

Primum Non Nocere:

Before working with Indigenous data,
the ACL must confront ongoing colonialism

Lane Schwartz
University of Alaska
loschwartz@alaska.edu



Primum Non Nocere:

Before working with Indigenous data,
the ACL must confront ongoing colonialism

Lane Schwartz
University of Alaska
loschwartz@alaska.edu



Land Acknowledgement

- The University of Alaska Fairbanks is on the ancestral lands of the **Dena** people of the lower Tanana River.
- The University of Illinois is on the lands of the **Peoria, Kaskaskia, Piankashaw, Wea, Miami, Mascoutin, Odawa, Sauk, Mesquaki, Kickapoo, Potawatomi, Ojibwe, and Chickasaw** Nations. These lands were the traditional territory of these Native Nations prior to their forced removal; these lands continue to carry the stories of these Nations and their struggles for survival and identity.
- The Native languages of Alaska are vital parts of the cultural heritage of Alaska's Native Peoples, and it is my great privilege to work with Akuzipik and the **St. Lawrence Island Yupik** people.



Parthenon



Athens

Parthenon Frieze



The British
Museum

Discover two million years of human history and
culture



London

The British
Museum

Discover two million years of human history and
culture

Elgin Marbles



London

“The Elgin Marbles are no longer part of the story of the Parthenon. They are now part of another story.”

