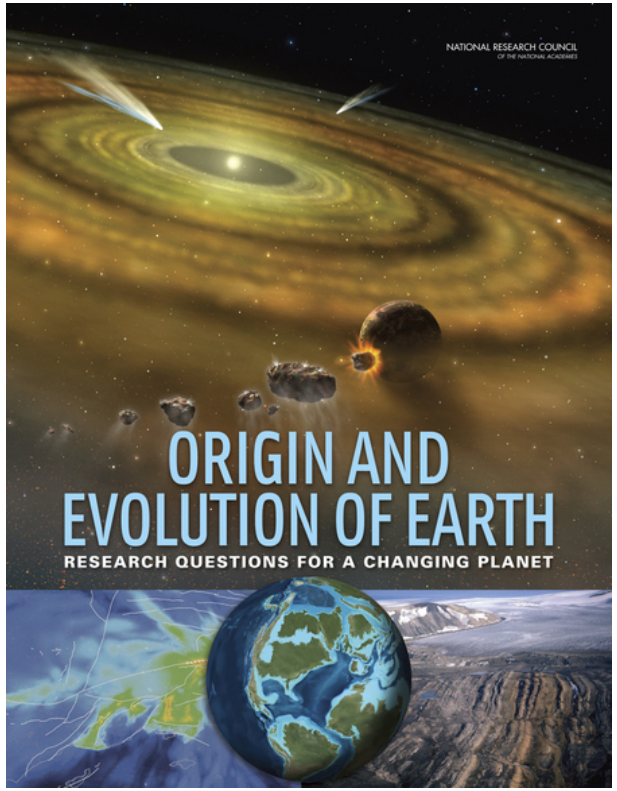
deeptimekb.org

A Knowledge Base of Deep Time to Assimilate Multi-disciplinary Datasets in the Study of Co-evolving Geosphere and Biosphere

Xiaogang (Marshall) Ma
Assistant Professor, Computer Science
Affiliate Faculty Member, Geology
University of Idaho, Moscow, ID, USA

Email: max@uidaho.edu | Twitter: @dtkb11

Co-evolution



ORIGINS

Question 1: How did Earth and other planets form? 7

Question 2: What happened during Earth's "dark age" (the first 500 million years)? 18

Question 3: How did life begin? 27

EARTH'S INTERIOR

Question 4: How does Earth's interior work, and how does it affect the surface? 35

Question 5: Why does Earth have plate tectonics and continents? 50

Question 6: How are Earth processes controlled by material properties? 60

A HABITABLE PLANET

Question 7: What causes climate to change—and how much can it change? 71

Question 8: How has life shaped Earth—and how has Earth shaped life? 84

HAZARDS AND RESOURCES

Question 9: Can earthquakes, volcanic eruptions, and their consequences be predicted? 95

Question 10: How do fluid flow and transport affect the human environment? 111

National Research Council, 2008
https://www.nap.edu/catalog/12161/origin-and-evolution-of-earth-research-questions-for-a-changing

