

Notes from Nature: a tool for citizen science engagement and transcription of natural history collection records.

Or

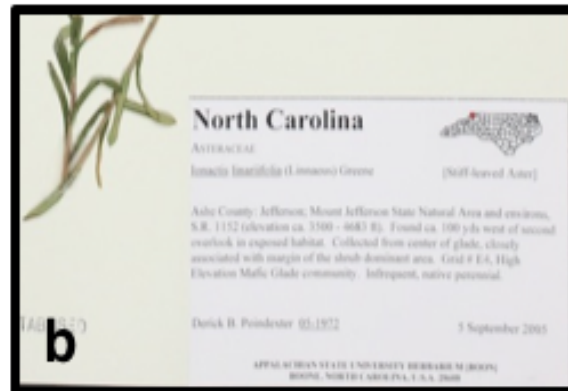
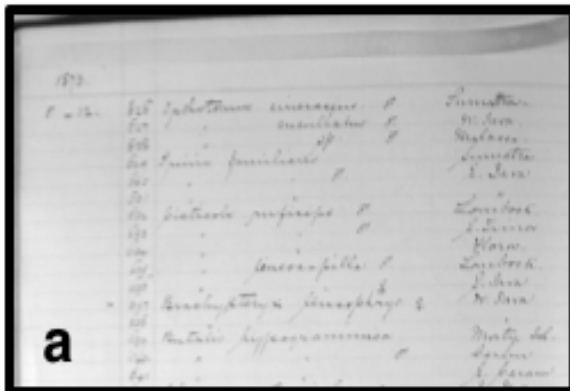
NFN: Where Now?



Digitization Challenges

Getting the job done this century

- Volume
- Data heterogeneity
- Practice heterogeneity
- QA/QC

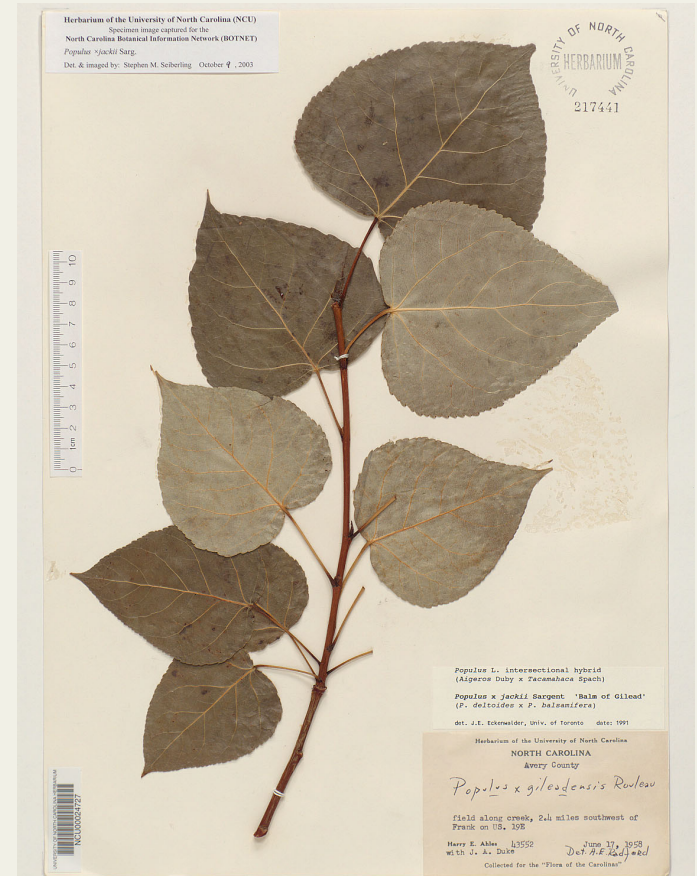


Digitization Challenges

Public engagement

Promote knowledge about:

- Organisms
- Museum collections
- Museum practices
- Use of specimens in research
- Collaborative digitization
- People, places, taxa




Transcribe museum records to

TAKE NOTES FROM NATURE

START TRANSCRIBING

ORGANIZATIONS INVOLVED





New York Botanical Garden

Macrofungi

START TRANSCRIBING

25%
DONE

The Macrofungi Collection Consortium is a partnership of 35 institutions across the U.S that collectively will digitize about 1.5 million specimens that have been collected the past 150 years.

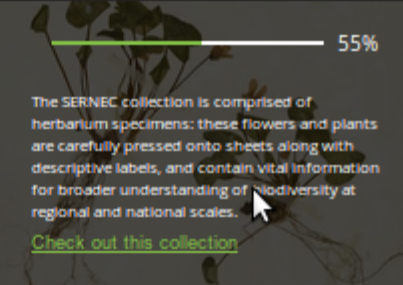
TEMPORAL RANGE
~1900 / ~2000

TAXONOMIC COVERAGE
Fungi

Organized By Collections

4 collections available

Browse through them and select your favorite



55%

The SERNEC collection is comprised of herbarium specimens: these flowers and plants are carefully pressed onto sheets along with descriptive labels, and contain vital information for broader understanding of biodiversity at regional and national scales.

