The advent of digital technology and its promise for biodiversity research

James Hanken
### State of the Art, 1974

<table>
<thead>
<tr>
<th>Location</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oaxaca</td>
<td>Oaxaca, 14.1 mi NE and 2.7 mi E, 8900' \n</td>
</tr>
<tr>
<td></td>
<td>Oaxaca, 3.5 mi S of Ranch el Pinto, 8350'</td>
</tr>
<tr>
<td></td>
<td>Rancho Mira Cumbre, 8600'</td>
</tr>
<tr>
<td></td>
<td>Cerro San Felipe \n</td>
</tr>
</tbody>
</table>

---

**Thorius sp.**

- 126294: Ixtlan de Juarez, 13.6 mi NE, 9900' \n- 126295: Cerro San Felipe, 9000' \n- 126296: Ixtlan de Juarez, 13.8 mi NE, 9800' \n- 126297: Oaxaca, 14.1 mi NE and 2.7 mi E, 8900' \n- 126845: Cofradia, ca. 8 mi SW San Vicente \n- 132095: Llano de los Flores; 2700 m \n- 132096: "Cerro Pelon" SW slope, 13 mi NE of Llano de los Flores (by road); 2700 m

**Thorius macdougalli**

- 118913: Llano de las Flores, 12 mi N of Ixtlan de Juarez, 9200' \n- 119705: Llano de las Flores, 2870 m \n
**Thorius narisovalis**

- 119702: 3.5 mi S of Ranch el Pinto, 8350' \n- 119703: Rancho Mira Cumbre, 8600' \n
**Thorius pulmonaris**

- 117247: Cerro San Felipe \n- 119701: 3.5 mi S of Rancho el Pinto, 8350' \n- 121886: ca. 10 mi NE of Cuidad Oaxaca
State of the Art, 1974
<table>
<thead>
<tr>
<th>Identification</th>
<th>Taxonomy</th>
<th>Location</th>
<th>Recorded By</th>
<th>Sex</th>
<th>Date</th>
<th>Map</th>
<th>Media</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCZ Herp A-137376</td>
<td>Amphibia: Thorius boreas</td>
<td>Mexico, Oaxaca: Cerro Pelon, 50 km N Guelatao [VERBATIM ELEVATION:2850...</td>
<td>Gabriela Parra-Olea</td>
<td></td>
<td>1997-10-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCZ Herp A-137378</td>
<td>Amphibia: Thorius macedougallii</td>
<td>Mexico, Oaxaca: 1 km E Cerro Machin</td>
<td>Gabriela Parra-Olea</td>
<td></td>
<td>1999-08-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCZ Herp A-137380</td>
<td>Amphibia: Thorius narisovalis</td>
<td>Mexico, Oaxaca: Cerro San Felipe, 9.3 km W La Cumbre [VERBATIM ELEVATI...</td>
<td>Gabriela Parra-Olea</td>
<td></td>
<td>1999-08-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVZ Amphibian and</td>
<td>Amphibia: Thorius pennatulus</td>
<td>Mexico, Veracruz: S-facing slope of River Canyon, 2 km N Teocelo, on r...</td>
<td>Collector(s): Samuel S. Sweet</td>
<td>undetermined</td>
<td>1972-12-19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reptile specimens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>106399</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVZ Amphibian and</td>
<td>Amphibia: Thorius pennatulus</td>
<td>Mexico, Veracruz: Barranca, 2 km N of Teocelo</td>
<td>Collector(s): Theodore J. Papenfuss</td>
<td>undetermined</td>
<td>1972-12-19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reptile specimens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>106409</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
53+ million pages
133,000+ titles | 220,000+ volumes
181+ million instances of taxonomic names
750+ in-copyright titles licensed for BHL
Agreements with 330+ licensors

*Stats as of February 2018
entirely that of the Plethodontidae. This was the less to have been anticipated, as the general characters of the only genus are those of the genus Spelerpes. The history of the metamorphosis is as yet unknown. The only known genus is Mexican.

THORIUS Cope.

Parietal and palatine bones rudimentary, represented by cartilage and membrane. Posterior nares therefore not separated from orbit; sphenoidal patches of teeth entirely united; tongue boletoid, free in front. Toes distinct, rudimental, 4—5.

The tarsal bones consist of astragalus, calcaneum, a scaphoid and three minute cuneiform bones. The metatarsals and phalanges are fully ossified, as are the corresponding elements of the fore limbs.

This genus is highly interesting, as indicating the lowest grade of ossific deposit found among the tailed Batrachians, accompanied by characters of full development in other respects. Thus, while the cranium is but imperfectly ossified, and less developed than in a comparatively early larval stage of Amblystoma, the tongue, vertebral column, and extremities have advanced far beyond its larval condition, which is permanent in the latter genus, and the branchial apparatus disappears while the individuals are but little more than half their adult size.

It is represented as yet by but one species, from Mexico, of terrestrial habits.

THORIUS PENNATUS Cope.

American Naturalist, 1869, 222.