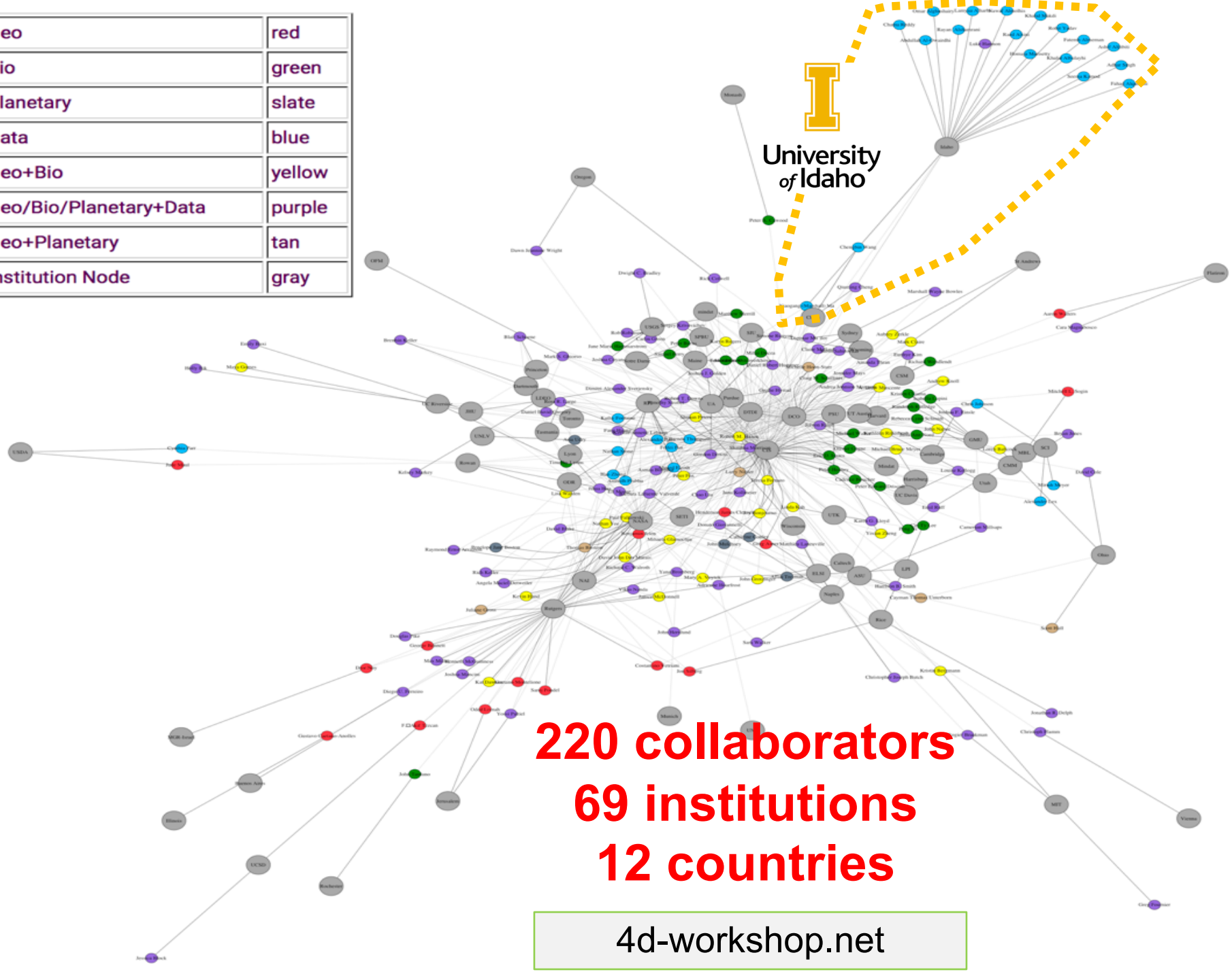
4D
Deep-time Data Driven Discovery

Geo	red
Bio	green
Planetary	slate
Data	blue
Geo+Bio	yellow
Geo/Bio/Planetary+Data	purple
Geo+Planetary	tan
Institution Node	gray

University of Idaho

2018-Present

The 4D Initiative: Deep-time Data-Driven Discovery



220 collaborators
69 institutions
12 countries

4d-workshop.net

Diagram from: Bob Hazen

The challenge: Abundant but heterogeneous data sources

- Paleontology
  - Paleobiology Database, Fossil Works
- Igneous
  - GeoKem, NAVDAT, NOAA Petrology Database
- Metamorphic
  - Metamorphic Petrology Database, Melts
- Sedimentology
- Petrology
  - NOAA Petrology Database, EarthChem, Metamorphic Petrology Database, Crystallography Database
- Mineralogy
  - MinDat, RRUFF
- Biology/Protein
  - Paleobiology Database, Protein Database, VAMPS
- Volcanology
  - NAVDAT, Smithsonian Institute Volcano Database
- Trace/Minor Elements
  - Database of Trace Element Utilization
- GeoChemistry
  - GeoKem, NOAA Petrology Database, Melts, EarthRef, GeoRoc, EarthChem, LEPR, ThermoCalc, Calcium Carbonate Database
- GeoMagnetism
  - MagIC