[Check out this collection](#)

Herbarium

SERNEC

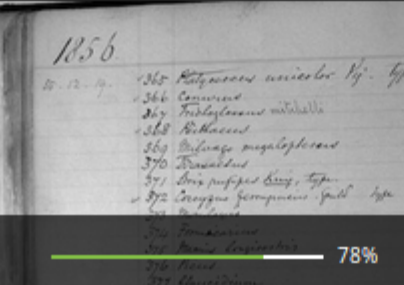


42%

Libellulidae

Calbug

Essig Museum Collections



78%

1856

365 *Polyporus versicolor* Nj. type
366 *Cantharellus*
367 *Trichoglossum mitchellii*
368 *Rhizoglyphus*
369 *Polyporus mycelioides*
370 *Boletus*
371 *Boletus purpuratus* type
372 *Boletus gemmatus* type
373 *Boletus*
374 *Boletus*
375 *Boletus*
376 *Boletus*
377 *Boletus*

Ornithological

Natural History Museum



1%

Macrofungi

New York Botanical Garden

Collections Pages

Notes from **Nature.**

[COLLECTIONS](#) [ABOUT](#) [DISCUSS](#) [BLOG](#) [ROBUR](#)



Essig Museum Collections
Calbug

START TRANSCRIBING **42% DONE**

Calbug is an effort to digitize terrestrial arthropod-specimen records with a focus on California. The project, funded by the National Science Foundation, represents a collaboration among eight entomological collections in California to capture data from specimen labels. The Calbug Science Team will then use the collection data to assess how arthropods have responded to climate change and habitat modification. Drawing from over a century of insect collecting in the region, our goal is to develop a database of over 1 million geographically referenced specimens.

TEMPORAL RANGE
1880 / Present

GEOGRAPHIC CONTEXT
California and surrounding region

TAXONOMIC COVERAGE
Insects | Spiders

PROGRESS - 31,558 / 75,000 **42%**

[START TRANSCRIBING](#)

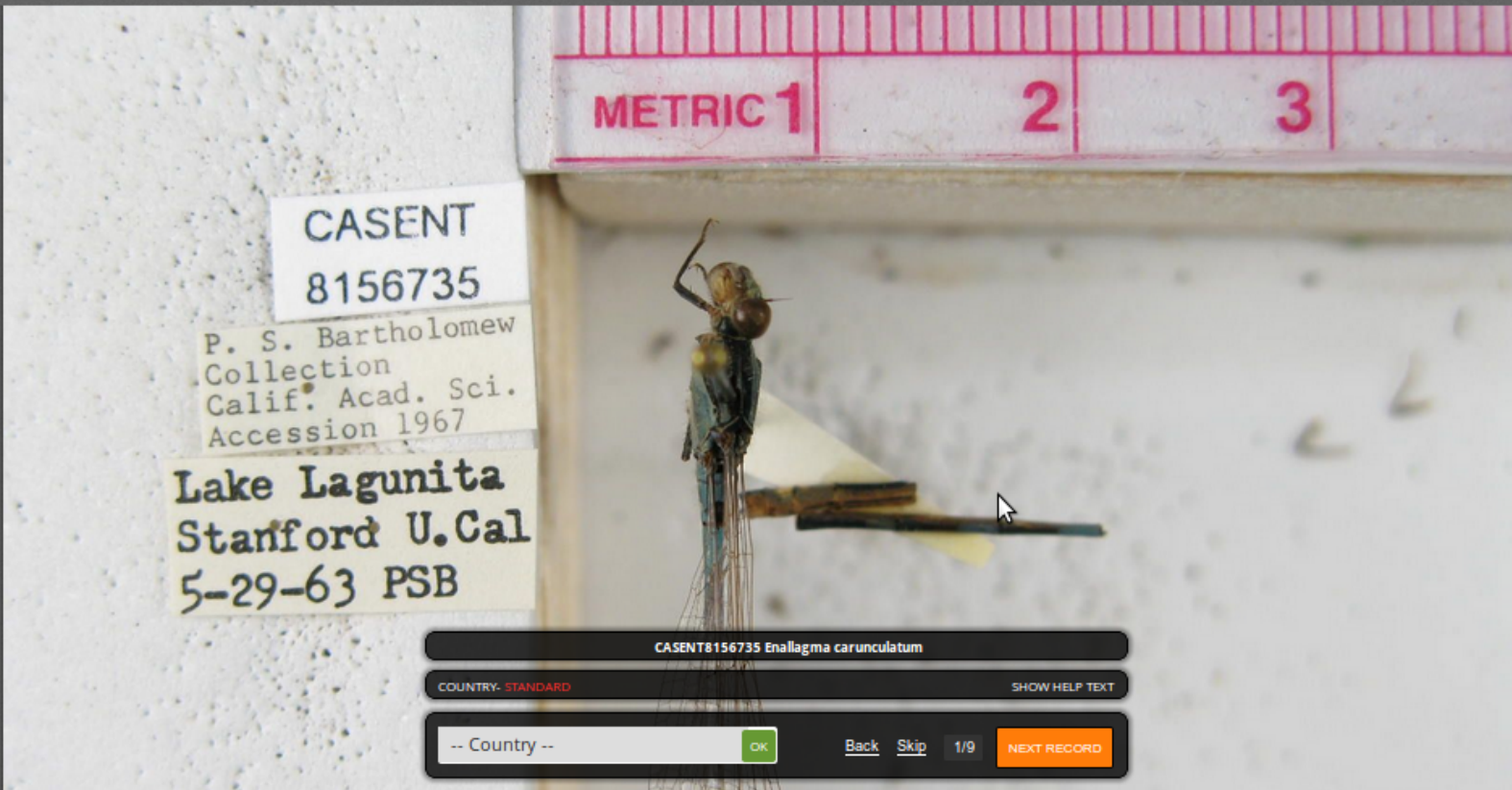
Transcribe to earn!



