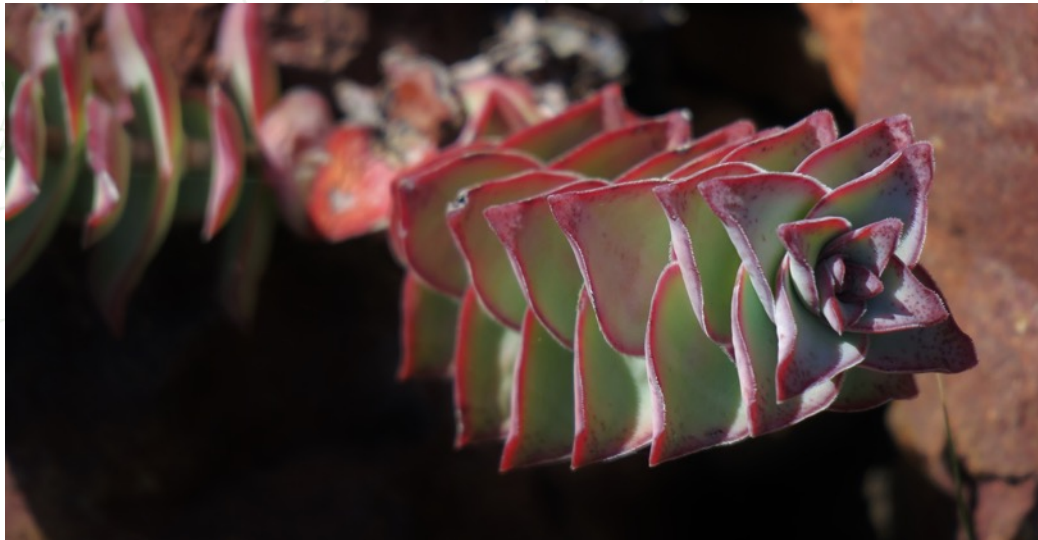




**GBIF**

Global Biodiversity  
Information Facility



## **GBIF meets GDPR: Best practices for data publishing institutions**

Kyle Copas | [kcopas@gbif.org](mailto:kcopas@gbif.org) | [@kylecopas](https://twitter.com/kylecopas) | 25 May 2020

---

## WHY SHOULD DATA PUBLISHING INSTITUTIONS CARE ABOUT GDPR?

- Pervasive international presence and influence of EU 'data subjects'
- GDPR grants important new rights to data subjects—and new responsibilities for data processors
- Sweeping language means **any information related to an identified or identifiable subject** is personal