Neil MacGregor, recent former Director, British Museum



Cultural strip mining



**Common NLP
practices perpetuate
colonialism**

Cultural **data** mining



How **not** to be an NLP Colonist



The British
Museum



Discover two million years of human history and
culture

***Primum non
nocere***

**Above all,
do no harm**

ABCs of ethical research

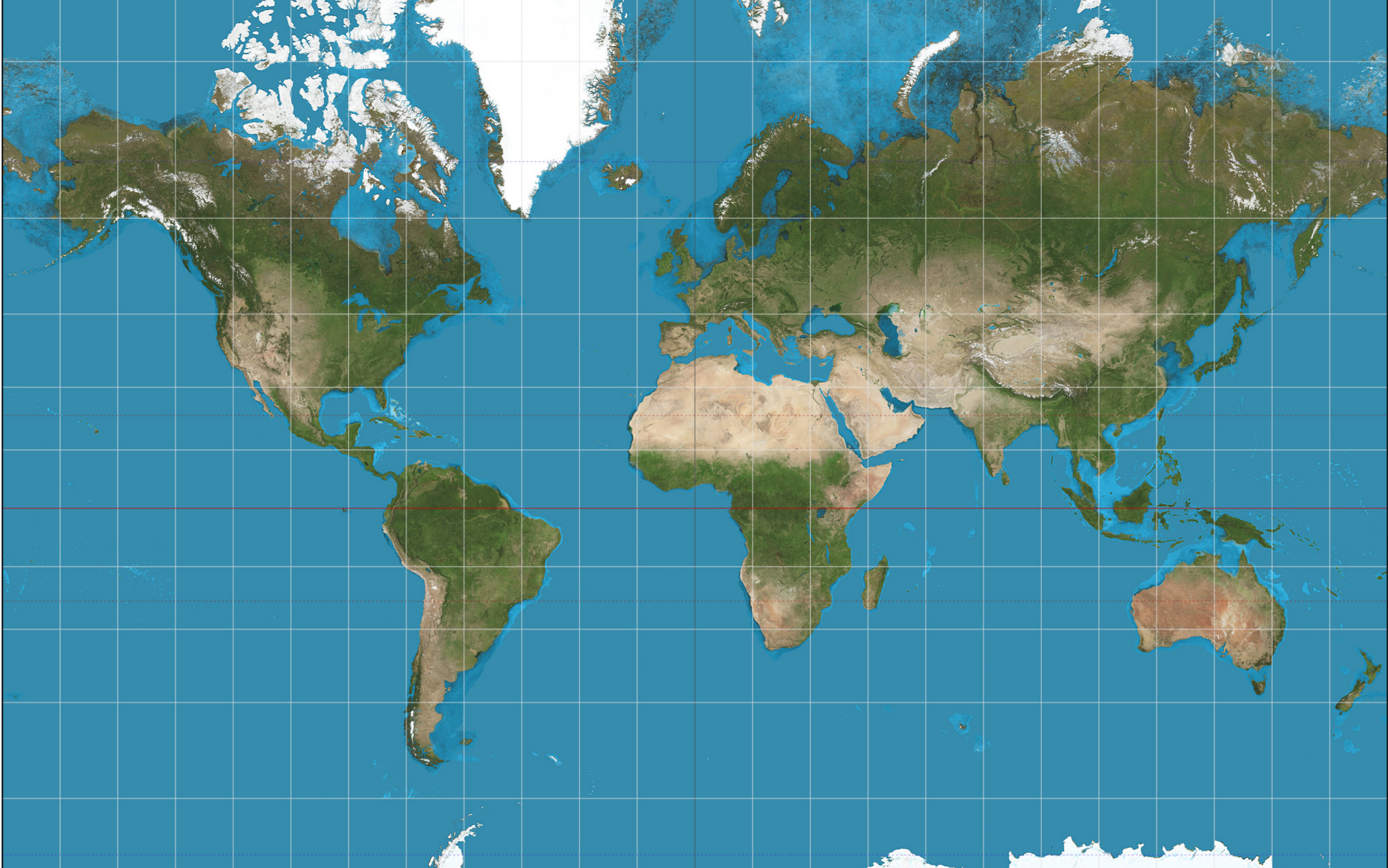
Accountability

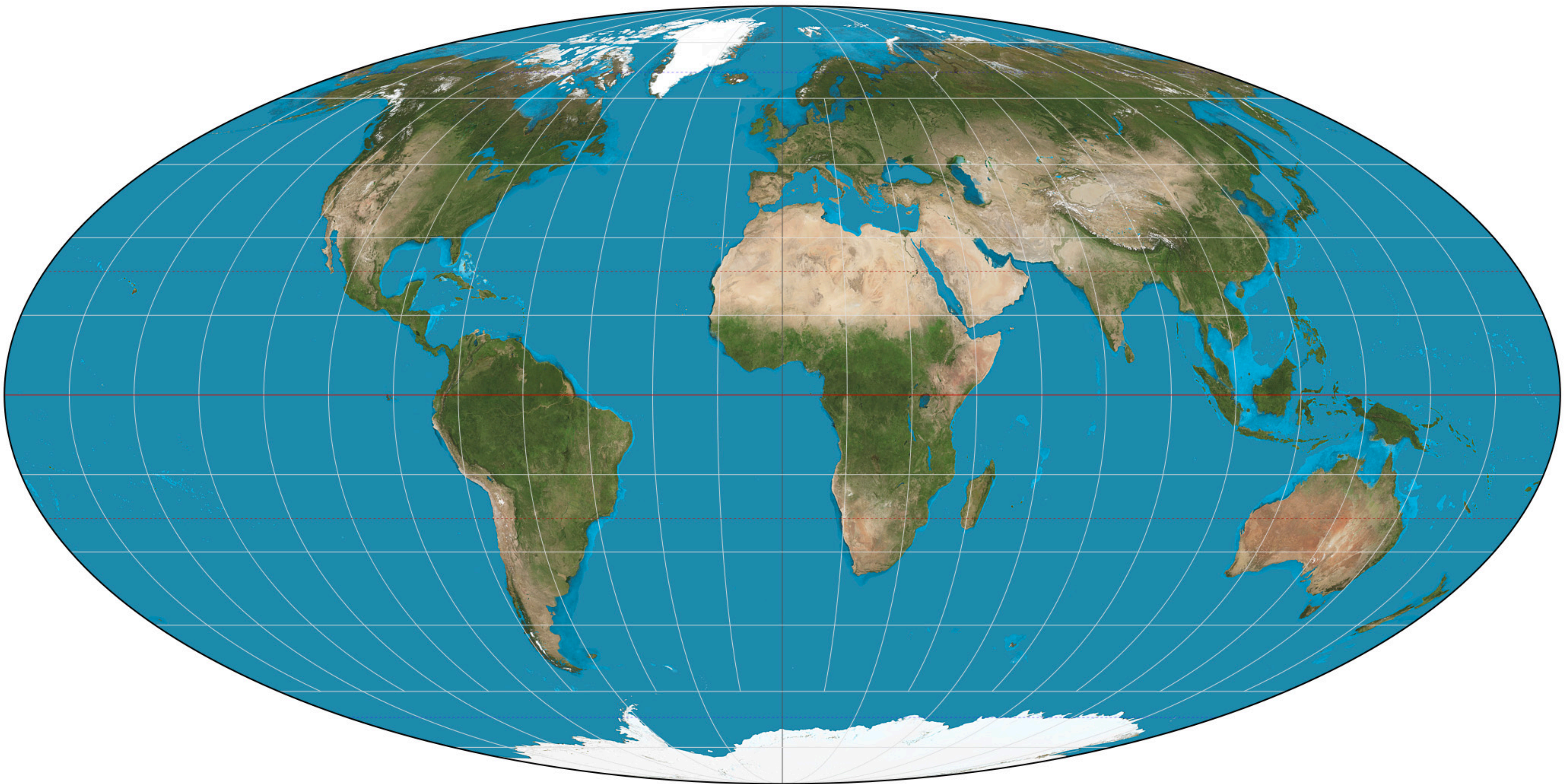
Beneficence

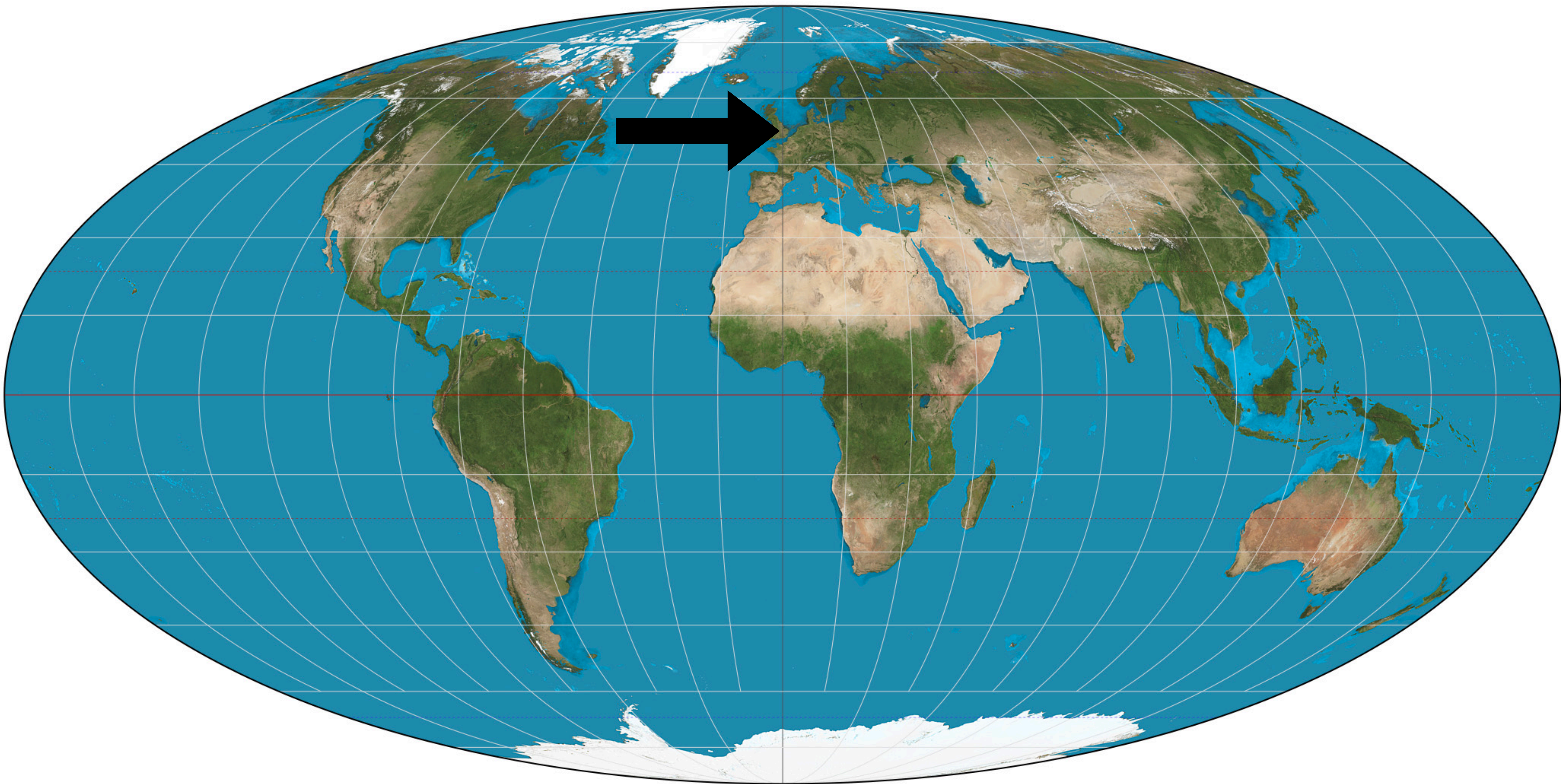
Cognizance

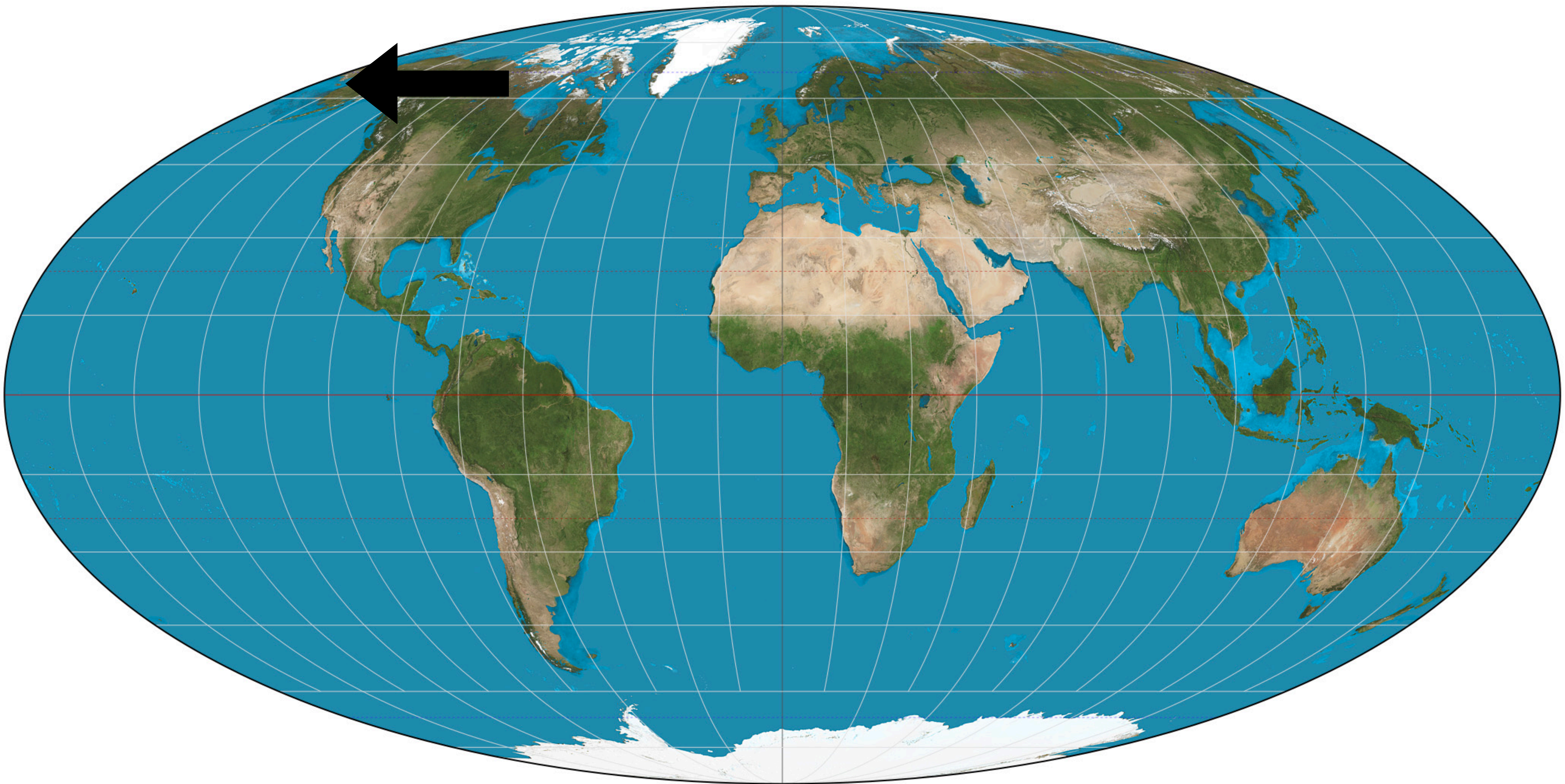
Do no harm

Centering

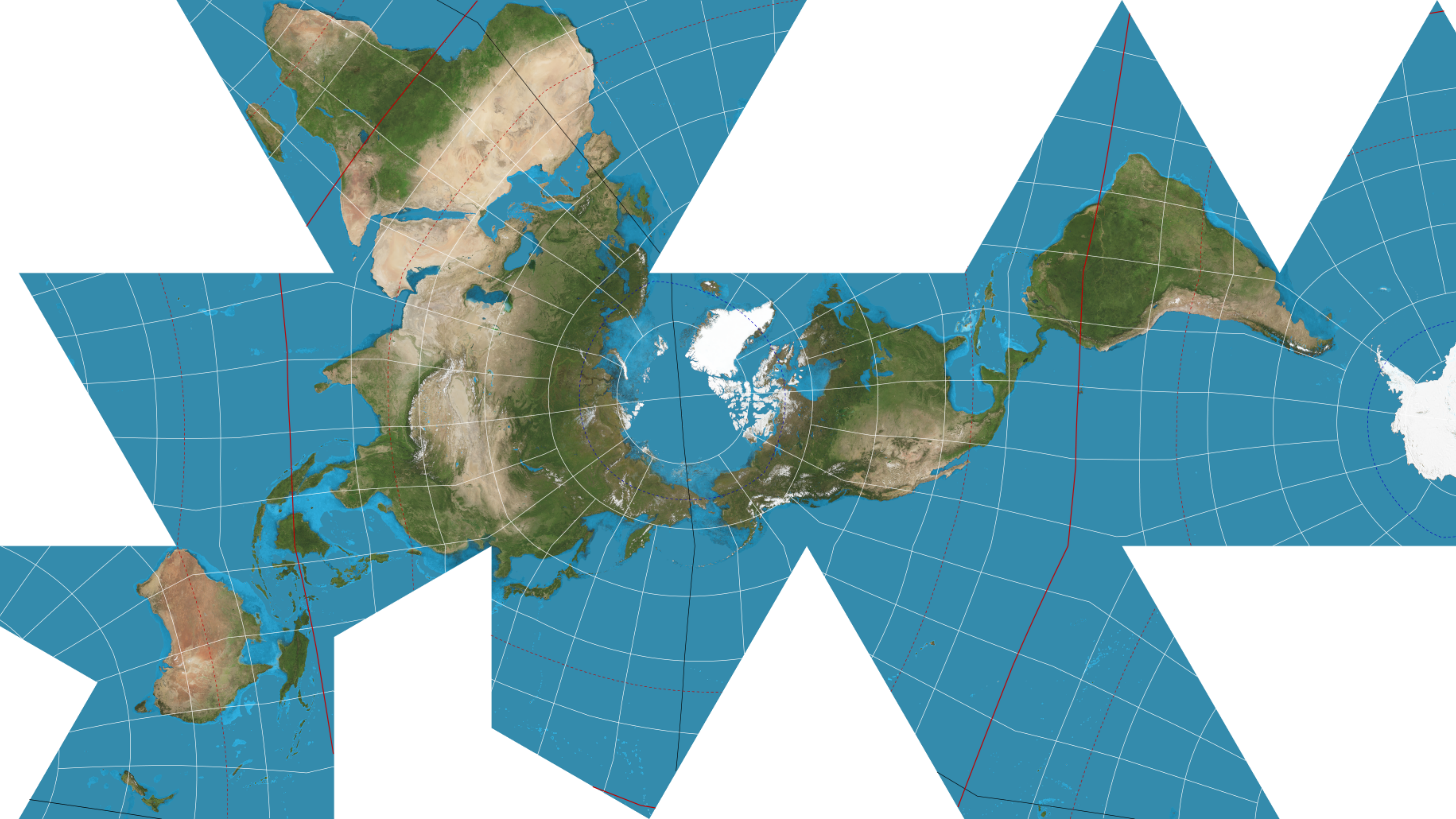


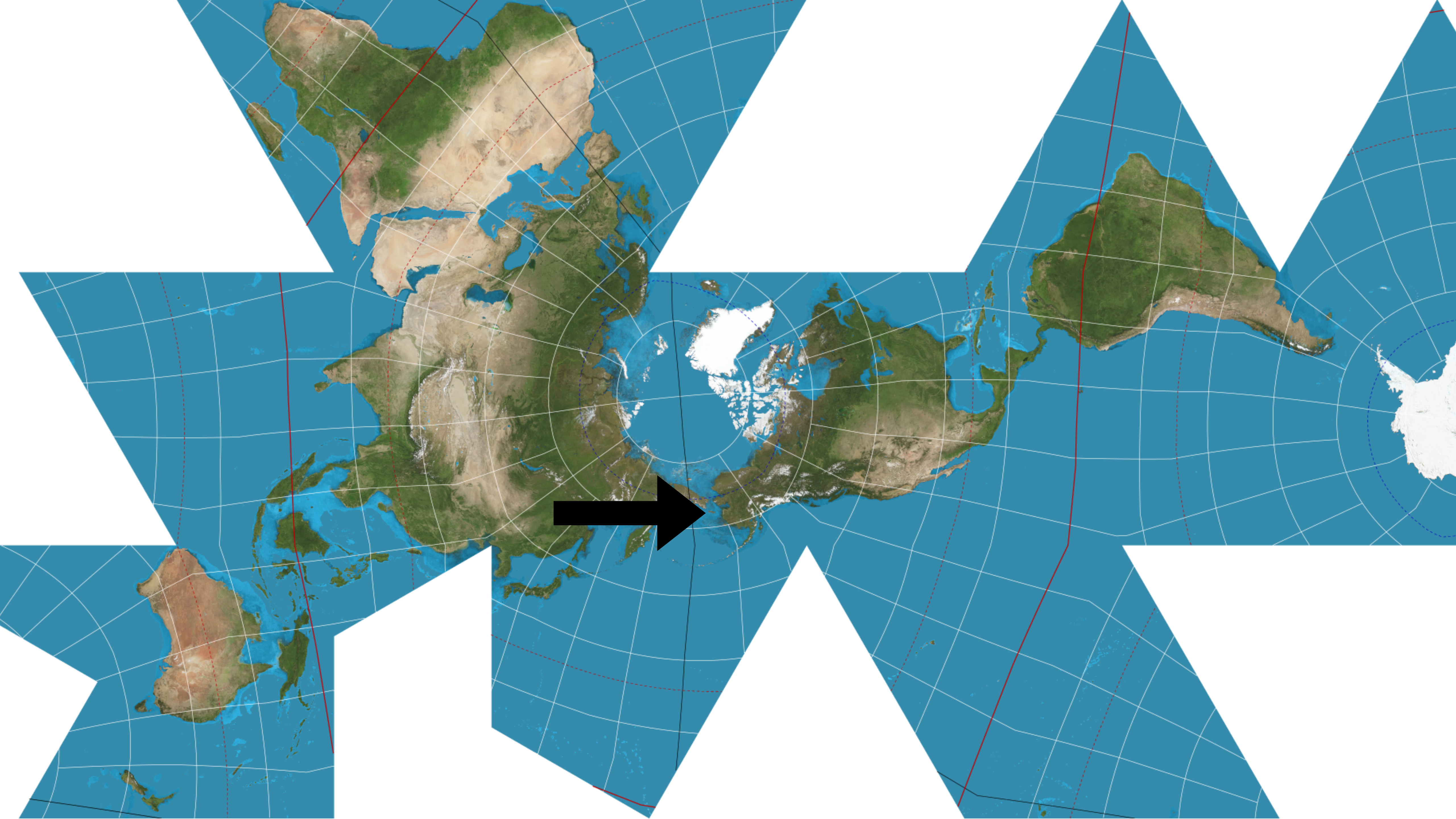






