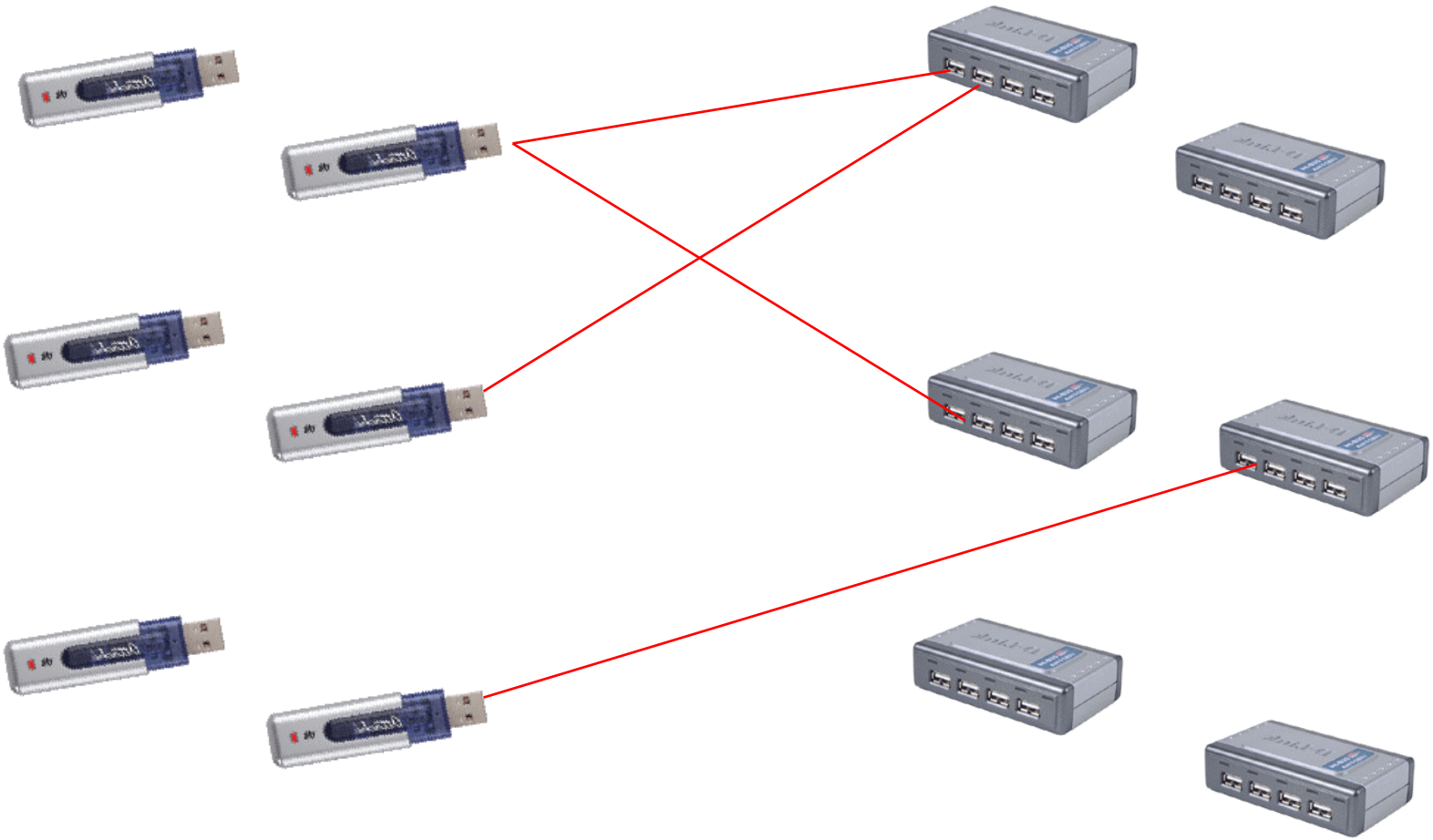


Bad Data-exchange Standard

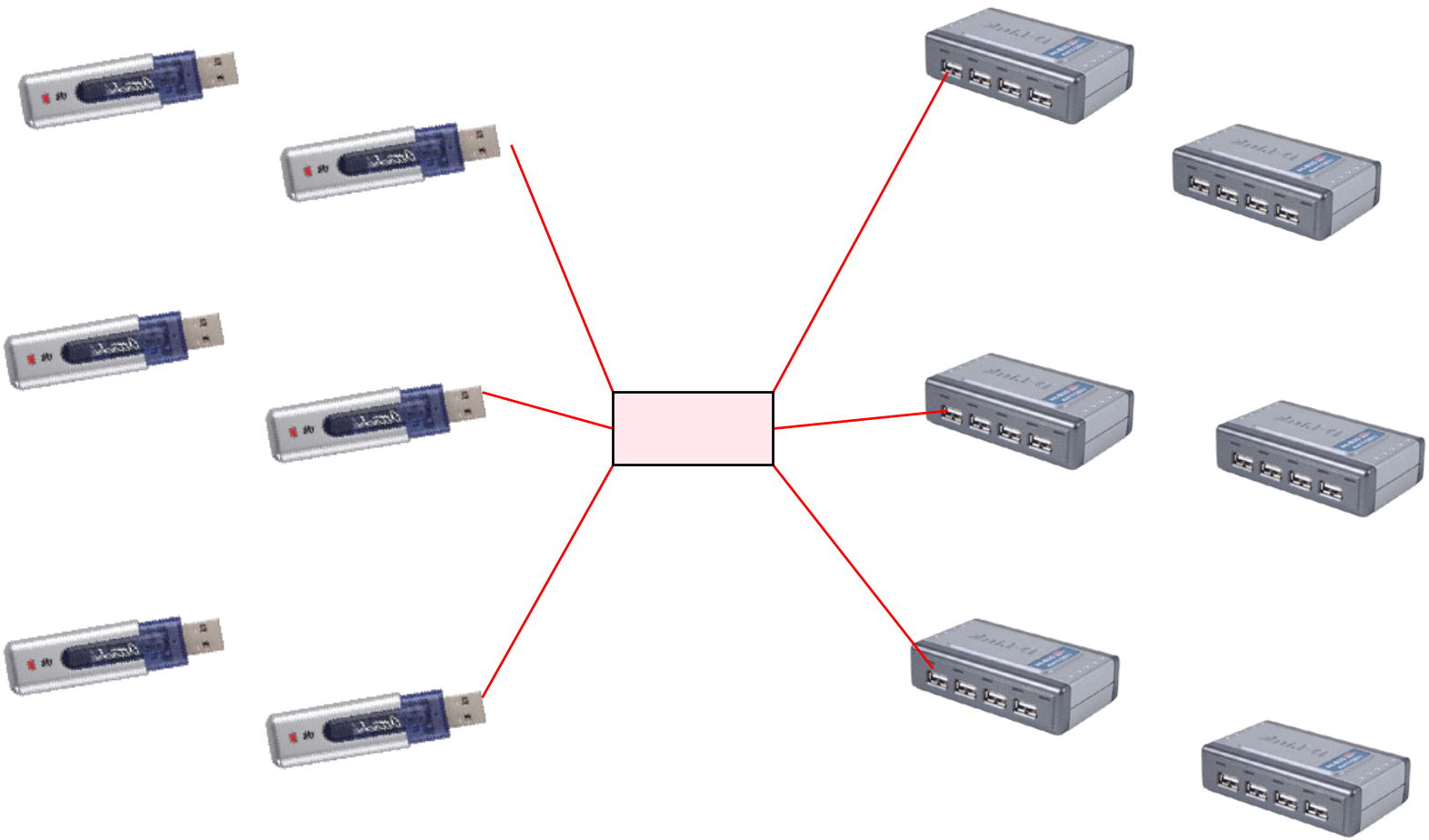


Good Data-exchange Standard

Good Data-exchange Standard

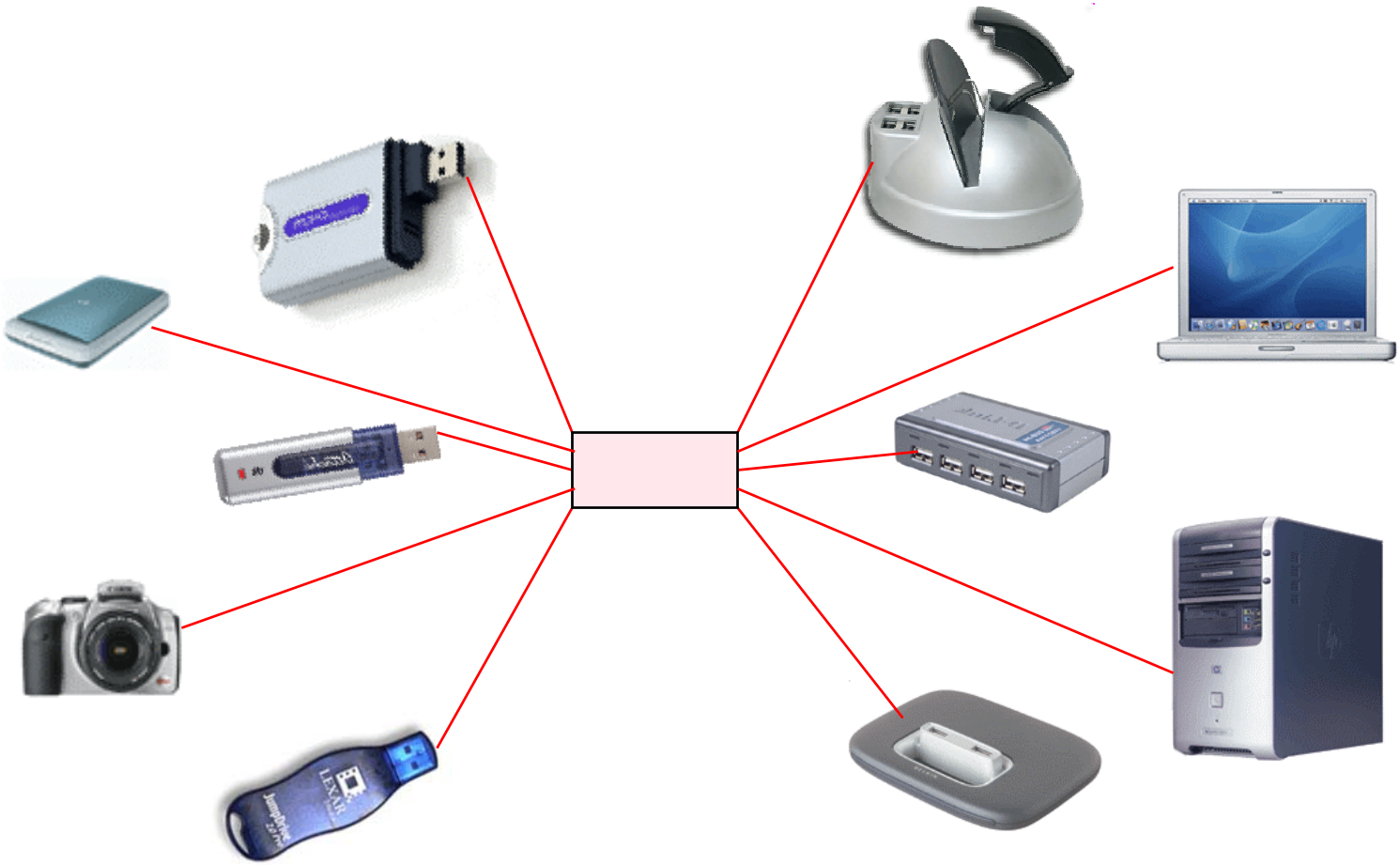


Good Data-exchange Standard



Good Data-exchange Standard

Good Data-exchange Standard



Challenges/Limits

Industry Trends

Industry Trends, I

- The advance of technology is relentless.

Industry Trends, I

- The advance of technology is relentless.
- New technology, new standards, new capabilities are constantly appearing.

Industry Trends, I

- The advance of technology is relentless.
- New technology, new standards, new capabilities are constantly appearing.
- Challenges once thought to be impossible yield to new solutions.

Industry Trends, I

- The advance of technology is relentless.
- New technology, new standards, new capabilities are constantly appearing.