This is a small species, with smooth skin, very weak limbs, and stout tail. The head is scarcely wider than the neck; it is not flattened, the loreal region is rather elevated and distinct, and the muzzle slightly prominent. The upper lip is sometimes truncate, with infranarial angle prominent, sometimes regularly rounded. The nostril is larger than any known salamander, its diameter equalling half that of the pupil.

The vomerine teeth are situated on a transverse, elevated crest, which is a little behind between the inner nares, and though curved backwards, is but little interrupted medially. Each half contains four teeth, perhaps five when complete. The sphenoidal series is large, pyriform, the anterior extremity narrowed and prolonged to opposite the middle of the orbits.

There are thirteen costal folds; three and one-half of their interspaces are by the extended hind limb, from its origin. The toes are very
Figure 1 Geographic distribution of Thorius in western and southern Oaxaca, Mexico. Type localities of six named species are denoted by open symbols; additional localities are denoted by closed symbols. Known localities of T. sp. 2 and T. sp. 3, two unnamed Oaxacan species, are also shown (Rovito et al., 2013). Small closed circles denote four localities where two or three species are sympatric or nearly sympatric (from left to right): Heroica Ciudad de Tlaxiaco, San Vicente Lachixio, Zaachila and Cerro San Felipe.

Figure 5 X-ray micro-computed tomography (μCT) scans of adult skulls. (A–C) Thorius pinicola, MCZ A-136429, paratype, male; (D–F) T. longicaudus, MCZ A-137819, holotype, female; (G–I) T. tlaxiactus, MVZ 183447, paratype, male; (J–L) T. minitusimus, IBH 23011, female; and (M–O) T. narisovalis, MVZ 162257, female. Each skull is shown in dorsal (left), ventral (middle) and left lateral views. The skeleton of the right hand is visible in G. Total length of each skull is only 3–4 mm; scale bar, 1 mm.
What happened, and when?
ORNIS2 PORTAL LAUNCHES
The ORNIS2 portal was launched.

FISHNET2 LAUNCHES
The FishNet2 portal was launched.

ORNIS FUNDED
ORNIS was funded by the National Science Foundation (NSF), DBI-0345448, 2004-2010.

HerpNET2 PORTAL LAUNCHES
The HerpNET2 portal was launched.

HERPNET FUNDED
HerpNET was funded by the National Science Foundation (NSF), DBI-0132303, 2002-2008.

MANIS FUNDED
MaNiS was funded by the National Science Foundation (NSF), DBI-0108161, 2001-2006.

FISHNET FUNDED
FishNet was funded by the National Science Foundation (NSF), DEB-9910159, 1998-2001.
iDigBio Portal has 109M records for 327M specimens with 23M associated media records

April 2018:
- iDigBio is working with 708 collections in 396 institutions
- 20 TCNs and 17 PENs in 50 states
- 1 coordinating center (“Hub”): iDigBio

Blue = NSF-funded, green = other.
Why?
Scientific collections provide an essential base for developing scientific evidence and are an important resource for scientific research, education, and resource management. Policies and procedures for maintaining, preserving, and developing Federal scientific collections while also increasing access to those collections for appropriate use are, therefore, central to their value.
The Network Integrated Biocollections Alliance will develop an inclusive, vibrant, partnership of U.S. biological collections that collectively will document the nation’s biodiversity resources and create a dynamic electronic resource that will serve the country’s needs in answering critical questions about the environment, human health, biosecurity, commerce, and the biological sciences.
Digitization is a Global Effort

- Atlas of Living Australia
- CONABIO (Mexico)
- Canadensys (Canada)
- India Biodiversity Portal
- Swedish LifeWatch
- PORBIOTA (Azores, Portugal)
- Edaphobase (Germany)
- Ocean Biographic Information System
Global Biodiversity Information Facility

June 1st 2018: 984,441,947 occurrence records
Data Error is a Serious Problem

“while the world’s [plant] collections have more than doubled since 1970, more than 50% of tropical specimens ... are likely to be incorrectly named. This finding has serious implications for the uncritical use of specimen data from natural history collections.”
Widespread sampling biases in herbaria revealed from large-scale digitization

“distance to herbaria explained 45% of the variance in collecting effort in AU, 29% in SA and 12.3% in NE, with a higher density of specimens closer to herbaria.”

There was a significant trend toward collection on weekends (Saturdays and Sundays) in NE ... and midweek in SA and AU.

Key words: collector bias, geographic bias, herbarium, regional flora, sampling bias, temporal bias, trait bias.
Most Basic Digitization Remains To be Done

• 3–4 billion specimens are housed in natural history collections.
• Core data are online for < 50%.
• Many fewer specimens are georeferenced, imaged.
• Linkages among data types are typically limited and inconsistent.
• Special collections are largely untouched.
Distributed System of Scientific Collections (DISSCo)

- 114 European museums in 21 countries
- 1.5 billion specimens!
- 2018-2022: Preparatory phase (Innovation and Consolidation programmes), e.g., *European Open Science Cloud Summit*, 11 June 2018
- 2019-2024: Construction phase/programme
- 2024-2025: Deployment phase

**Our mission**
To unify European natural sciences collections, effectively transforming a scattered and fragmented access model to a central facility with an integrated data-driven pan-European research infrastructure.