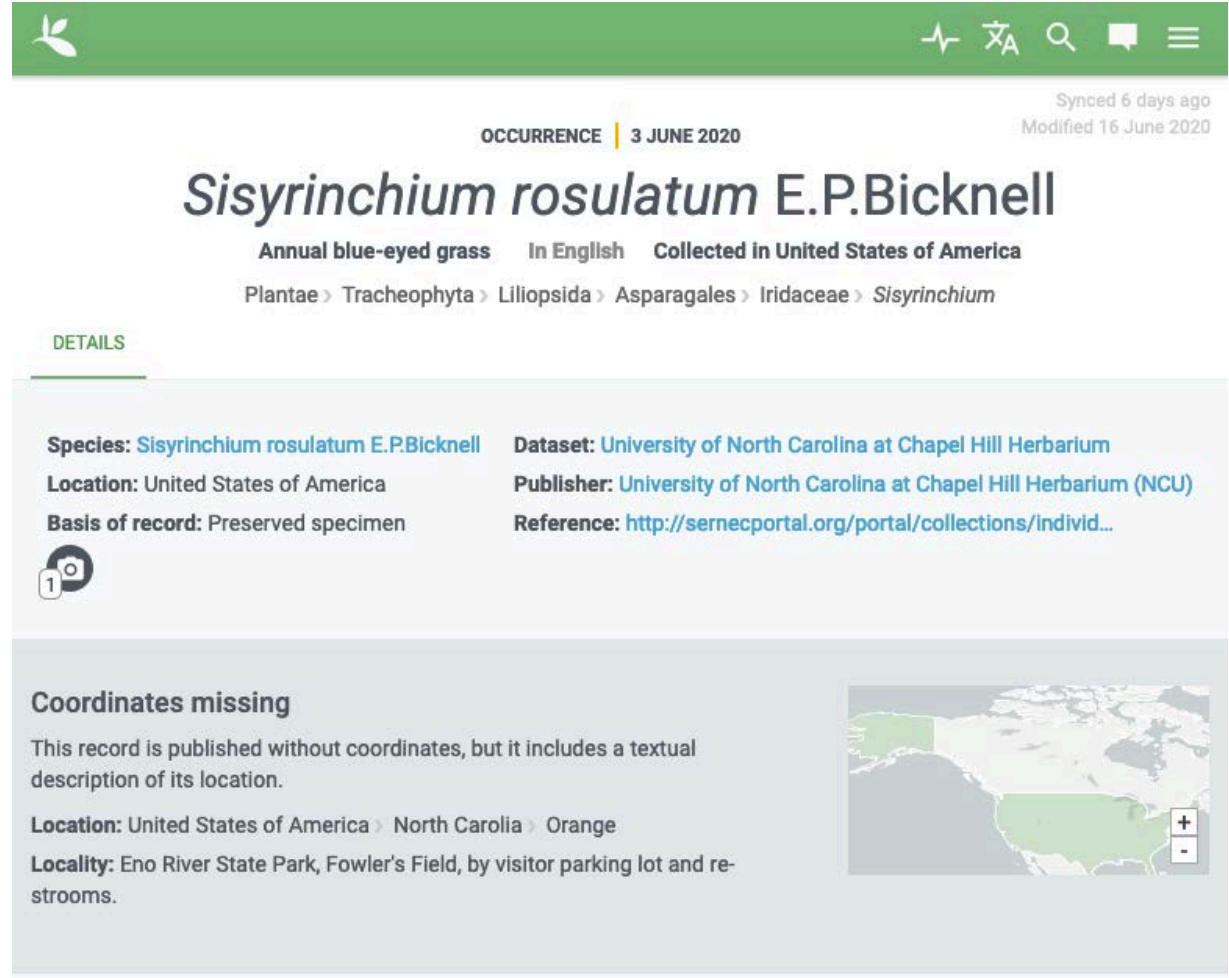
# TAKING IT PERSONALLY

Personally identifiable data in records may include:



# TAKING IT PERSONALLY

Personally identifiable data in records may include:



The screenshot shows a GBIF occurrence record for *Sisyrinchium rosulatum* E.P.Bicknell. The record is dated 3 June 2020 and was collected in the United States of America. The specimen is a preserved specimen. The record is published by the University of North Carolina at Chapel Hill Herbarium (NCU). The location is Orange, North Carolina, at Eno River State Park, Fowler's Field, by visitor parking lot and restrooms. The record is missing coordinates. The GBIF logo is visible in the top right corner of the page.

OCCURRENCE | 3 JUNE 2020 Synced 6 days ago  
Modified 16 June 2020



## *Sisyrinchium rosulatum* E.P.Bicknell

Annual blue-eyed grass In English Collected in United States of America

Plantae > Tracheophyta > Liliopsida > Asparagales > Iridaceae > *Sisyrinchium*

### DETAILS

**Species:** *Sisyrinchium rosulatum* E.P.Bicknell **Dataset:** University of North Carolina at Chapel Hill Herbarium  
**Location:** United States of America **Publisher:** University of North Carolina at Chapel Hill Herbarium (NCU)  
**Basis of record:** Preserved specimen **Reference:** <http://serneportal.org/portal/collections/individ...>


 

**Coordinates missing**

This record is published without coordinates, but it includes a textual description of its location.