- Challenges once thought to be impossible yield to new solutions.
- Newly developed technologies, like web-services and XML-schema data systems make SpeciesBANK a real possibility.

But always remember,

In fifteen years, today's technology will seem as hopelessly dim and inadequate as 1990s technology seems today.

To build SpeciesBANK, we must USE current technology but we must be careful not to DEPEND on that technology.

Industry Trends, II

- As technology matures ease of use become more and more important.

Industry Trends, II

- As technology matures ease of use become more and more important.
- Real user value occurs when technology is engineered away to invisibility.

SpeciesBANK

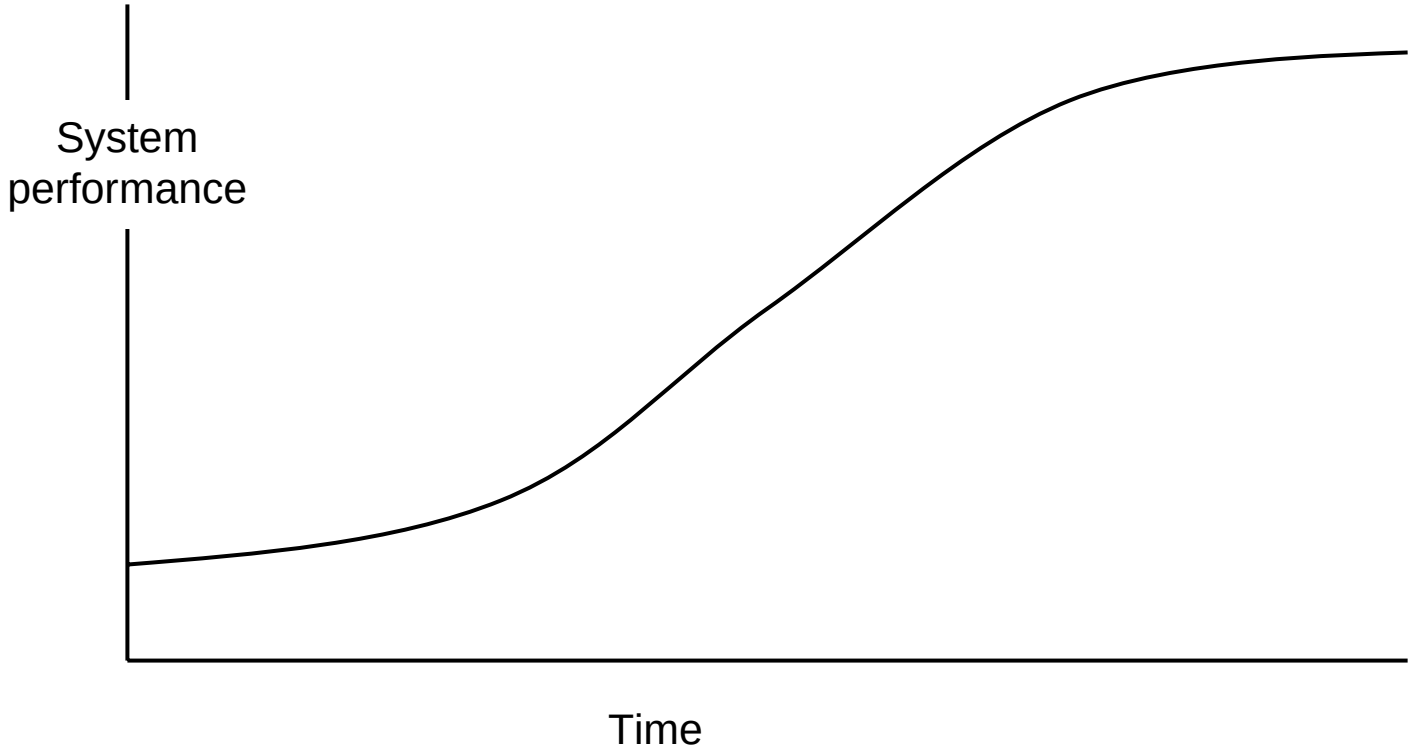
To build truly useful SpeciesBANK systems,

We must appreciate and effectively use advanced technology.