Transcription Interface

Discuss

Skip Record



CASENT
8156735
P. S. Bartholomew
Collection
Calif. Acad. Sci.
Accession 1967
Lake Lagunita
Stanford U. Cal
5-29-63 PSB

METRIC 1 2 3

CASENT8156735 *Enallagma carunculatum*

COUNTRY: STANDARD SHOW HELP TEXT

-- Country -- OK Back Skip 1/9 NEXT RECORD

Engagement

- Blog posts
 - Content
 - Social Media
 - Profiles
 - http://blog.notesfromnature.org/2013/05/11/profile-of-notes-from-nature-citizen-scientist-el_lion/
- Talk page (super important!)
- Novel incentives, connections
 - iDigBio/NfN Hackathon
 - Species naming/NfN event in Nov.
- A Transcriber's Life

Notes from Nature Forum

(and a huge advantage of the Zooniverse platform)

Featured discussions

Welcome to Notes from Nature!

Posted in Just chatting

94 posts / 51 participants

Popular hashtags

- handwriting
- error
- unclear
- collector
- cross
- uncleardate
- unclearlocation
- incomplete
- time
- no-collector
- date
- multiple
- helpme
- nobug
- moderator
- unclearabbreviation
- genus
- paratype
- genusonly
- malaise-trap

Currently Online

2 users

Recent Object Comments...



Cropped at right edge at "curtisi", appears to be

by Mr. Kevvy 2 hours ago



Teichalm forest, Styria, Austria

by darryluk 6 hours ago



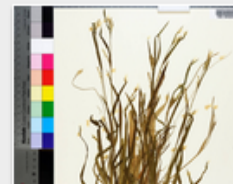
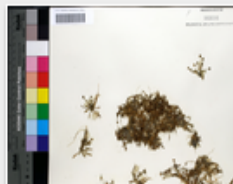
Collector R C Reed, in Chile

by darryluk 6 hours ago



Not only the labels are almost impossible to read,

by T-M 16 hours ago



Each Registered User Gets A Profile Page

The screenshot displays a user profile for 'robgur'. The main section, titled 'A transcriber's life', shows a vertical timeline of transcription activity. The right sidebar contains statistics for 'Entries transcribed' (287) and 'Collections' (2), along with sections for 'Herbarium' (64 transcriptions done), 'Calbug' (223 transcriptions done), and 'Your Badges' (Transcribe to earn badges).

robgur

A transcriber's life

See robgur activity

Time Ago	Activity
an hour ago	You transcribed 103 items
6 months ago	You transcribed 98 items
7 months ago	You transcribed 96 items
7 months ago	You transcribed 86 items
8 months ago	You transcribed 84 items
8 months ago	You transcribed 73 items