**Location:** United States of America > North Carolina > Orange

**Locality:** Eno River State Park, Fowler's Field, by visitor parking lot and restrooms.



Annual blue-eyed grass (*Sisyrinchium rosulatum*), Eno River State Park, NC, USA, 3 June 2020. Record by Carol Ann McCormick, Milo Guthrie Pyne, Julie Thomson via University of North Carolina at Chapel Hill Herbarium, licensed under CC BY-NC 4.0. <https://www.gbif.org/occurrence/2640532582>

# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders

Occurrence			
Term	Interpreted	Original	Remarks
Catalogue number	NCU00434509	NCU00434509	
Occurrence ID	fa15acc3-e962-44fc-abeb-2301607efb19	fa15acc3-e962-44fc-abeb-2301607efb19	
Occurrence remarks	Common Name: "Fairy Stars"... Minute pale blue tepals (red at base interior corolla)	Common Name: "Fairy Stars"... Minute pale blue tepals (red at base interior corolla)	
Other catalogue numbers	673010	673010	
Record number	s.n.	s.n.	
Recorded by	McCormick, Carol Ann; Pyne, Milo Guthrie; Thomson, Julie	McCormick, Carol Ann; Pyne, Milo Guthrie; Thomson, Julie	

Event			
Term	Interpreted	Original	Remarks
Day	3	3	
Month	6	6	
Year	2020	2020	
Event date	2020-06-03T00:00:00	2020-06-03	Altered
Habitat	common in mown lawn area	common in mown lawn area	
Start day of year	155	155	

Identification			
Term	Interpreted	Original	Remarks
Date identified	2020-06-03T00:00:00	2020-06-03	Altered
Identified by	McCormick, Carol Ann	McCormick, Carol Ann	

# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers

Occurrence			
Term	Interpreted	Original	Remarks
Catalogue number	NCU00434509	NCU00434509	
Occurrence ID	fa15acc3-e962-44fc-abeb-2301607efb19	fa15acc3-e962-44fc-abeb-2301607efb19	
Occurrence remarks	Common Name: "Fairy Stars"... Minute pale blue tepals (red at base interior corolla)	Common Name: "Fairy Stars"... Minute pale blue tepals (red at base interior corolla)	
Other catalogue numbers	673010	673010	
Record number	s.n.	s.n.	
Recorded by	McCormick, Carol Ann; Pyne, Milo Guthrie; Thomson, Julie	McCormick, Carol Ann; Pyne, Milo Guthrie; Thomson, Julie	

Event			
Term	Interpreted	Original	Remarks
Day	3	3	
Month	6	6	
Year	2020	2020	
Event date	2020-06-03T00:00:00	2020-06-03	Altered
Habitat	common in mown lawn area	common in mown lawn area	
Start day of year	155	155	

Identification			
Term	Interpreted	Original	Remarks
Date identified	2020-06-03T00:00:00	2020-06-03	Altered
Identified by	McCormick, Carol Ann	McCormick, Carol Ann	

# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers
- observation/collection dates

## Occurrence

Term	Interpreted	Original	Remarks
Catalogue number	NCU00434509	NCU00434509	
Occurrence ID	fa15acc3-e962-44fc-abeb-2301607efb19	fa15acc3-e962-44fc-abeb-2301607efb19	
Occurrence remarks	Common Name: "Fairy Stars"... Minute pale blue tepals (red at base interior corolla)	Common Name: "Fairy Stars"... Minute pale blue tepals (red at base interior corolla)	
Other catalogue numbers	673010	673010	
Record number	s.n.	s.n.	
Recorded by	McCormick, Carol Ann; Pyne, Milo Guthrie; Thomson, Julie	McCormick, Carol Ann; Pyne, Milo Guthrie; Thomson, Julie	

## Event

Term	Interpreted	Original	Remarks
Day	3	3	
Month	6	6	
Year	2020	2020	
Event date	2020-06-03T00:00:00	2020-06-03	Altered
Habitat	common in mown lawn area	common in mown lawn area	
Start day of year	155	155	

## Identification

Term	Interpreted	Original	Remarks
Date identified	2020-06-03T00:00:00	2020-06-03	Altered
Identified by	McCormick, Carol Ann	McCormick, Carol Ann	

# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers
- observation/collection dates
- location information

Location			
Term	Interpreted	Original	Remarks
Country or area	United States of America	United States	Altered
Country code	US		Inferred
County	Orange	Orange	
Locality	Eno River State Park, Fowler's Field, by visitor parking lot and restrooms.	Eno River State Park, Fowler's Field, by visitor parking lot and restrooms.	
State province	North Carolina	North Carolina	