Centering





**“Always name
the languages
you work with”**

Emily Bender

ACL abstracts

- TACL (2013-2021)
- ACL, NAACL, EMNLP, EACL, AACL, JCL (2017-2021)

9602 *ACL abstracts

ACL abstracts

- TACL (2013-2021)
- ACL, NAACL, EMNLP, EACL, AACL, JCL (2017-2021)

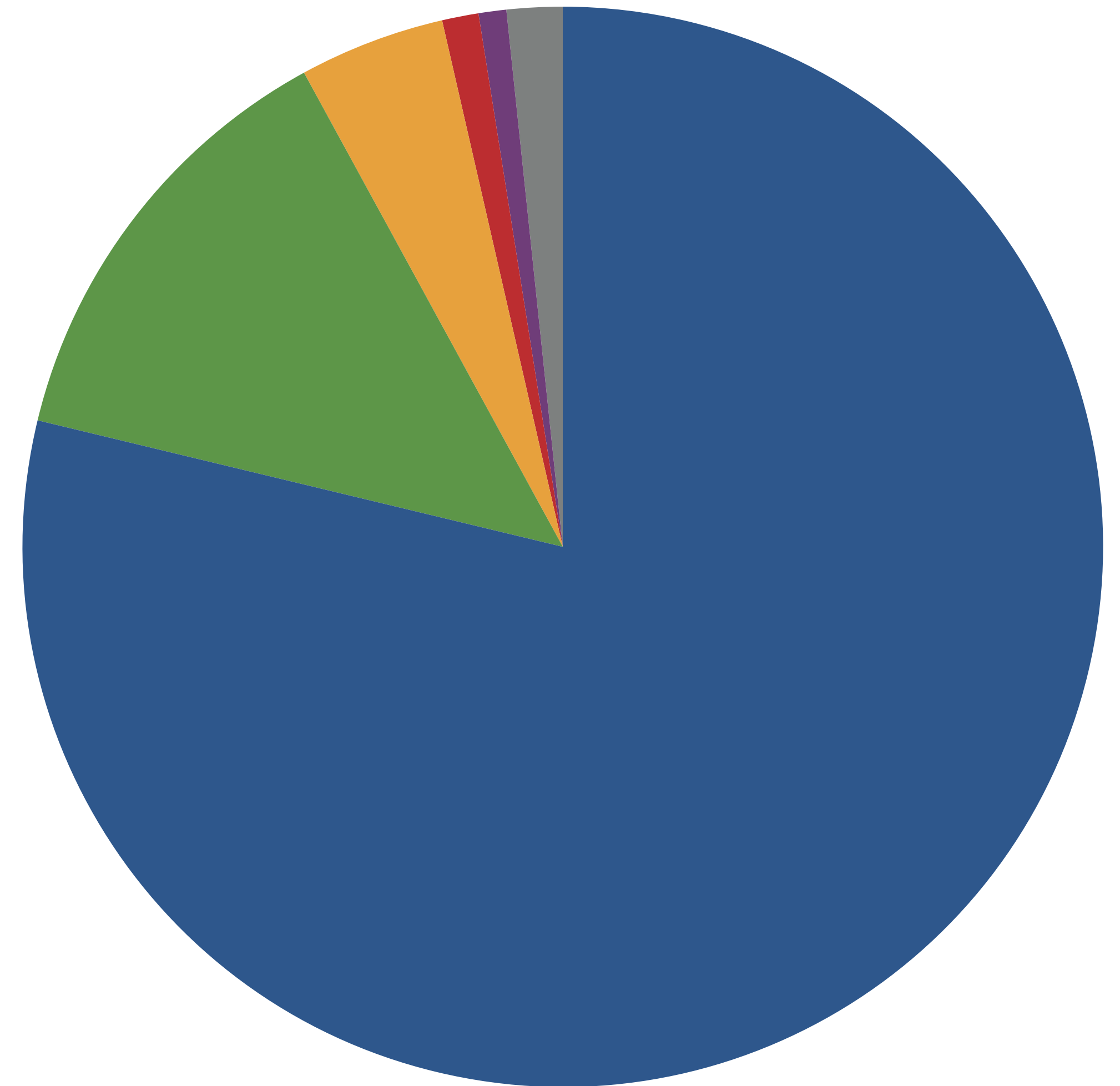
18 language families represented!