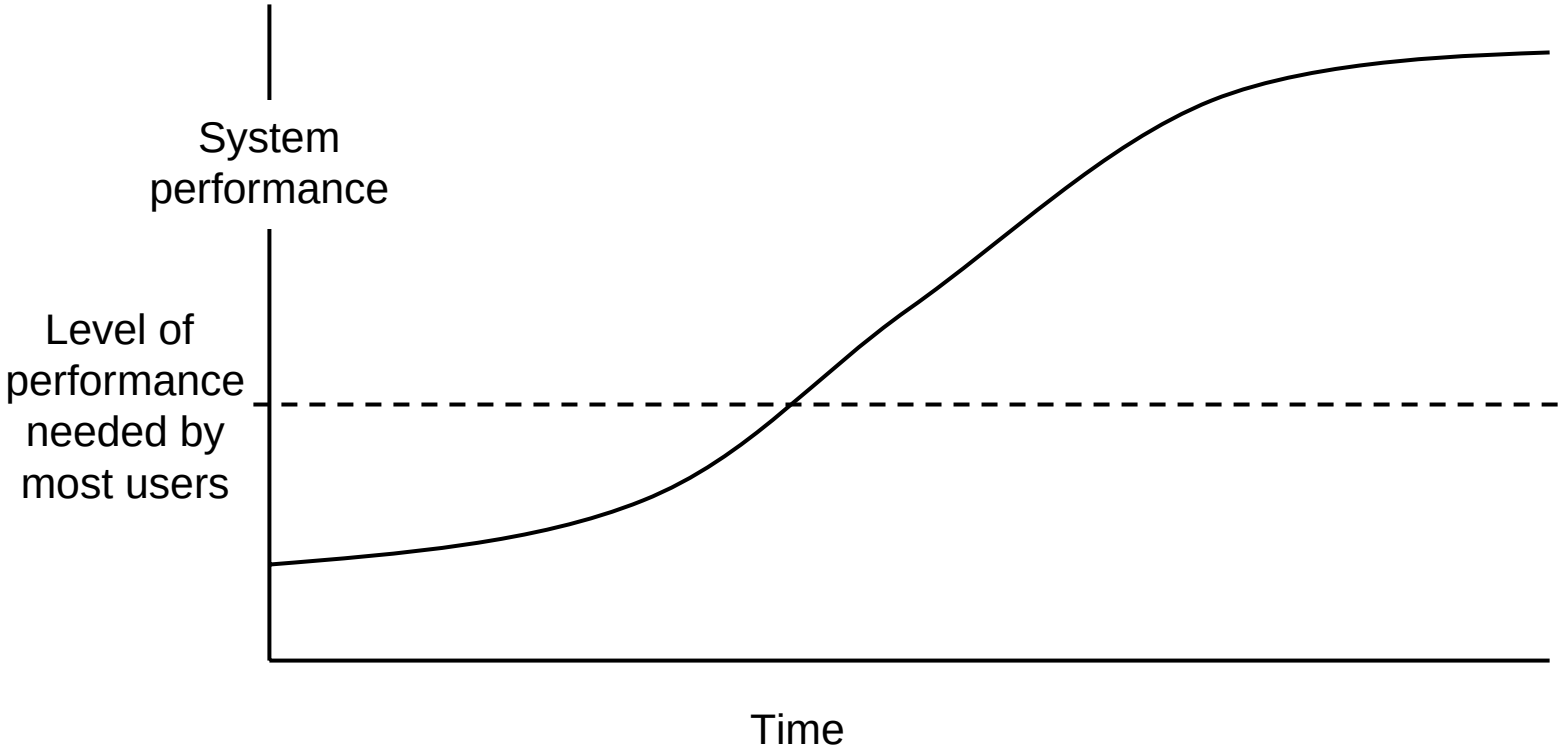
But, we must never allow ourselves to become enamored of that technology.

Our success will depend on our knowledge of the process and practice of science than on our expertise with information technology.