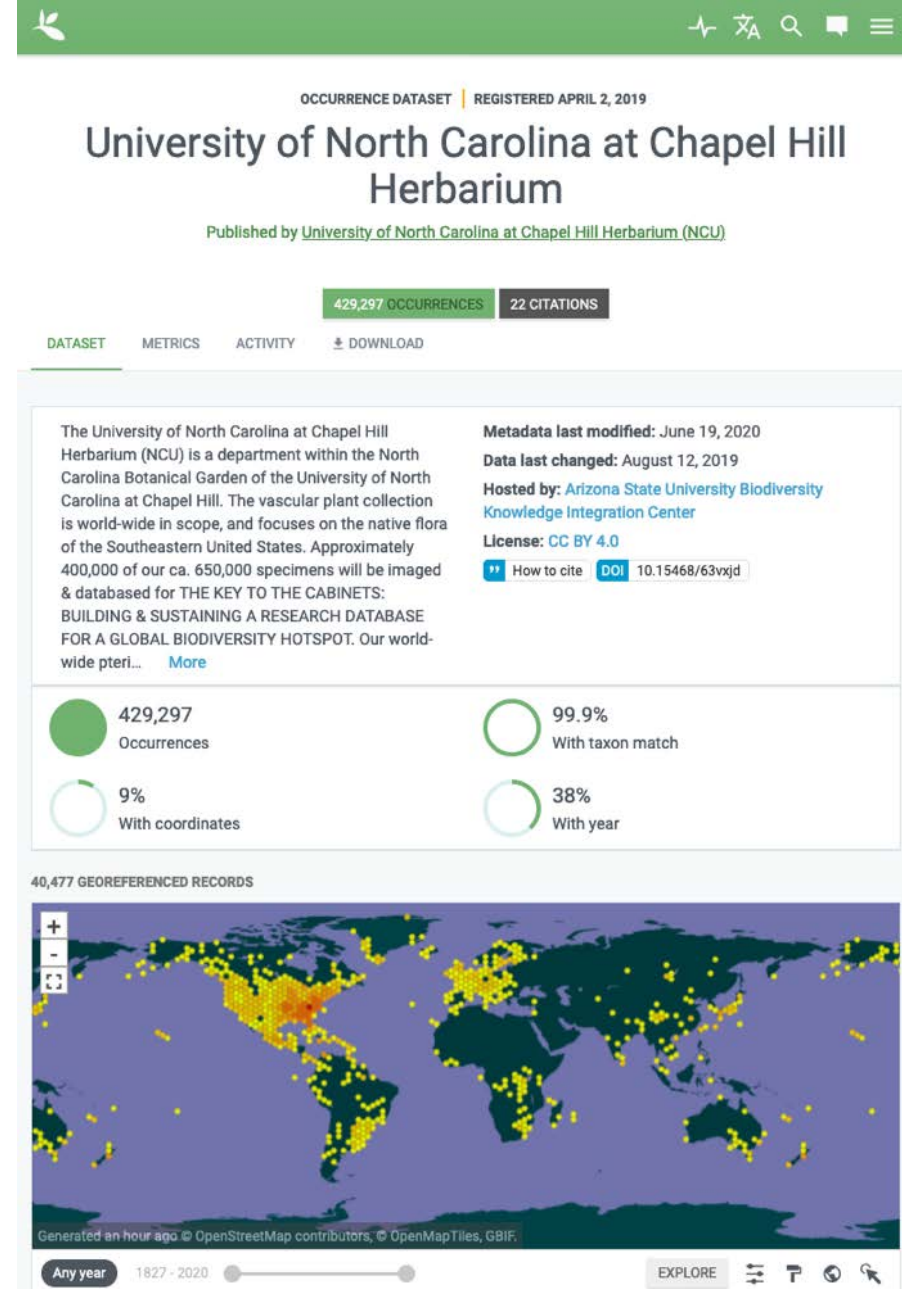


# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers
- observation/collection dates
- location information

Other names names (for many reasons):