ACL abstracts



- TACL (2013-2021)
- ACL, NAACL, EMNLP, EACL, AACL, JCL (2017-2021)

9602 *ACL abstracts

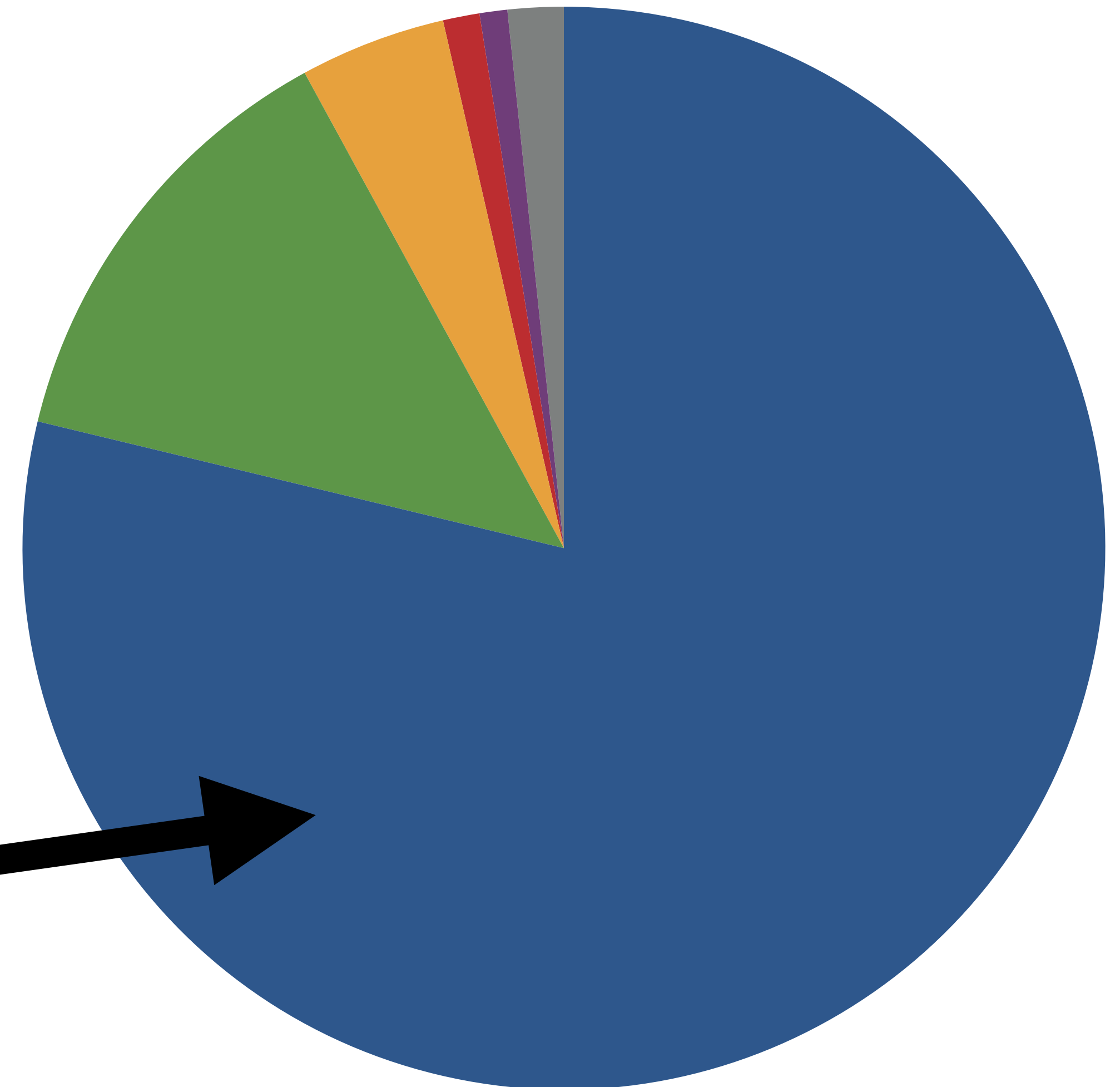


ACL abstracts



- TACL (2013-2021)
- ACL, NAACL, EMNLP, EACL, AACL, JCL (2017-2021)