287
Entries transcribed

2
Collections

Herbarium

64 transcriptions done

Calbug

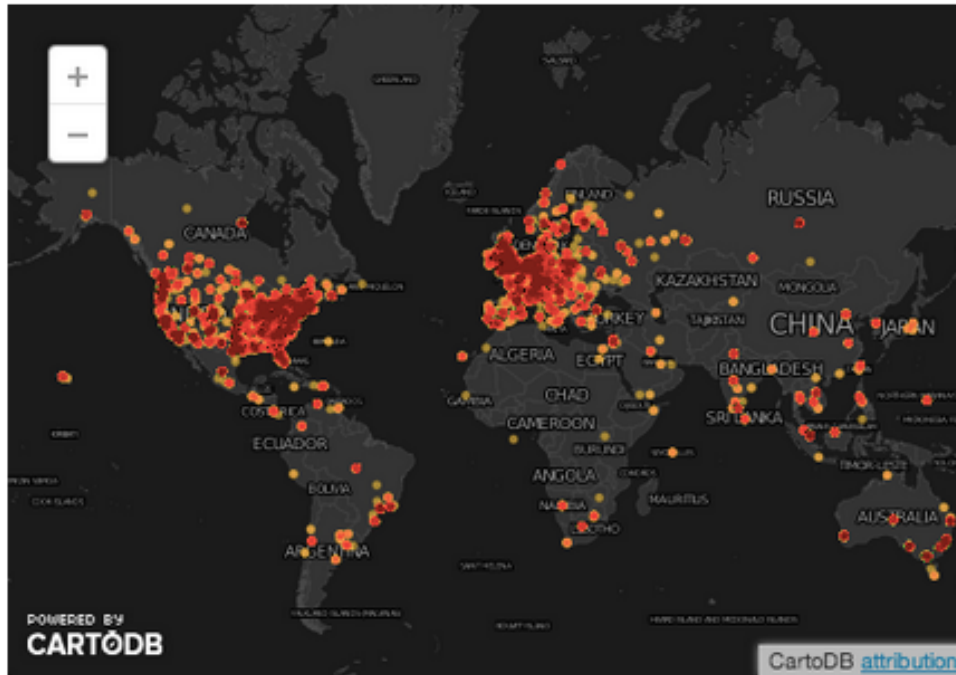
223 transcriptions done

Your Badges

Transcribe to earn badges

The badges are arranged in a 2x3 grid. The top row includes a snake, a butterfly, and a leaf with a drop. The bottom row includes a leaf on a hill, a leaf with a drop, and a leaf on a hill.

SUCCESS!



- ~400,000 transcriptions
- ~5000 volunteers signed up
- Worldwide distribution of transcribers
- New collections (bird ledgers, macrofungi)
- New and exciting engagement ideas

- http://blog.notesfromnature.org/2013/07/01/stats_first_look/

Press

- Wired
 - <http://www.wired.com/wiredscience/2013/05/museum-collection-insects/?pid=7002>
- Daily Mail
 - <http://www.dailymail.co.uk/sciencetech/article-2330424/Can-decipher-Darwins-handwriting-Museum-calls-help-deciphering-handwritten-labels-collection.html>
- BBC
 - Webscape, Kate Russell
- NPR story today!



EEK! SUCCESS!

- Growing pains
- Dealing with “onboarding” new collections
- Dealing with transcribed data on the back end
- Continuing to stay involved with transcription community
- Organization models.
- Funding growth
- Sustainability
- Dealing with all the stuff we wished we did better

Rush. Biv. Prov. Pk.
 15 mi. SE Kenora, Ont.
 CAN. VI-28-1970
 W.O. Powers, Collr.

516d7fa6e a3052046a 00000a	EMEC224981 Cicindela limbalis.jpg	4/26/2013 0:19:00	6/28/ 1970	W.O. Powers	Canada		Rush Biv Prov Pk 15 mi SE Kenora	Ontari o
516d7fa6e a3052046a 00000a	EMEC224981 Cicindela limbalis.jpg	5/25/2013 22:00:00	6/28/ 1970	W.O. Powers	Canada		Rushing River Provincial Park	Ontari o
516d7fa6e a3052046a 00000a	EMEC224981 Cicindela limbalis.jpg	6/1/2013 3:45:00	6/28/ 1970	W. O. Powers	Canada	Kenora	Rushing River Provincial Park, 15 miles southeast of Kenora	Ontari o
516d7fa6e a3052046a 00000a	EMEC224981 Cicindela limbalis.jpg	7/1/2013 12:08:00	6/28/ 1970	W.O. Powers	Canada	Kenora District	Rush. Biv. Prov. Pk. 15mi. SE Kenora	Ontari o

IN PRACTICE: RECONCILING TRANSCRIPTIONS

1. Label image

Label: 3 mi. E. Cresta Powerhouse. Hwy. 70



2. Transcriptions *via* Notes from Nature

Transcriber 1: 3 mi. E Cresta Powerhouse. Hwy. 70

Transcriber 2: 3 mi. East Cresta Powerhouse. Hwy. 70

Transcriber 3: 3 miles E. Cresta Powerhouse. Hwy. 70

Transcriber 4: 3 mi. E. Cresta Powerhouse

Character consensus method

3. Align transcriptions using alignment algorithms

```
3_mi._._E._._Cresta_Powerhouse._Hwy._70
3_mi._._East_Cresta_Powerhouse._Hwy._70
3_miles_E._._Cresta_Powerhouse._Hwy._70
3_mi._._E._._Cresta_Powerhouse_-----
```