Annual blue-eyed grass (*Sisyrinchium rosulatum*), Eno River State Park, NC, USA, 3 June 2020. Record by Carol Ann McCormick, Milo Guthrie Pyne, Julie Thomson via University of North Carolina at Chapel Hill Herbarium, licensed under CC BY-NC 4.0. <https://www.gbif.org/occurrence/2640532582>

# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers
- observation/collection dates
- location information

Other names names (for many reasons):

- Institutional contacts for data publishers, hosts



OCCURRENCE DATASET | REGISTERED APRIL 2, 2019

## University of North Carolina at Chapel Hill Herbarium

Published by [University of North Carolina at Chapel Hill Herbarium \(NCU\)](#)

### Description

The University of North Carolina at Chapel Hill Herbarium (NCU) is a department within the North Carolina Botanical Garden of the University of North Carolina at Chapel Hill. The vascular plant collection is world-wide in scope, and focuses on the native flora of the Southeastern United States. Approximately 400,000 of our ca. 650,000 specimens will be imaged & databased for THE KEY TO THE CABINETS: BUILDING & SUSTAINING A RESEARCH DATABASE FOR A GLOBAL BIODIVERSITY HOTSPOT. Our world-wide pteridophyte collection will be imaged & cataloged for the PTERIDOPHYTE COLLECTIONS CONSORTIUM. NCU also curates plant fossils (esp. Devonian and collections of Patricia Gensel), fungi (esp. collections of W.C. Coker, J.N. Couch, A.B. Seymour), lichens (esp. collections of Gary Perlmutter), bryophytes, and algae (esp. red marine algae of temperate zones collections of Max Hommersand & coralline algae collections of Paul Gabrielson). NCU is open to the public & welcomes researchers; contact Curator to reserve on-campus parking permit.

### Contacts

#### SERNEC

Originator • Metadata author

[herrick.brown@gmail.com](mailto:herrick.brown@gmail.com)

<http://semecportal.org/portal/index.php>

University of North Carolina at Chapel  
Hill Herbarium

Administrative point of contact •

Content provider

[mccormick@unc.edu](mailto:mccormick@unc.edu)

[1] 919 962 6931

<https://ncbg.unc.edu/research/unc-herbarium/>

Any year

1827 - 2020

EXPLORE



# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers
- observation/collection dates
- location information

Other names names (for many reasons):

- Institutional contacts for data publishers, hosts
- Network members (participant and node staff)

Annual blue-eyed grass (*Sisyrinchium rosulatum*), Eno River State Park, NC, USA, 3 June 2020. Record by Carol Ann McCormick, Milo Guthrie Pyne, Julie Thomson via University of North Carolina at Chapel Hill Herbarium, licensed under CC BY-NC 4.0. <https://www.gbif.org/occurrence/2640532582>



OCCURRENCE DATASET | REGISTERED APRIL 2, 2019

## University of North Carolina at Chapel Hill Herbarium

Published by [University of North Carolina at Chapel Hill Herbarium \(NCU\)](#)

### Description

The University of North Carolina at Chapel Hill Herbarium (NCU) is a department within the North Carolina Botanical Garden of the University of North Carolina at Chapel Hill. The vascular plant collection is world-wide in scope, and focuses on the native flora of the Southeastern United States. Approximately 400,000 of our ca. 650,000 specimens will be imaged & databased for THE KEY TO THE CABINETS: BUILDING & SUSTAINING A RESEARCH DATABASE FOR A GLOBAL BIODIVERSITY HOTSPOT. Our world-wide pteridophyte collection will be imaged & cataloged for the PTERIDOPHYTE COLLECTIONS CONSORTIUM. NCU also curates plant fossils (esp. Devonian and collections of Patricia Gensel), fungi (esp. collections of W.C. Coker, J.N. Couch, A.B. Seymour), lichens (esp. collections of Gary Perlmutter), bryophytes, and algae (esp. red marine algae of temperate zones collections of Max Hommersand & coralline algae collections of Paul Gabrielson). NCU is open to the public & welcomes researchers; contact Curator to reserve on-campus parking permit.