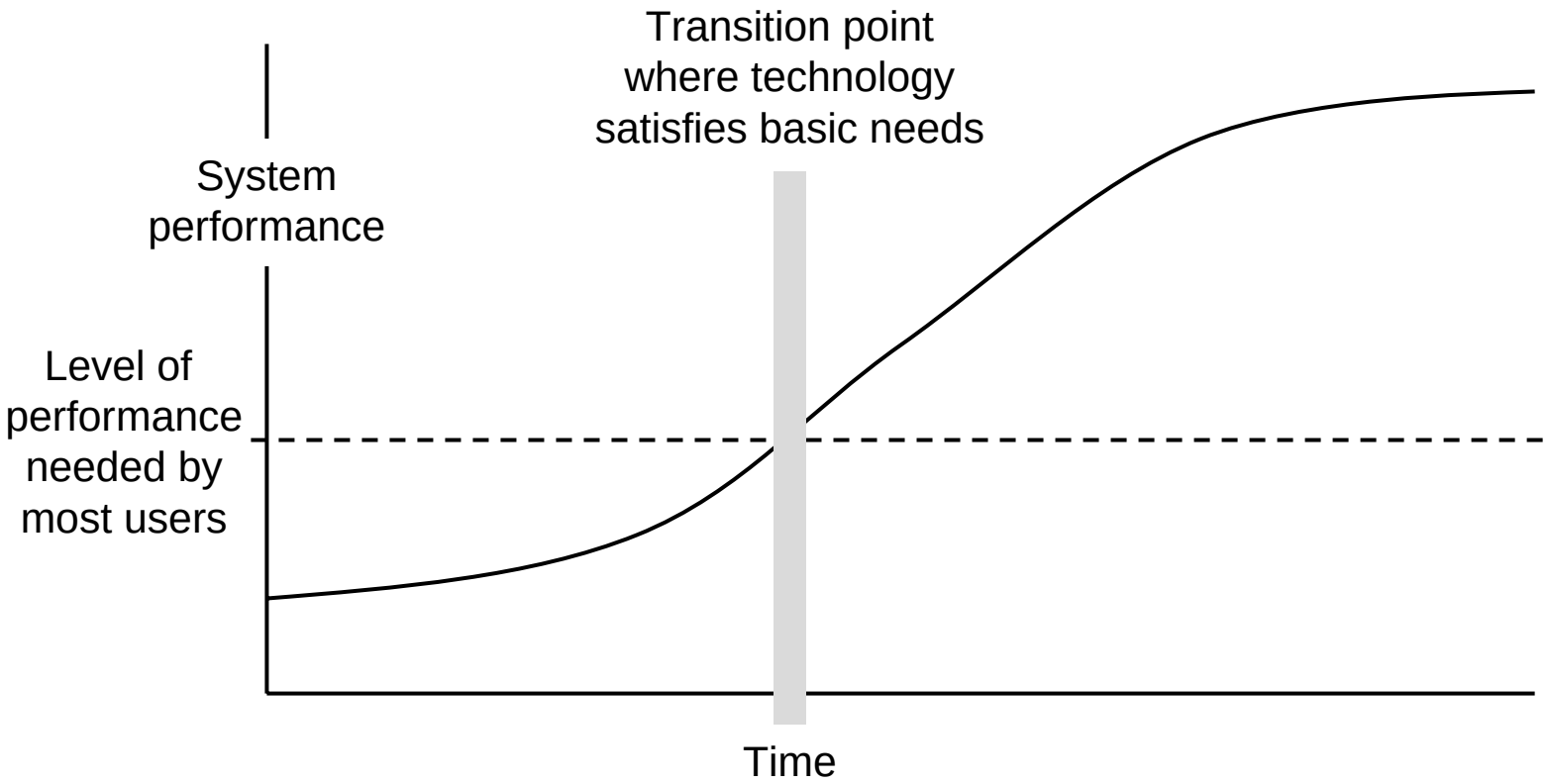
Industry Trends



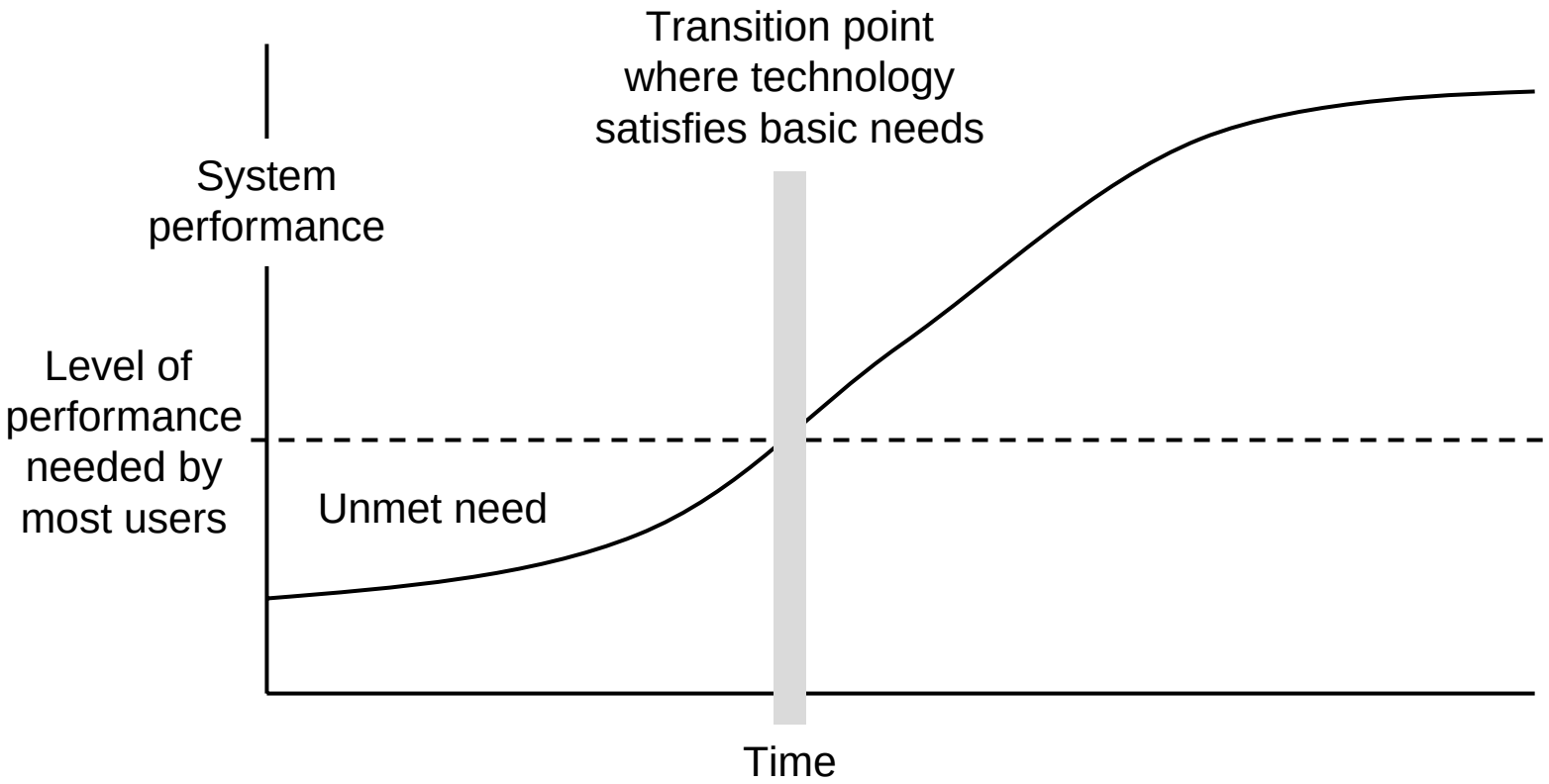
Industry Trends



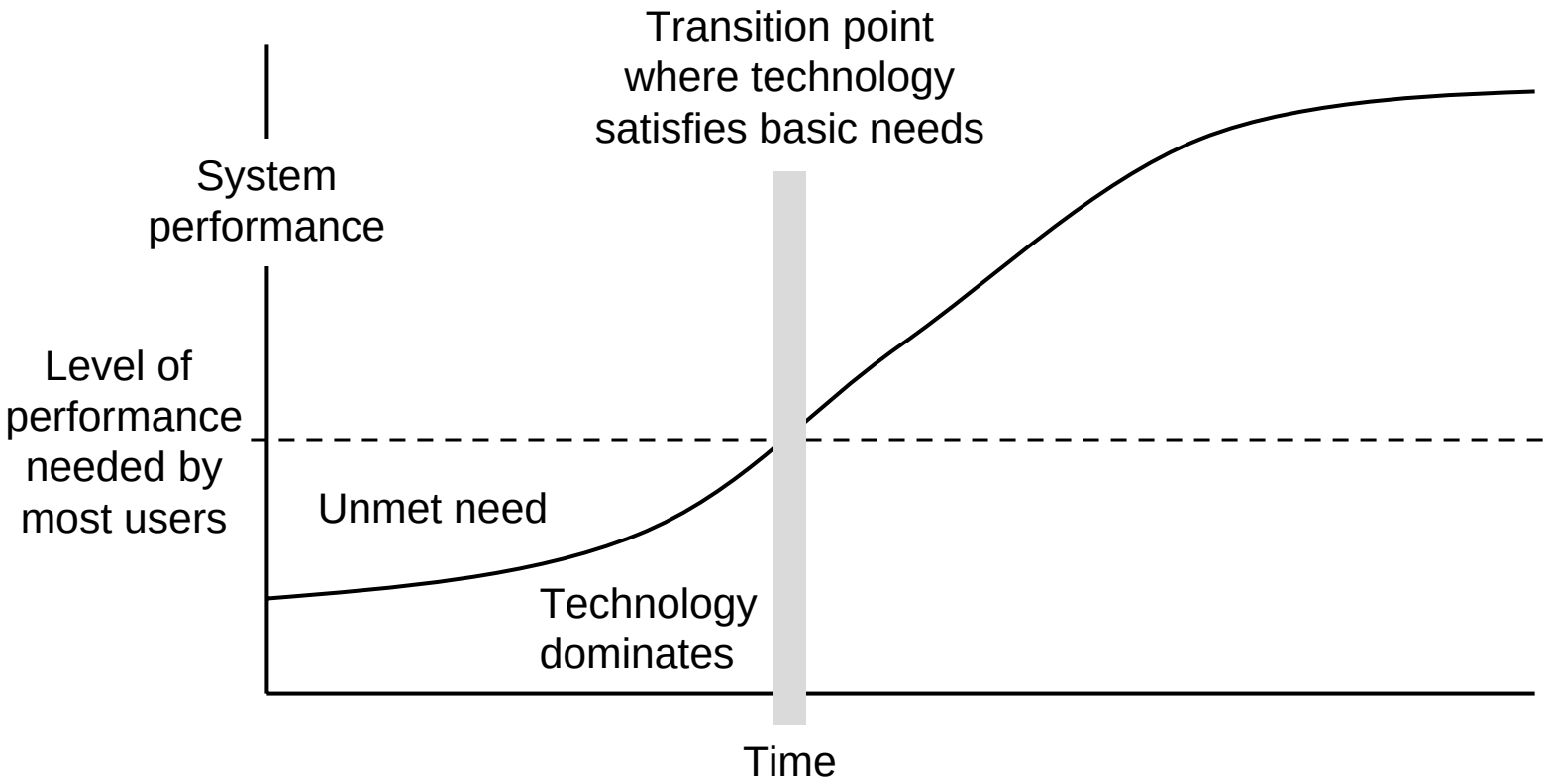
Industry Trends



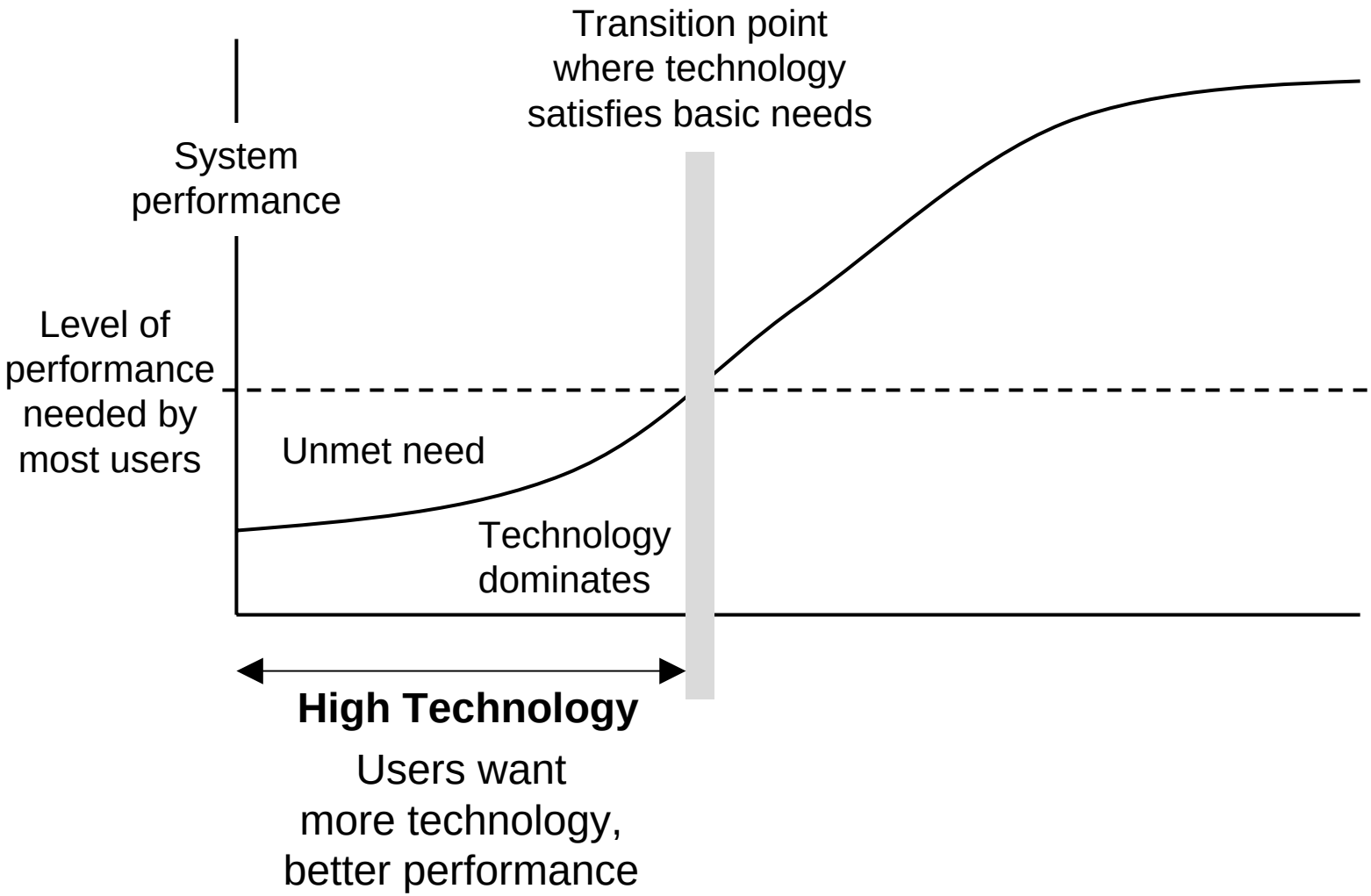
Industry Trends



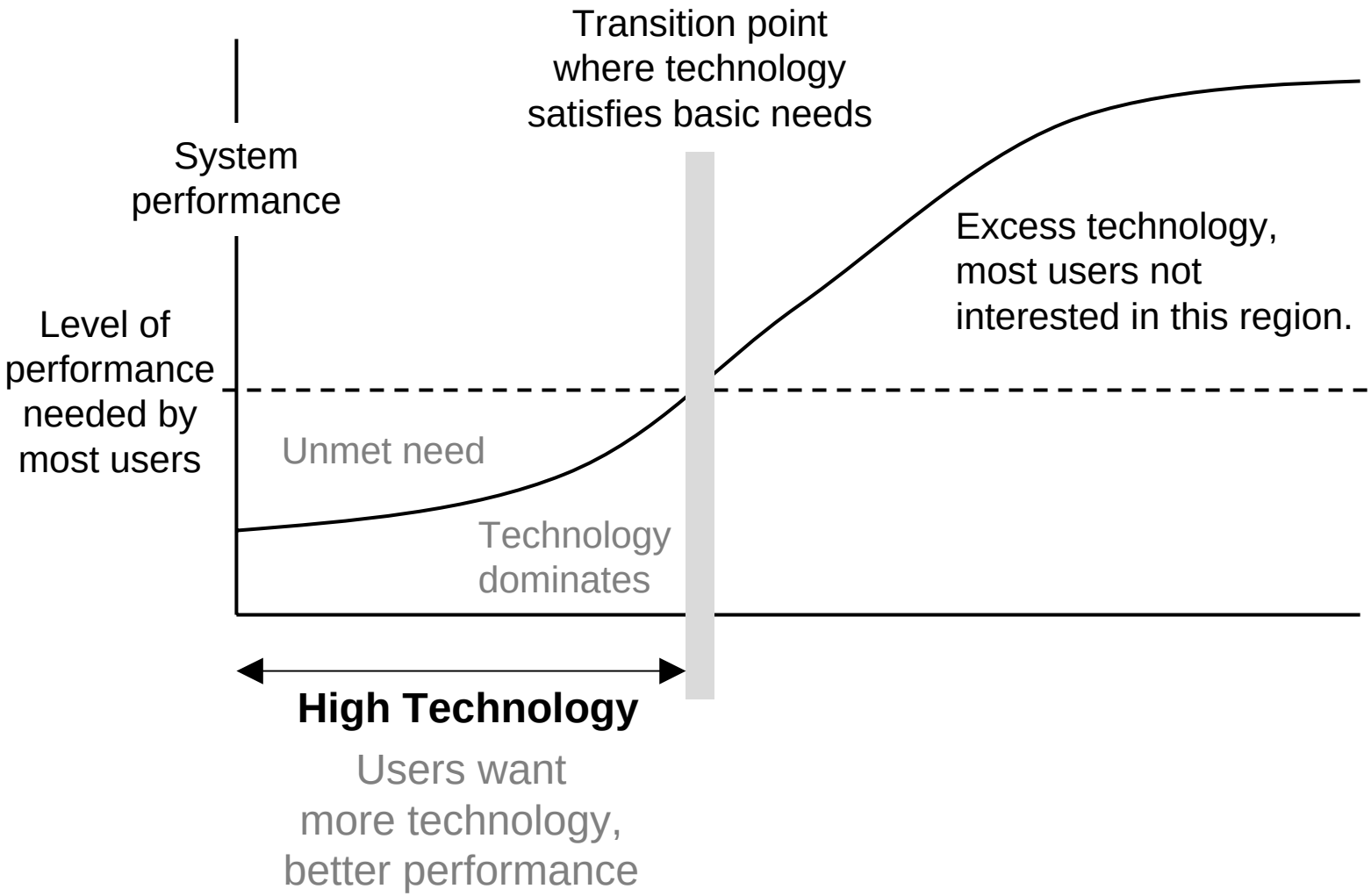
Industry Trends



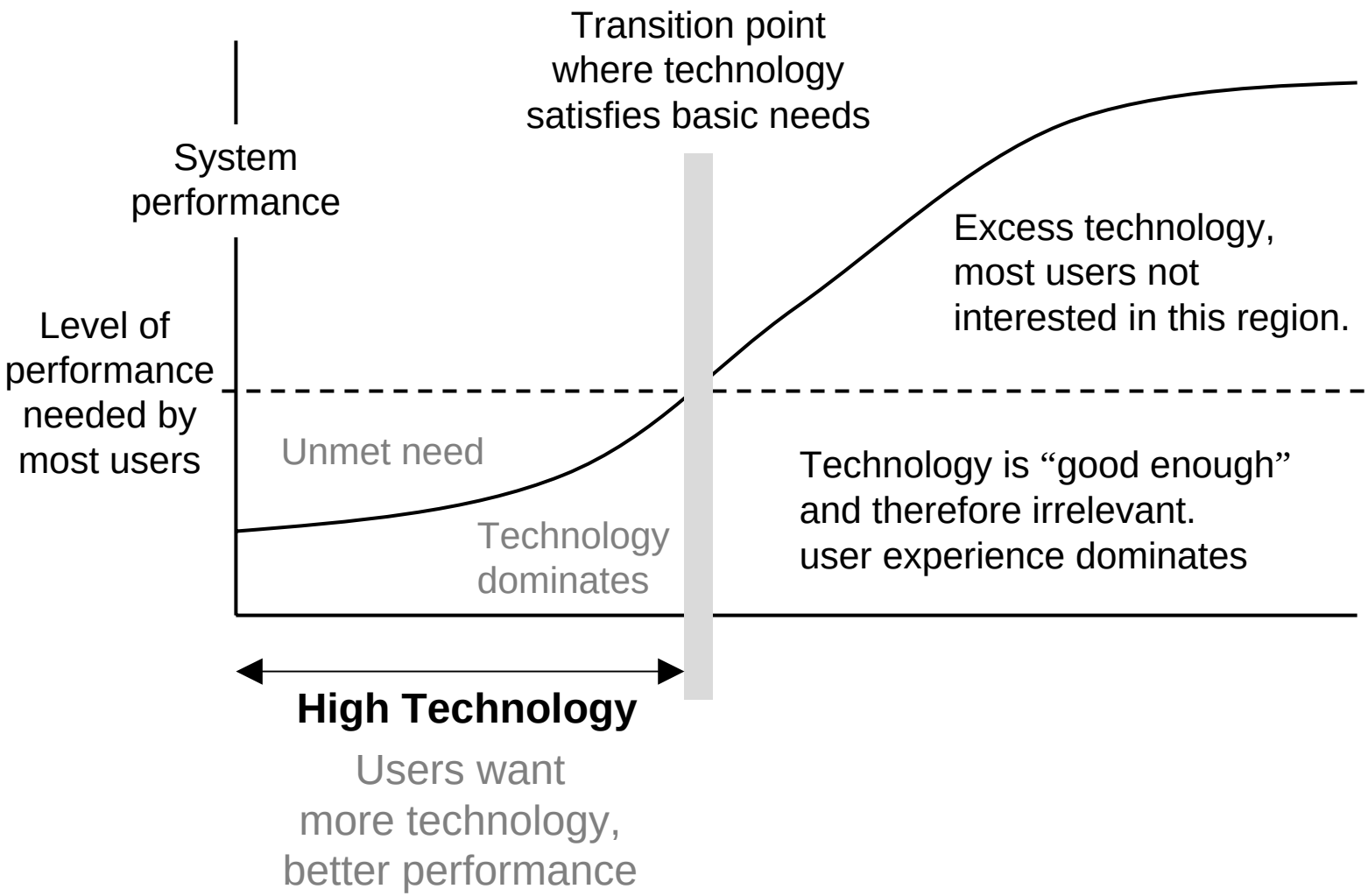
Industry Trends



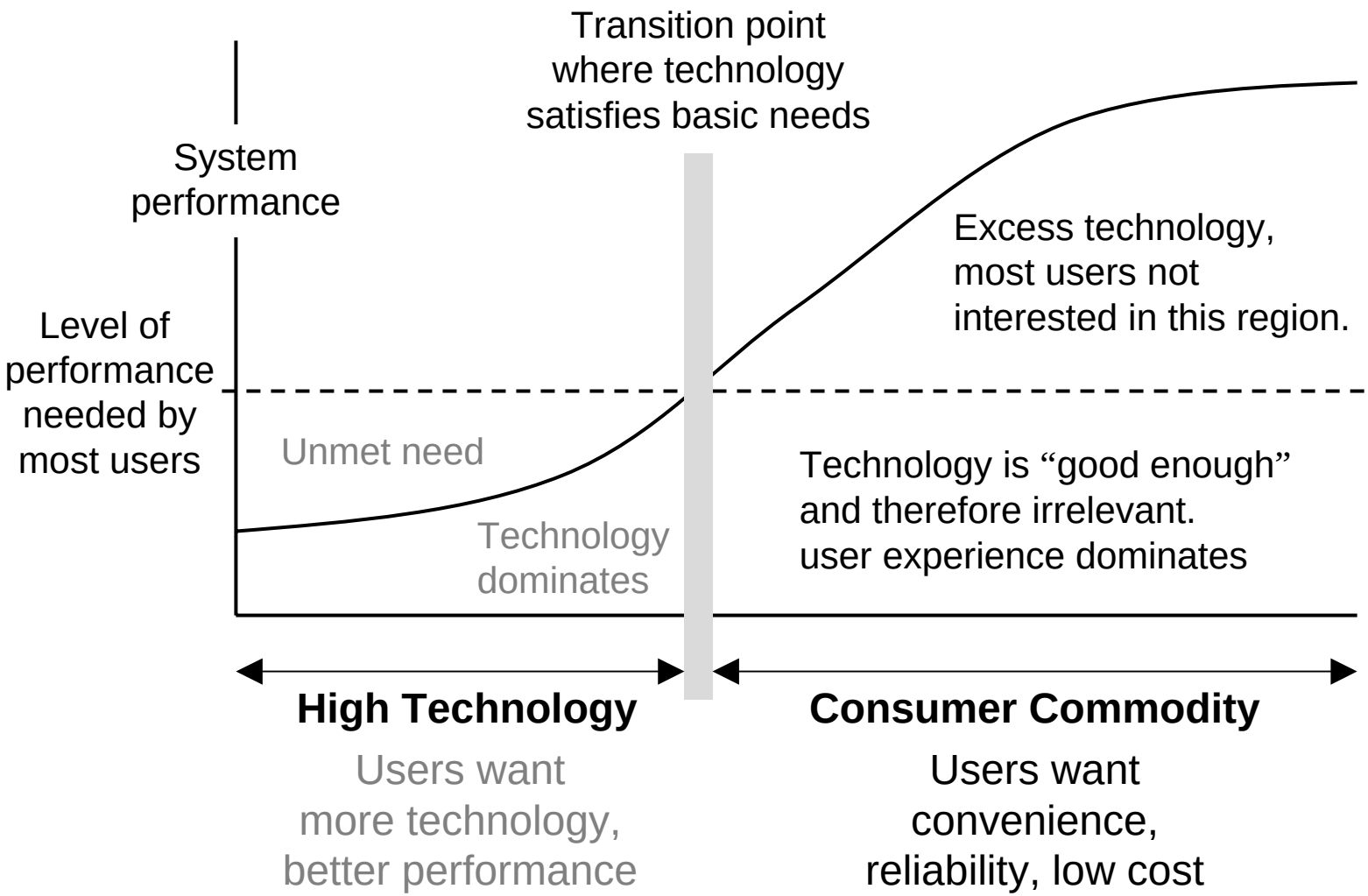
Industry Trends



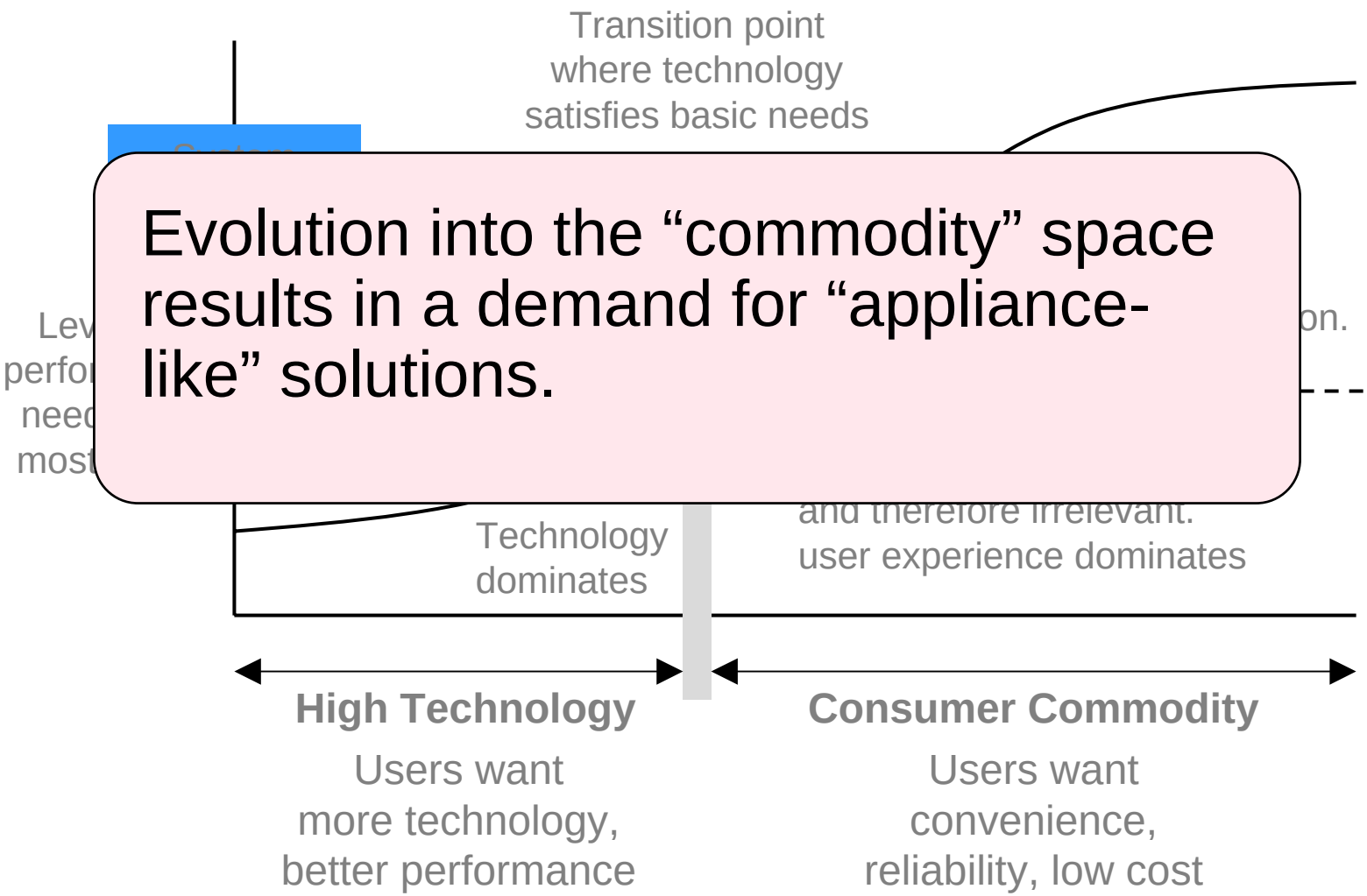
Industry Trends



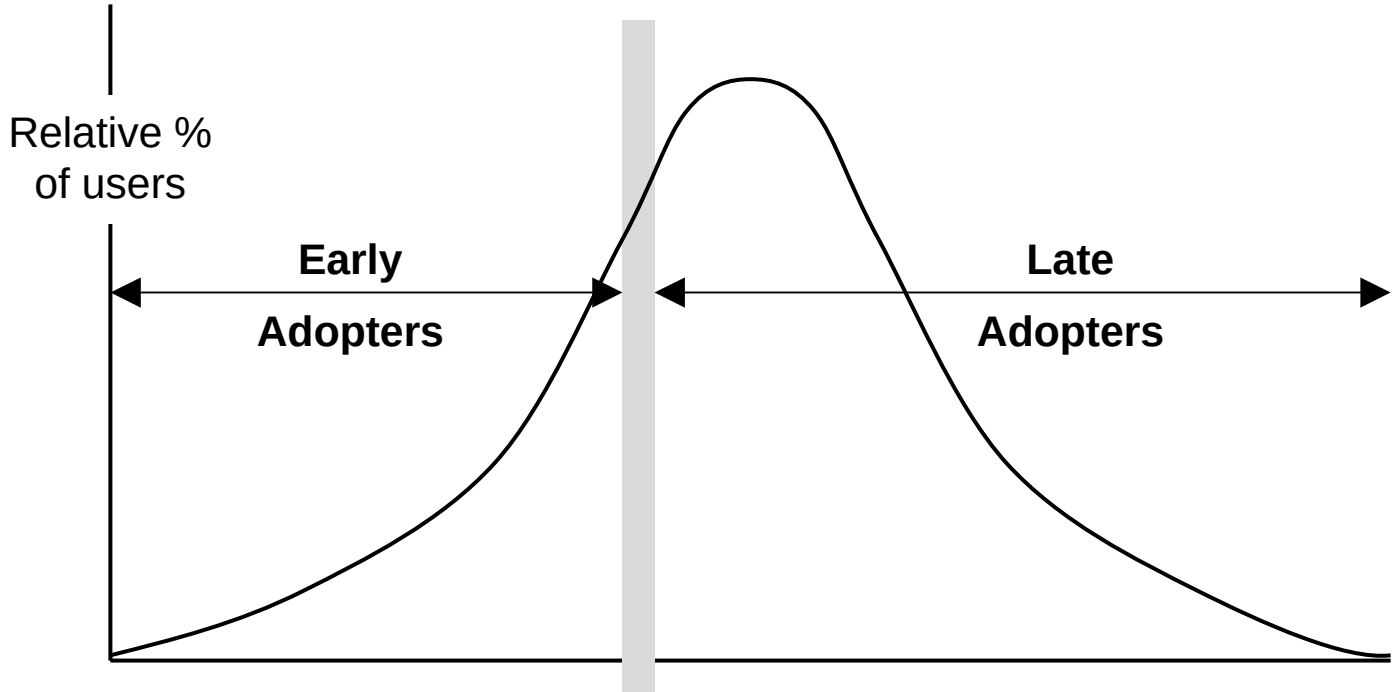
Industry Trends



Industry Trends

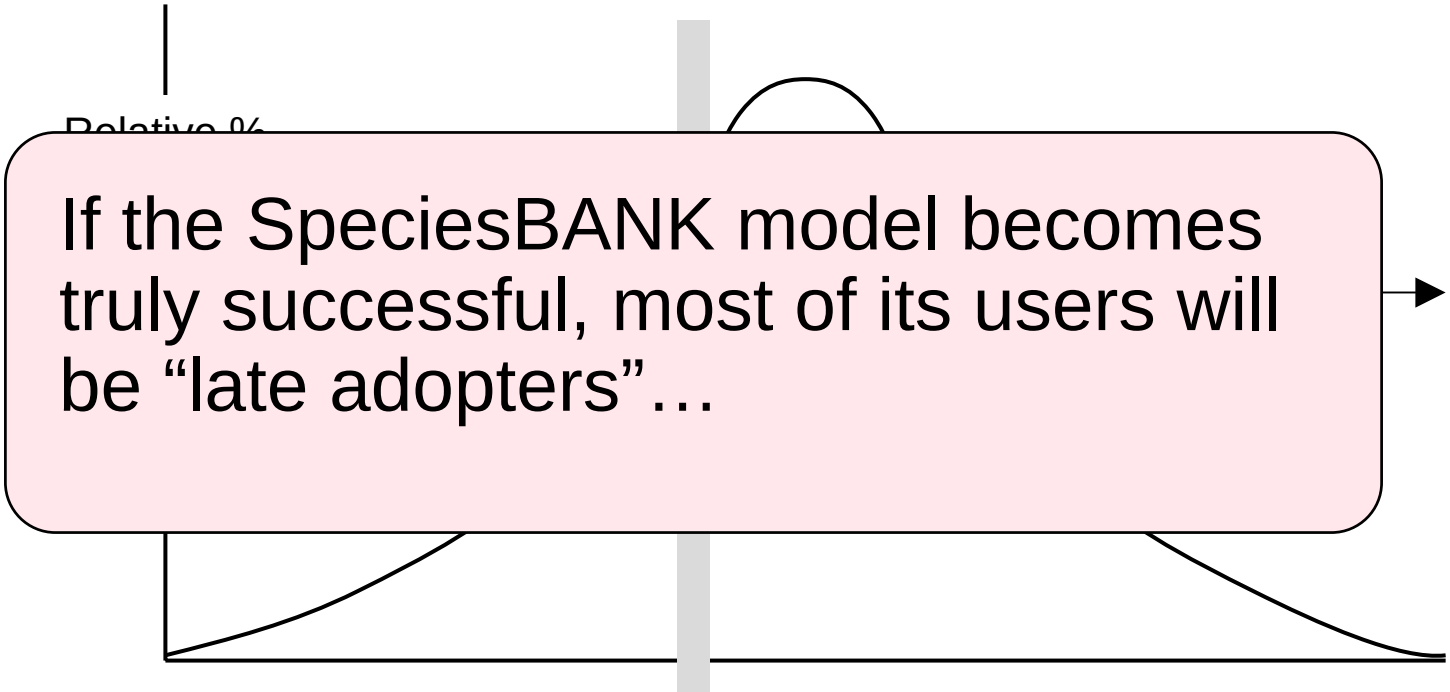


Industry Trends



Early adopters drive the technical capabilities of the system, forcing the bar of acceptable performance upward. However, at some point the bar stabilizes and late adopters come to dominate the market for (and hence the design of) technology products.

Industry Trends



Early adopters drive the technical capabilities of the system, forcing the bar of acceptable performance upward. However, at some point the bar stabilizes and late adopters come to dominate the market for (and hence the design of) technology products.

Making it Work

Universal Interoperability

Universal Interoperability

- Hard…

Making it Work

Logical Simplicity