4. Find consensus

```
3_mi._._E._._Cresta_Powerhouse._Hwy._70
```

5. Remove gaps

3 mi. E. Cresta Powerhouse. Hwy. 70

Token consensus method

3. Tokenize and align transcriptions

```
A_B_C_D_E_F_G
A_B_H_D_E_F_G
A_I_J_D_E_F_G
A_B_J_D_E
```

4. Find consensus

```
A_B_J_D_E_F_G
```

5. Un-tokenize

3 mi. E. Cresta Powerhouse. Hwy. 70

BETTER WAYS TO UNDERSTAND MOTIVATIONS OF CITIZEN SCIENTISTS



Competitors

Defined by:
A focus on winning, rank, and direct peer-to-peer competition.

Engaged by:
Leaderboards, Ranks



Achievers

Defined by:
A focus on attaining status and achieving preset goals quickly and/or completely.

Engaged by:
Achievements



Socialites

Defined by:
A focus on socializing and a drive to develop a network of friends and contacts.

Engaged by:
Newsfeeds, Friends Lists, Chat



Explorers

Defined by:
A focus on exploring and a drive to discover the unknown.

Engaged by:
Obfuscated Achievements

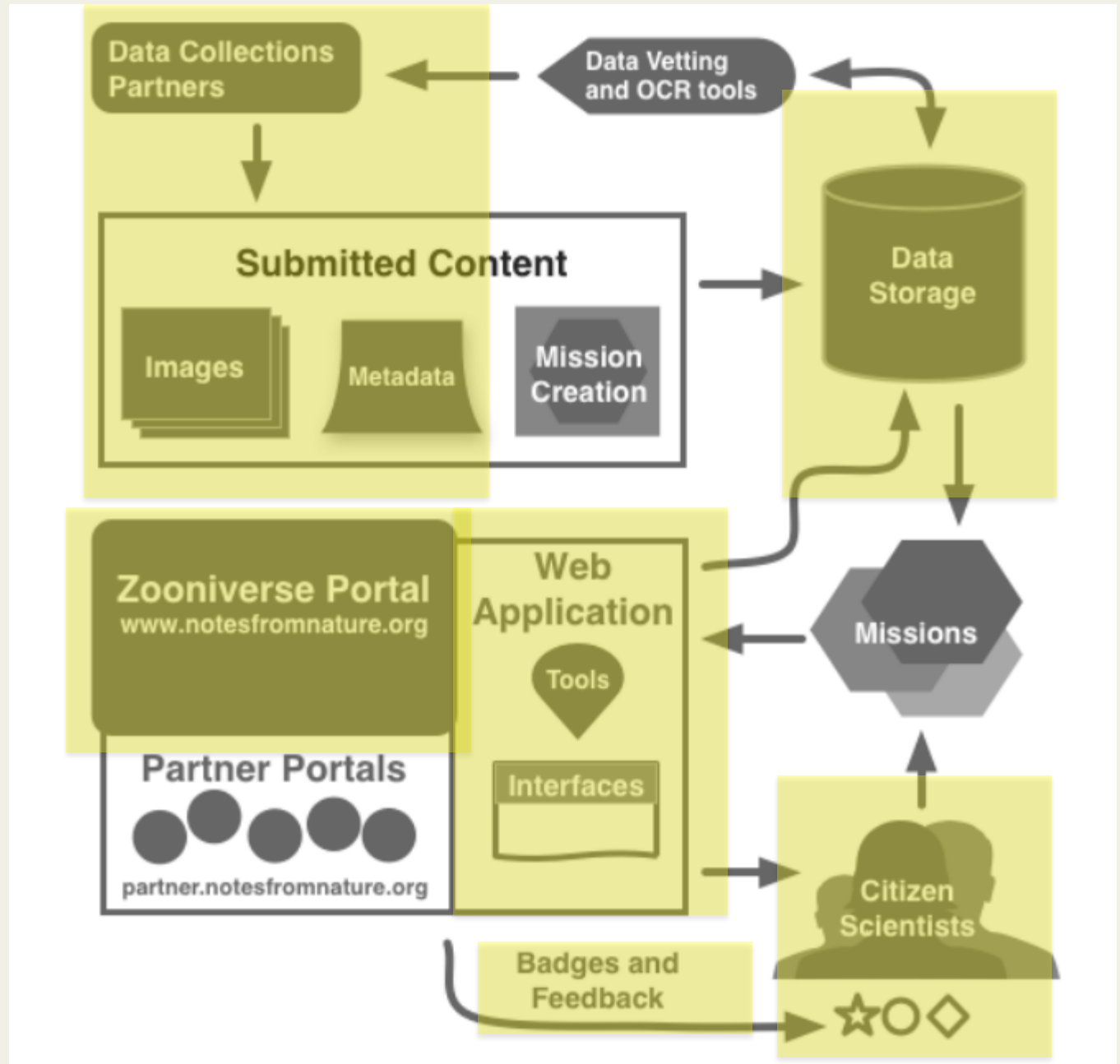
B
A
D
G
E
S



NOTES
FROM
NATURE:

SNAPSHOT
OF THE
PRESENT

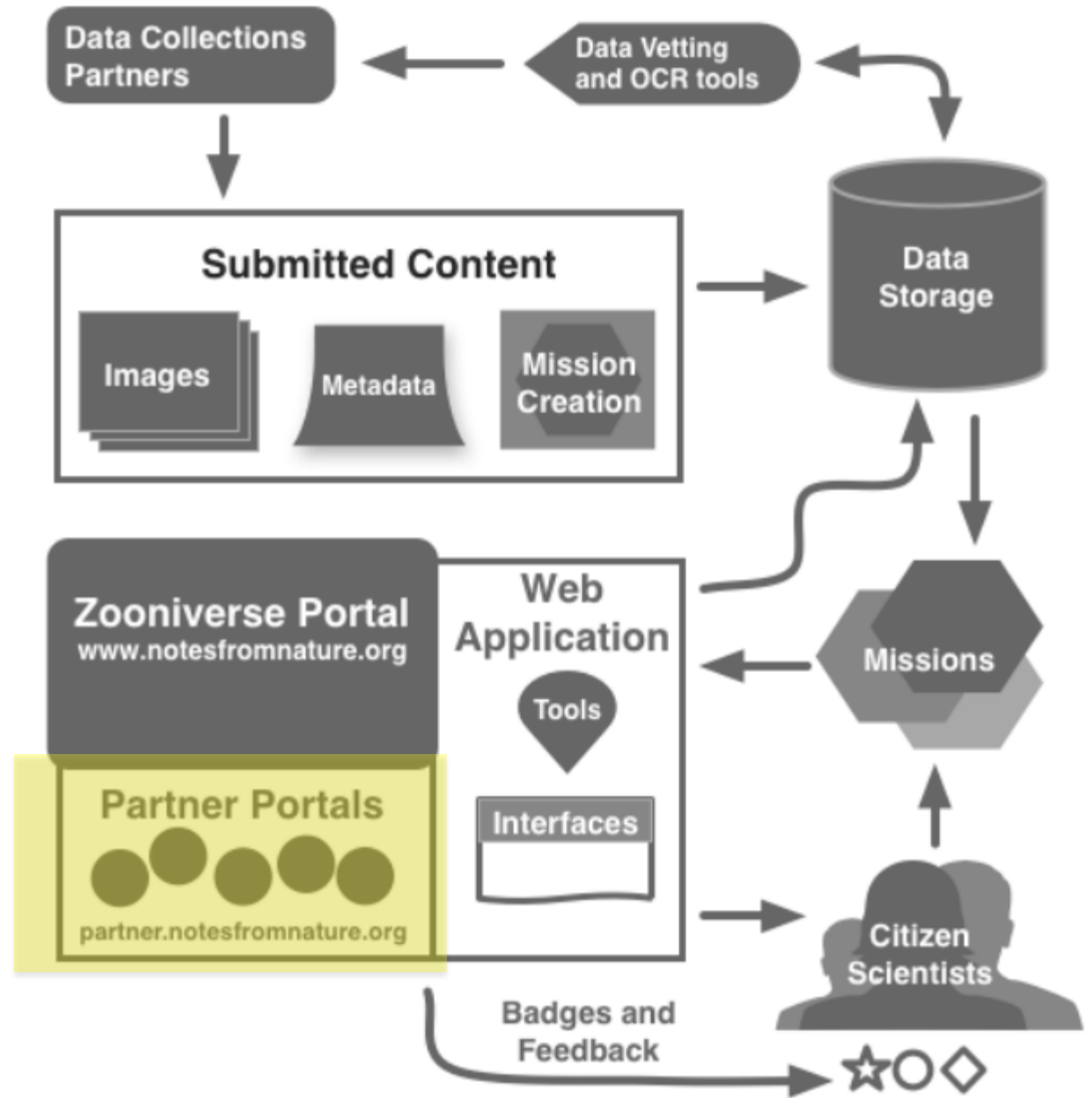
(all the
highlited
items)



NOTES FROM NATURE: MISSING PIECES

#1 – Partner Portals.