### Contacts

#### SERNEC

Originator • Metadata author

[herrick.brown@gmail.com](mailto:herrick.brown@gmail.com)

<http://semecportal.org/portal/index.php>

University of North Carolina at Chapel  
Hill Herbarium

Administrative point of contact •

Content provider

[mccormick@unc.edu](mailto:mccormick@unc.edu)

[1] 919 962 6931

<https://ncbg.unc.edu/research/unc-herbarium/>

Any year

1827 - 2020

EXPLORE



# TAKING IT PERSONALLY

Personally identifiable data in records may include:

- the name of data collectors/recorders
- the name of identifiers
- observation/collection dates
- location information

Other names names (for many reasons):

- Institutional contacts for data publishers, hosts
- Network members (participant and node staff)
- Data users

Annual blue-eyed grass (*Sisyrinchium rosulatum*), Eno River State Park, NC, USA, 3 June 2020. Record by Carol Ann McCormick, Milo Guthrie Pyne, Julie Thomson via University of North Carolina at Chapel Hill Herbarium, licensed under CC BY-NC 4.0. <https://www.gbif.org/occurrence/2640532582>



OCCURRENCE DATASET | REGISTERED APRIL 2, 2019

## University of North Carolina at Chapel Hill Herbarium

Published by [University of North Carolina at Chapel Hill Herbarium \(NCU\)](#)

### Description

The University of North Carolina at Chapel Hill Herbarium (NCU) is a department within the North Carolina Botanical Garden of the University of North Carolina at Chapel Hill. The vascular plant collection is world-wide in scope, and focuses on the native flora of the Southeastern United States. Approximately 400,000 of our ca. 650,000 specimens will be imaged & databased for THE KEY TO THE CABINETS: BUILDING & SUSTAINING A RESEARCH DATABASE FOR A GLOBAL BIODIVERSITY HOTSPOT. Our world-wide pteridophyte collection will be imaged & cataloged for the PTERIDOPHYTE COLLECTIONS CONSORTIUM. NCU also curates plant fossils (esp. Devonian and collections of Patricia Gensel), fungi (esp. collections of W.C. Coker, J.N. Couch, A.B. Seymour), lichens (esp. collections of Gary Perlmutter), bryophytes, and algae (esp. red marine algae of temperate zones collections of Max Hommersand & coralline algae collections of Paul Gabrielson). NCU is open to the public & welcomes researchers; contact Curator to reserve on-campus parking permit.

### Contacts

#### SERNEC

Originator • Metadata author

[herrick.brown@gmail.com](mailto:herrick.brown@gmail.com)

<http://semecportal.org/portal/index.php>

University of North Carolina at Chapel  
Hill Herbarium

Administrative point of contact •

Content provider

[mccormick@unc.edu](mailto:mccormick@unc.edu)

[1] 919 962 6931

<https://ncbg.unc.edu/research/unc-herbarium/>

Any year

1827 - 2020

EXPLORE



# WHAT IS PERSONAL DATA UNDER GDPR?

## *Article 4*

### **Definitions**

For the purposes of this Regulation:

(1) ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

# SENSITIVE v. ORDINARY PERSONAL DATA

## *Article 4*

### **Definitions**

For the purposes of this Regulation:

(1) ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to **the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;**

Note special categories of ‘sensitive data’—all others are ‘ordinary’ personal data