**83% of *ACL papers
fail the #BenderRule**

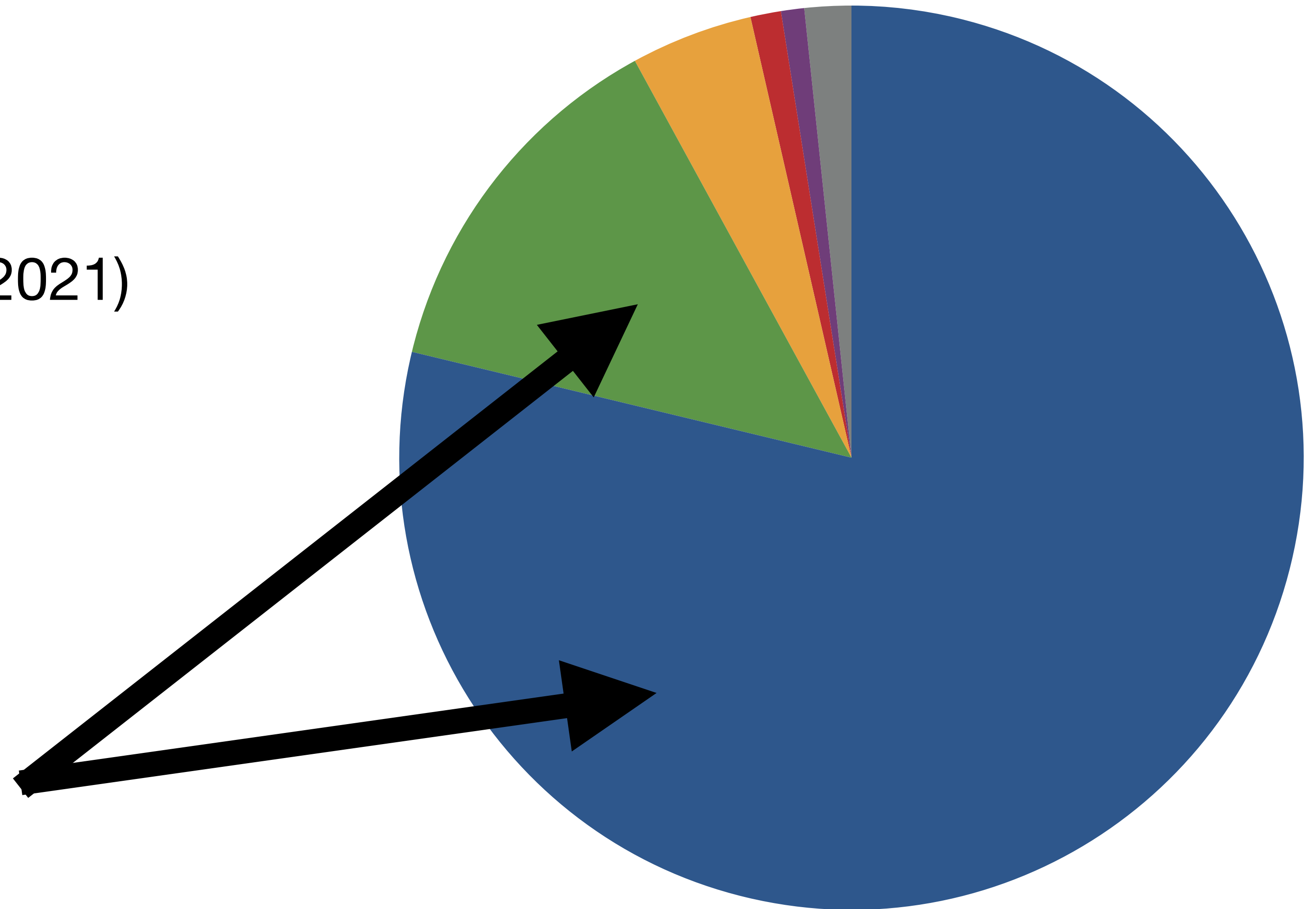


ACL abstracts

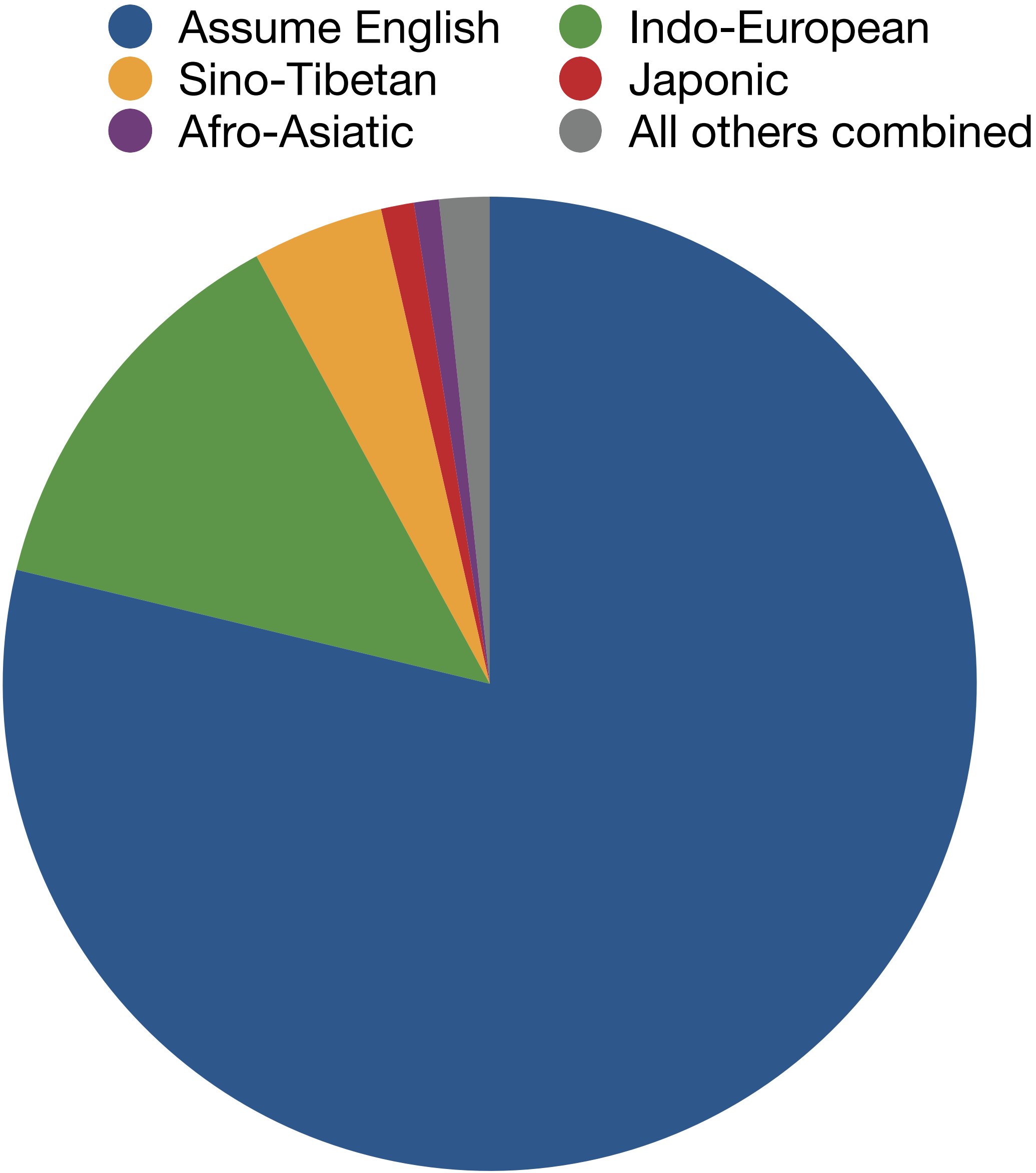


- TACL (2013-2021)
- ACL, NAACL, EMNLP, EACL, AACL, JCL (2017-2021)