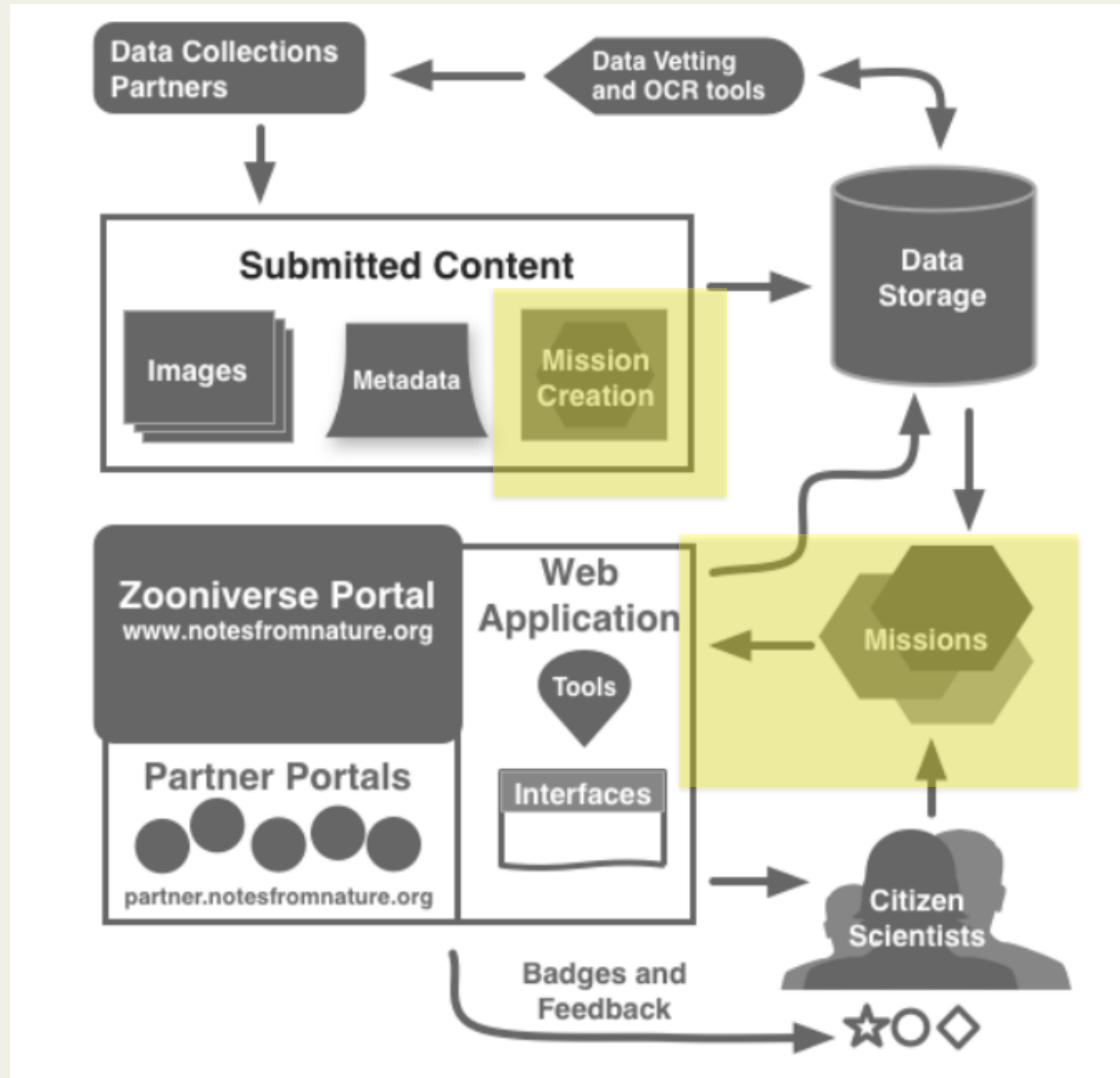
Partial independence from Zooniverse (which everyone wants)