Logical Simplicity

- In a federated, component-based environment, the biggest challenge is managing complexity.
- This requires a commitment to simplicity.
- Components must be entirely self-contained.
- All inter-component communication occurs only through well defined interfaces.
- Systems must be designed to accommodate change.

Logical Simplicity

- In a federated, component-based environment, the biggest challenge is managing complexity.
- This requires a commitment to simplicity.
- Components must be entirely self-contained.
- All inter-component communication occurs only through well defined interfaces.
- Systems must be designed to accommodate change.

Driving Assumption

- Many use case requirements across the federation will be inconsistent and some will be genuinely contradictory.

Driving Assumption

- Many use case requirements across the federation will be inconsistent and some will be genuinely contradictory.
- The federation must work anyway.

Making it Work

Social Scalability

Social Scalability

- In a truly federated environment, long term success for a federated security model will depend upon social scalability.
- Social scalability CANNOT be achieved through normative pronouncements.
- Experience suggests that social scalability is best achieved through a combination of pure laissez faire individualism and social consequences – i.e., social contracts.

Social Scalability

- In a truly federated environment, long term
- [illegible]
- [illegible]est achieved through a combination of pure laissez faire individualism and social consequences – i.e., social contracts.

Negotiated social contracts – not mandated technical solutions – drive the emergence of standards in a federation.

Social Consequences

- Every individual is free to do whatever he/she chooses.

Social Consequences

- Every individual is free to do whatever he/she chooses.
- Every other individual is free to respond however he/she chooses.

Social Consequences

- Every individual is free to do whatever he/she chooses.
- Every other individual is free to respond however he/she chooses.
- Interactive relationships then sort things out.

Social Consequences

- Every individual is free to do whatever he/she chooses.
- Every other individual is free to respond however he/she chooses.
- Interactive relationships then sort things out.
- Examples:

  One cuts, the other chooses.

Social Consequences

- Every individual is free to do whatever he/she chooses.
- Every other individual is free to respond however he/she chooses.