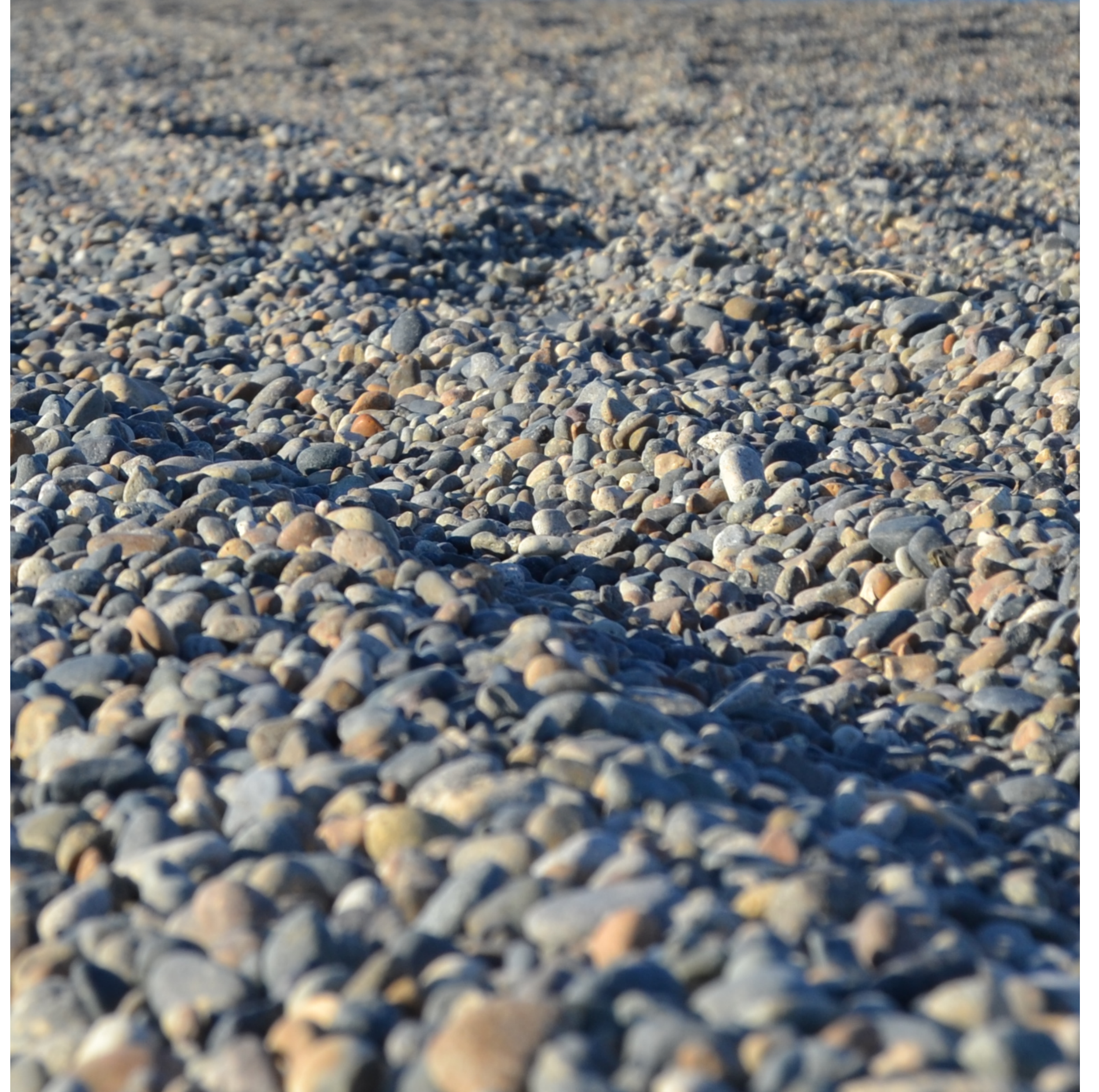
**97% of *ACL papers
Indo-European**



83.26%	7995	Implicitly assume English
13.70%	1315	Indo-European (incl. English)
4.50%	432	Sino-Tibetan
1.12%	108	Japonic
0.85%	82	Afro-Asiatic
0.41%	39	Turkic
0.26%	25	Koreanic
0.25%	24	Austroasiatic
0.24%	23	Dravidian
0.22%	21	Uralic
0.21%	20	Austronesian
0.09%	9	Basque
0.09%	9	Atlantic-Congo
0.07%	7	Na-Dene
0.05%	5	Kra-Dai
0.02%	2	Arnhem
0.02%	2	Iroquoian
0.02%	2	Inuit-Yupik-Unangan
0.01%	1	Sumerian