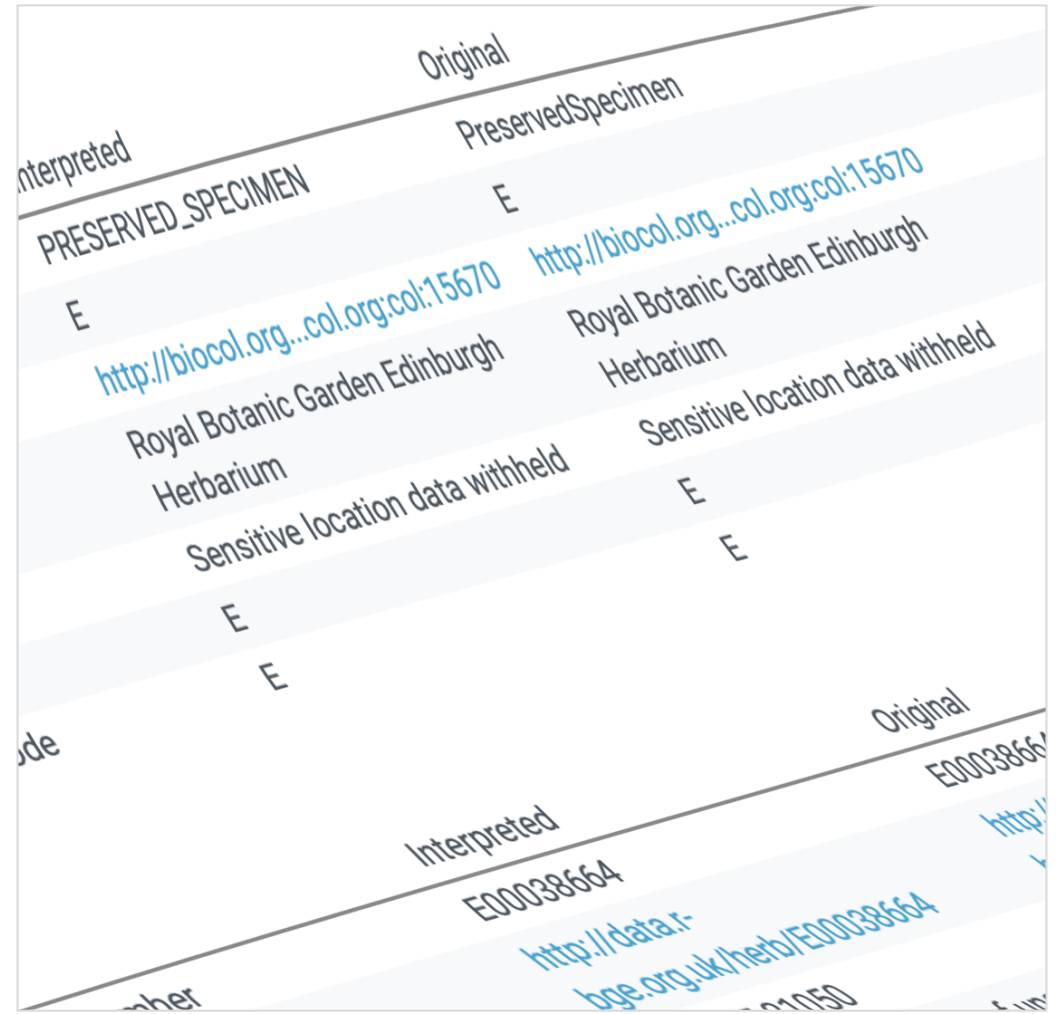
# WE'RE IN THE U.S., NOT EUROPE—SO WE'RE GOOD, RIGHT?

- U.S. data privacy / protection framework is a patchwork that includes
  - Privacy Act of 1974
  - Electronic Communications Privacy Act 1986
  - HIPAA 1996
  - FERPA
  - U.S. Safe Harbor → EU-US Privacy Shield 2016
  - State regulations
- Operative concept of 'Personally identifiable information' (PID) is different
- Globalized (and justifiable) concerns about individual rights to privacy