**NOTES
FROM
NATURE:
MISSING
PIECES**

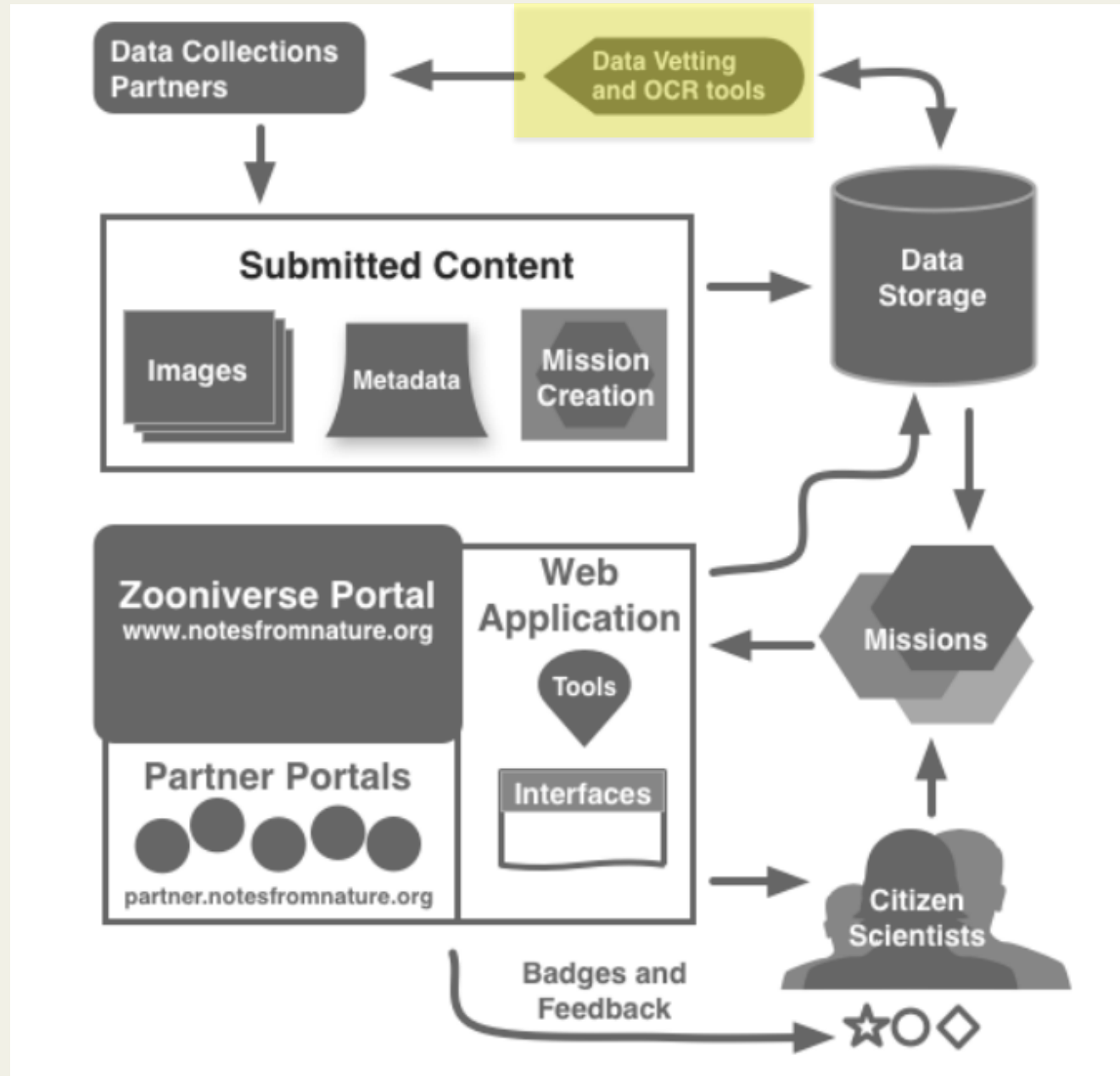
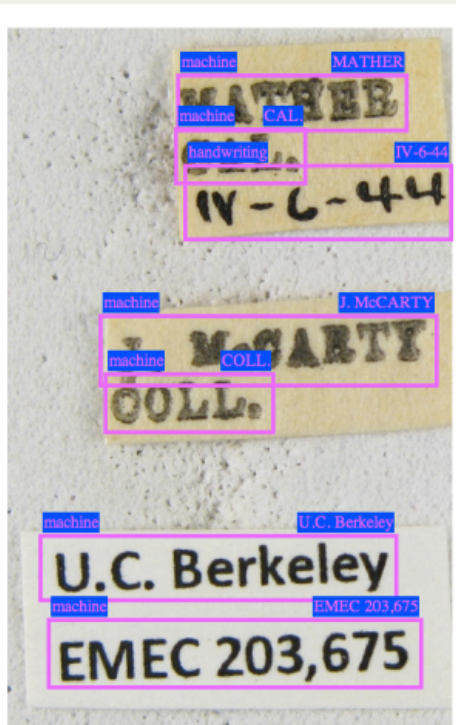
**#2 – Divide
collections into
missions**

**Single most
Important thing
for many providers
(has links to
motivation)**



NOTES FROM NATURE: MISSING PIECES

#3 – Data vetting and OCR





Notes from Nature Governance:

- Steering Committee (Chaired by NEON's own Michael Denslow)
- Needed for joint governance issues
- Includes Zooniverse members
- Provides a single voice to the community