- Interactive relationships then sort things out.
- Examples:

  I am free to suppress my caller ID; if I do, you are free to refuse to answer my calls.

Social Consequences

- Every individual is free to do whatever he/she chooses.
- Every other individual is free to respond however he/she chooses.
- Interactive relationships then sort things out.
- Examples:

  You are free to run your systems in as stupid and incoherent manner as you choose; if you do, I am free to refuse to have anything to do with your systems.

Logical Issues

- Rules governing behavior can be permissions or prohibitions.

Logical Issues

- Rules governing behavior can be permissions or prohibitions.
- The union set of contradictory permissions is a very flexible environment.

Logical Issues

- Rules governing behavior can be permissions or prohibitions.
- The union set of contradictory permissions is a very flexible environment.
- The union set of contradictory prohibitions is the null set.

Logical Issues

- Rules governing behavior can be permissions or prohibitions.
- The union set of contradictory permissions is a very flexible environment.
- The union set of contradictory prohibitions is the null set.
- Use case requirements across a federation will be contradictory.

Logical Issues

- If a federated information system is to deliver services greater than the null set, it must be technically implemented on the aggregation of permissions, not prohibitions.
- Behavioral constraints should be achieved on a virtual organization basis, through negotiated social contracts.

Logical Issues

- For example, the components of a federated information system should make it easy for users to behave according to common standards, but it should not mandate that they do so.

Social Scalability: Required Reading

James Madison
Alexander Hamilton
John Jay

The Federalist Papers

Social Scalability: Required Reading

James Madison
Alexander Hamilton
John Jay

The Federalist Papers

There is no better source of ideas on how to build systems that work in a decentralized social environment.

Remember, you can't change human nature, so you must design systems that work **despite** human nature.

Social Scalability: Required Reading

Al
Ja
Jo

THEOREM:

When there is no authority to **compel** participation in standard systems, then one must **entice** participation in standard systems.

Social Scalability: Required Reading

Ale
Ja
Jo

OUR TASK:

To devise an effective and enticing SpeciesBANK system…

SpeciesBANK

http://www.esp.org/rjr/RJR-GBIF.pdf