# SO THEN EVERYONE NEED EVERYONE'S PERMISSION?

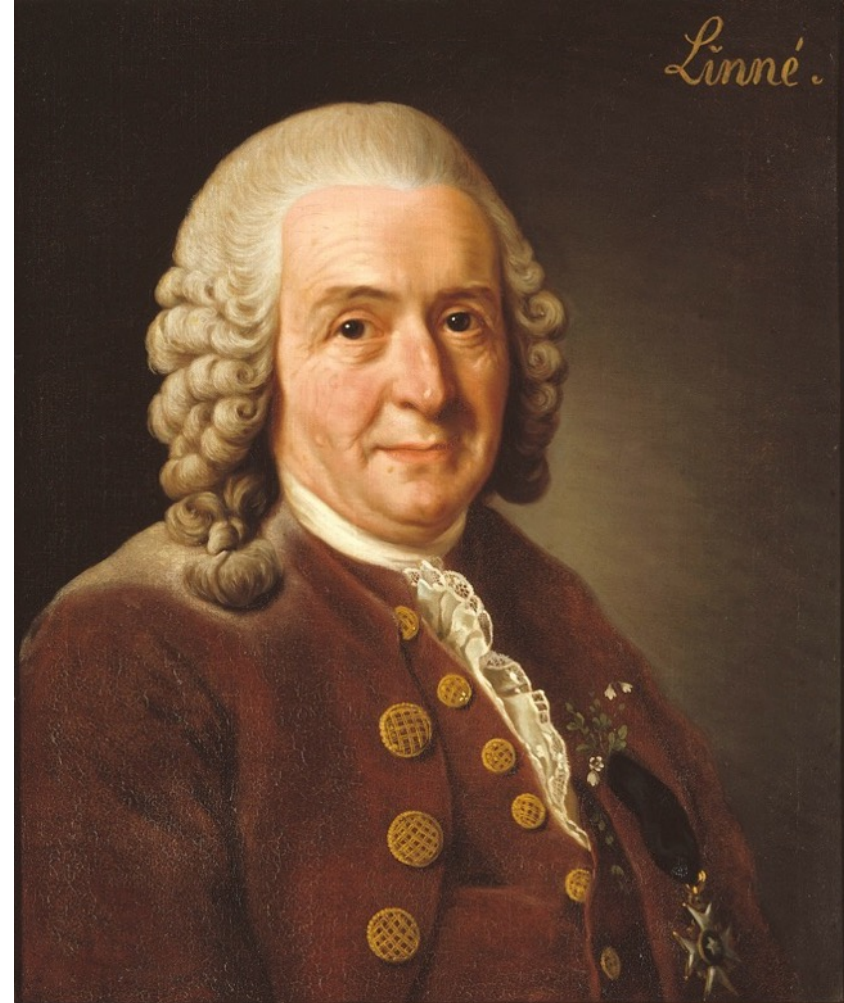
- Not exactly...
- **Consent is only one lawful basis** among six for processing personal data under GDPR—and potentially the weakest
- Others are:
  - Contract
  - Legal obligation
  - Vital interests
  - Public interests
  - **Legitimate interests**





## ARE LINNAEUS'S HEIRS GOING TO SUE US?

- With condolences to the Carlsson family members, **the dead have no rights under GDPR**



# BALANCING TESTS OF DATA SUBJECT'S RIGHTS AND LEGITIMATE INTERESTS

GBIF's position is that the inclusion of this personal information in occurrences serves both

- the public interest
- purposes of scientific and historical research

GBIF does not (and will not) obtain consent directly from recorders and identifiers

If challenged, three tests are necessary:

- **Purpose test:** is there a legitimate interest behind the processing?
- **Necessity test:** is the processing necessary for that purpose?
- **Balancing test:** is the legitimate interest overridden by the individual's interests, rights or freedoms?



## WHAT HAS GBIF ALREADY DONE

- Revised privacy policy to clarify legitimate interests in processing personally identifiable information
- Update describes and differentiates most common relationships with data subjects:
  - network members: participant and node staff
  - data publishers
  - data users
  - data recorders and identifiers



## NEXT STEPS FOR GBIF

- **Work with nodes** to ensure consistent communication with data-publishing networks
- **Engage data publishers directly** with twice-yearly emails focused on implementing policies and best practices, starting for processing personal data
- Provide publishers with guidance for **updating policies and practices**
- **Build trust** with stakeholders on data privacy issues



# NEXT STEPS FOR DATA PUBLISHERS

- **Review disclosures** of collection and use of personal data made during data acquisition/ collections accession processes
- [Update Write privacy policy...]
- **Update privacy policy** to disclose sharing of data to third parties (GBIF, national and thematic networks)
- **Assess processes** for collecting and storing data
- **Inform collectors, recorders and identifiers** of the changes that may impact them



# THANK YOU

Kyle Copas  
kcopas@gbif.org  
@kylecopas

<https://bit.ly/gbifjune20>



Neon cuckoo bee (*Thyreus nitidulus*), Papua New Guinea. Photo 2019 bdriscoll via iNaturalist Research-grade Observations (<https://www.gbif.org/occurrence/2331956823>), licensed under CC BY-NC 4.0 (<http://creativecommons.org/licenses/by-nc/4.0/>).