Homogeneity
makes for bad science



**Unquestioned
assumptions
make for unethical science**



ABCs of ethical research

Accountability

Beneficence

Cognizance

Do no harm

The ethical foundation

Do no harm

Duty to educate yourself

Cognizance

Your work must benefit the community

Beneficence

Meaningful reporting and respect of sovereignty

Accountability

ABCs of ethical research

Accountability

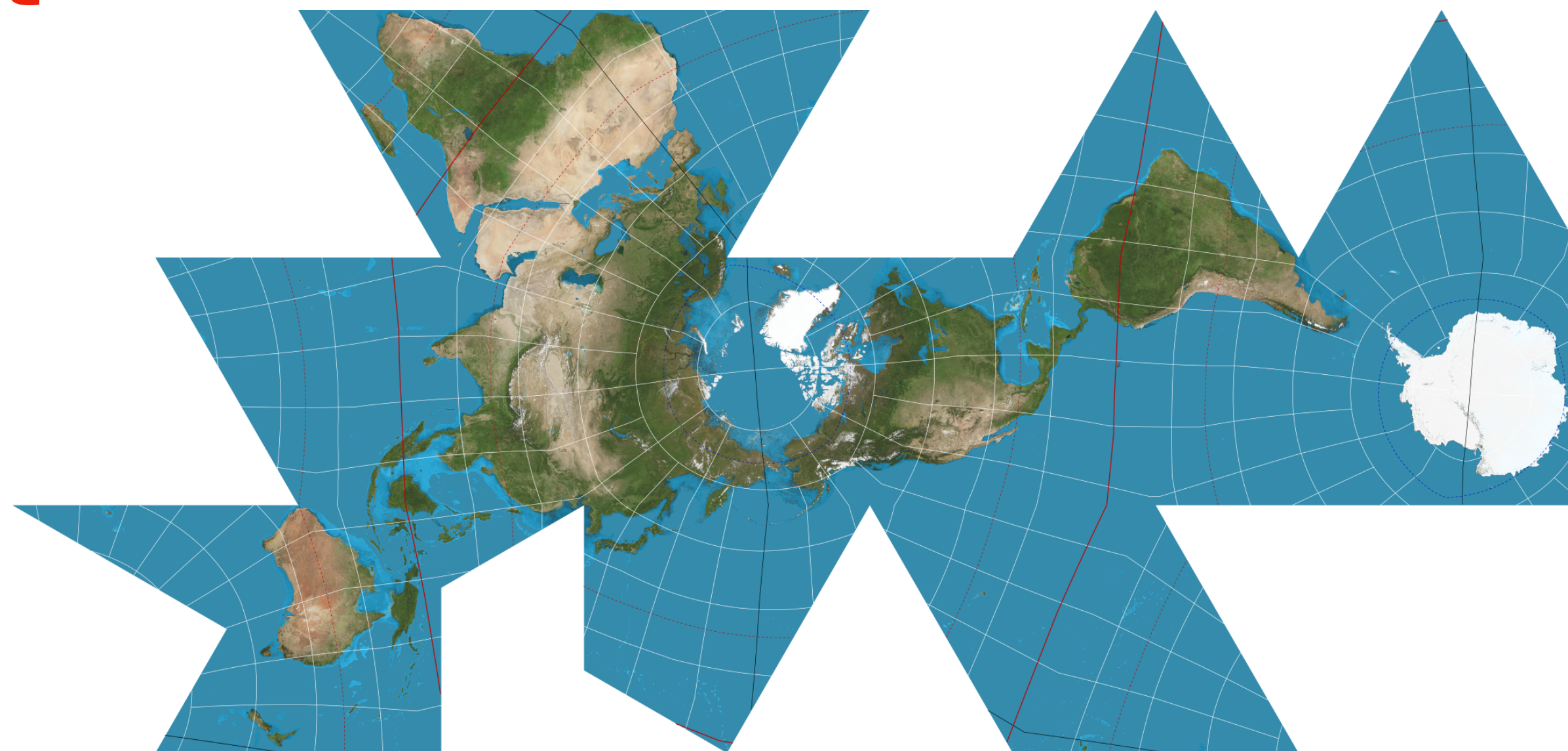
Beneficence

Cognizance

Do no harm

Respect

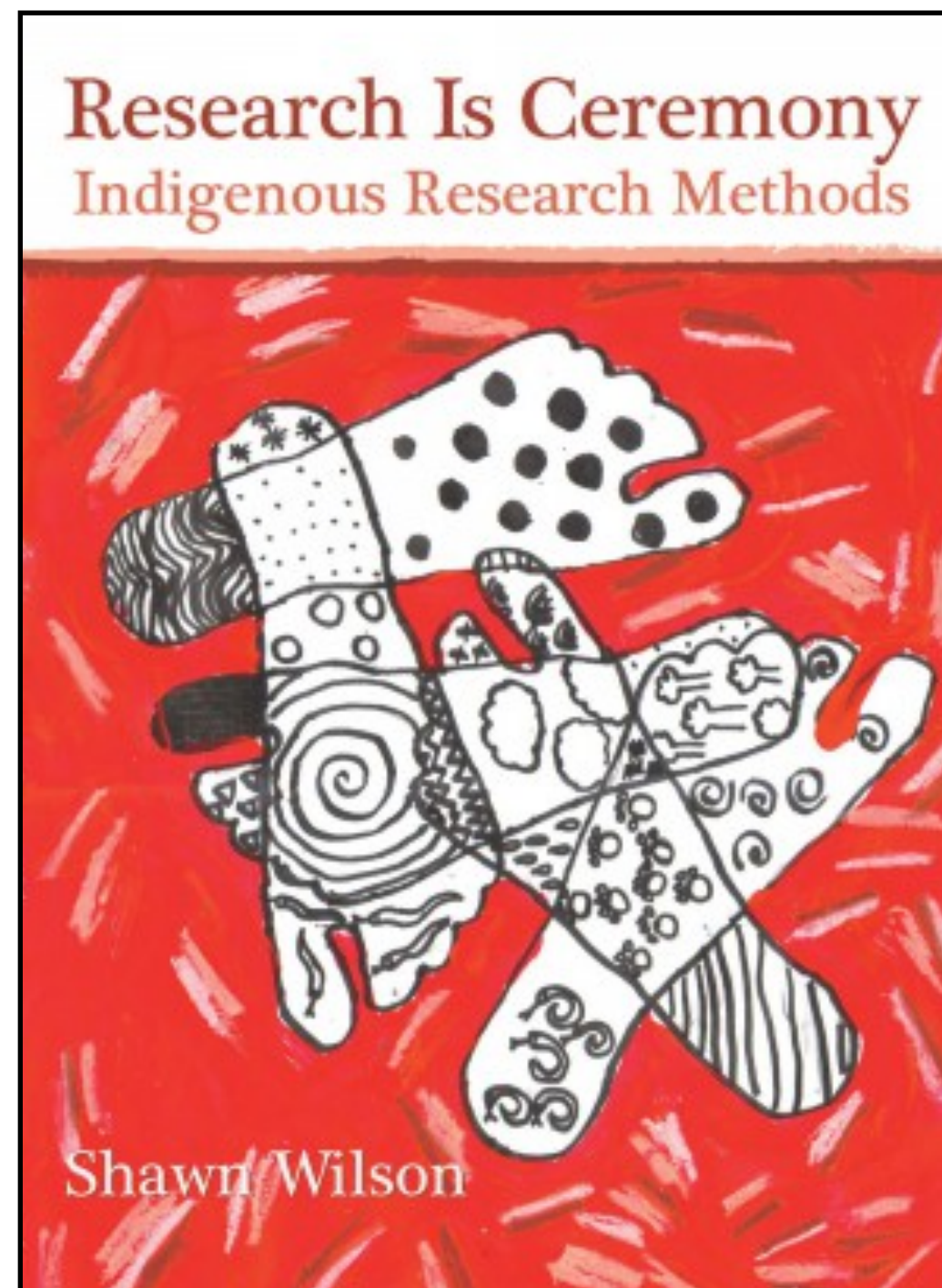
Responsibilities



Reciprocity

Relationships

Recommended readings



Decolonising Speech and Language Technology

Steven Bird
Northern Institute
Charles Darwin University



The #BenderRule:
On Naming the
Languages We Study
and Why It Matters

14.SEP.2019

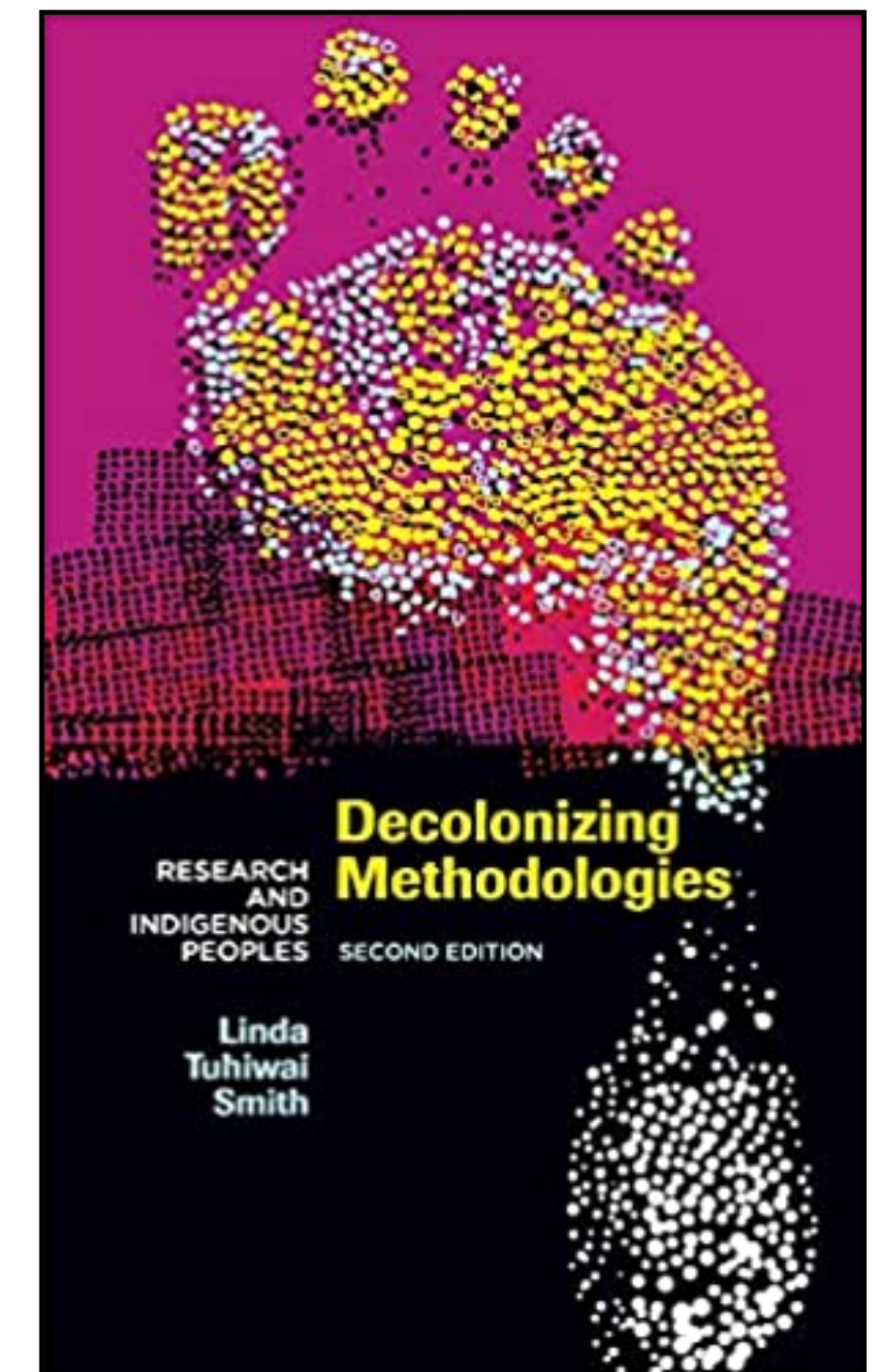


Image credits

- Frame of traditional Yupik skin boat above the west beach of Gambell, Alaska by Lane Schwartz
- Strip mining near Globe, Arizona, 1990 by Phillip Capper
- UK - London - Bloomsbury: British Museum - Elgin Marbles - Southeast metope by Wally Gobetz
- The Parthenon in Athens by Steve Swayne
- Buckminster Fuller's Dymaxion map by Justin Kunimune
- Mercator and Tobler hyperelliptical map projections by Daniel R. Strebe

Primum Non Nocere:

Before working with Indigenous data,
the ACL must confront ongoing colonialism

Lane Schwartz