- 1.5 billion specimens held in European facilities
- 80% of global biodiversity described held by European collections
- 100 collaborative projects
- 5,000 scientists
- 15,000 visiting scientists annually
- 3,000 scientific publications annually
- 10 million public visitors annually
- 25 million web visitors annually
oVert TCN
Microcomputed Tomography (µCT)
FMNH: ca. 400,000 slides, including 10,000 slides of type specimens.
MCZ: > 32,000 slides of the Harvard Embryological Collection (Charles Minot).
HUH: > 15,000 slides depicting wood anatomy, fungi, diatoms, insect pests, etc.
NMNH: > 2,300 types/paratypes on slides and > 10,000 histological slides just for IZ.
32,000 slides comprising 2,535 series of paraffin-embedded, stained and sectioned vertebrate embryos
Resurrecting embryos of the tuatara, *Sphenodon punctatus*, to resolve vertebrate phallus evolution

Thomas J. Sanger¹, Marissa L. Gredler² and Martin J. Cohn¹,²,³

¹Department of Molecular Genetics and Microbiology, ²Department of Biology, and ³Howard Hughes Medical Institute, University of Florida, PO Box 103610, Gainesville, FL 32610, USA

The breadth of anatomical and functional diversity among amniote external genitalia has led to uncertainty about the evolutionary origins of the phallus. In several lineages, including the tuatara, *Sphenodon punctatus*, adults lack an intermittent phallus, raising the possibility that the amniote ancestor lacked external.

---

Fishing for jaws in early vertebrate evolution: a new hypothesis of mandibular confinement

Tetsuto Miyashita*

*Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada
How the brain is wired

Welcome to SlideAtlas

Girder: a data management platform
Automated Slide Scanner

Batch-process

• Scan file: ~ 2.5 Gb per slide (1½" x 3").
• Scan time: ~ 25 min per slide.
• Maximum 80 slides per run.
Dark Data No More
Ultimate Goal: Linking Research to Collections*

* Wingfield (2012) BIO Advisory Committee

Protein → alignment → phylogenetics → sequence → GenBank → vouchered specimen → museum
Challenges Facing Humanity in the Next 50 Years

BIOLOGICAL SCIENCES

- Environment & Economy
- Biodiversity
- Science Workforce
- Natural Resources
- Public Education
- Synthetic Biology
- Human Systems
- Bio/Nano Systems
- Adaptation & Sustainability
- Climate

Wingfield (2012)
BIO Advisory Committee
“e-research infrastructure for biodiversity knowledge generation”

“integration of various taxon-level data types (genome, morphology, distribution and species interactions) within a phylogenetic and environmental framework.... Living in this future will require the adoption of new ways of integrating scientific knowledge into societal decision making.”
The macroscope can be considered the symbol of a new way of seeing, understanding, and acting.

Joël de Rosnay (1975)
a "macroscope" ... unlike the microscope to see the very small, or the telescope that can see far away, ... is a system of software and algorithms to bring all of Earth's complex data together to analyze it by space and time for meaning.... The time has come for us to truly see more of the world thanks to being able to digitize and collect new sources of data from millions of connected objects....
An integrated assessment of the vascular plant species of the Americas

We present an integrated assessment of all known native species of vascular plants in the Americas. Twelve regional and national checklists ... were merged into a single list. Our publicly searchable checklist includes 124,993 species..., which correspond to 33% of the 383,671 vascular plant species known worldwide.

“...and the new go out of his mind if the wonders don’t cease soon.”
—Alexander von Humboldt (1)

A marínez de Oviedo’s chronicles (2) from 1526 contain the first European accounts...
Fig. 4. Species described per year. The number of plant species (basionyms) described per year from 1753 to 2015 for the Americas list (blue bars), and the cumulative number of accepted species (black line).
We assemble a census of the biomass of all kingdoms of life. This analysis provides a holistic view of the composition of the biosphere and allows us to observe broad patterns over taxonomic categories, geographic locations, and trophic modes.

One of the most fundamental efforts in biology is to describe the composition of the living world. Centuries of research have yielded an increasingly detailed picture of the species that inhabit our planet and their respective roles in global ecosystems. In describing a complex system like the biosphere, it is critical to quantify the abundance of individual components of the system (i.e., species, broader taxonomic groups). A quantitative description of the distribution of biomass is essential for taking stock of biosequestered carbon (1) and modeling global biogeochemical cycles (2), as well as for understanding the historical effects and future impacts of human activities.
Global biomass of different forms of life (A) and different animal taxa (B)*

* "We estimate that the contribution of reptiles and amphibians to the total animal biomass is negligible...."
Building—and Sustaining—the Macroscope Will Require Targeting Diverse User Communities Beyond Natural History & Systematics

- Conservation and land-use planning
- Climate change and invasive species
- Medicine (infectious disease, drug discovery, etc.)
- Agriculture and forestry
- Physiology, developmental biology, ecology, etc.
- Biologically inspired engineering
- K-12 education