A full list: https://tw.rpi.edu//web/project/DTDI/WorkingGroups/dtdidatabases

Time can be used as a common reference in geoscience data integration

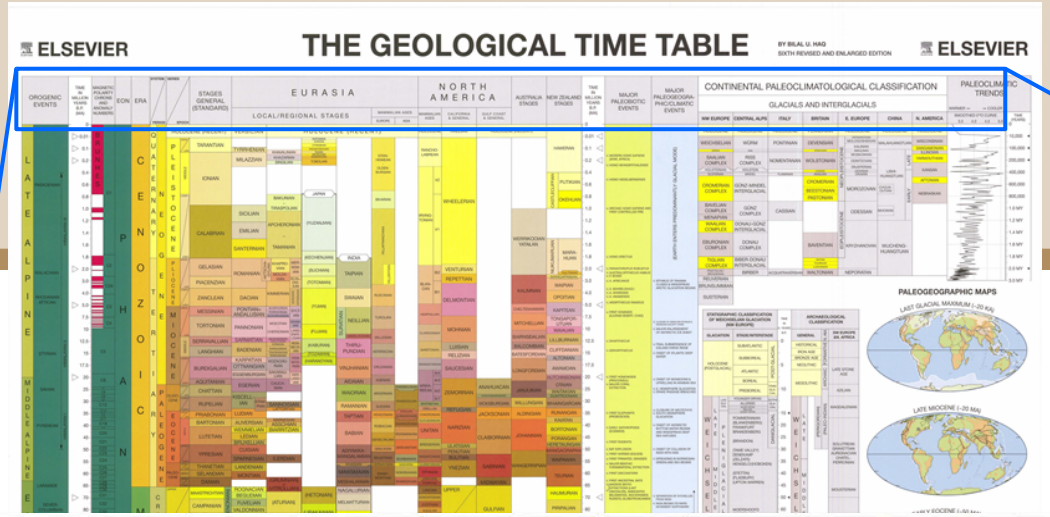
- Time, location, and redox state are three key subjects that can be applied as axes to assimilate those datasets
- Then we can work on focused scientific questions:
  - What was the geological and geochemical context for life's origins?
  - Did biological catalysis follow from the chemistry of rocks and minerals?
  - How did plate tectonics begin and to what extent are Earth's surface and deep interior linked?
  - When, and at what rate, did photosynthesis modify our planet's near-surface environment?
  - More

Question list courtesy of:
Bob Hazen and 4D group

- But…

Heterogeneous geologic time terminology in deep time open data

Reginal and national geologic time standards and their usage in various databases



OROGENIC EVENTS	TIME IN MILLION YEARS B.P. (MA)	MAGNETIC POLARITY CHRONS AND ANOMALY NUMBERS	EON	ERA	SYSTEM / PERIOD	SERIES / EPOCH	STAGES GENERAL (STANDARD)	EURASIA			NORTH AMERICA		
								LOCAL/REGIONAL STAGES	MAMMALIAN AGES		MAMMALIAN AGES	CALIFORNIA & GENERAL	GULF COAST & GENERAL
									EUROPE	ASIA			

AUSTRALIA STAGES	NEW ZEALAND STAGES	TIME IN MILLION YEARS B.P. (MA)	MAJOR PALEOBIOTIC EVENTS	MAJOR PALEOGEOGRA-PHIC/CLIMATIC EVENTS	CONTINENTAL PALEOCLIMATOLOGICAL CLASSIFICATION							PALEOCLIM TRENDS
					GLACIALS AND INTERGLACIALS							WARMER ← → COOLER
					NW EUROPE	CENTRAL ALPS	ITALY	BRITAIN	E. EUROPE	CHINA	N. AMERICA	SMOOTHED $\delta^{18}O$ CURVE 3.5 4.0 4.5 5.0

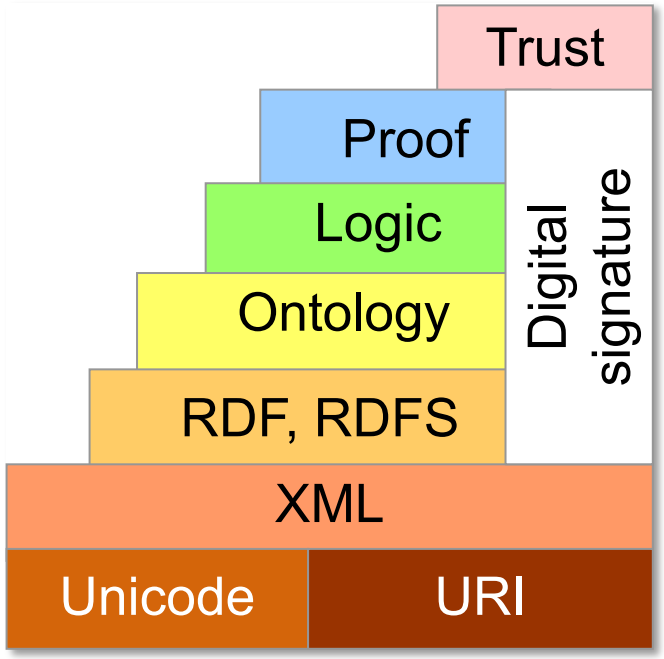
(Haq, 2007)

Data Interoperability vs. Data Heterogeneity

Background image courtesy of: British Geological Survey

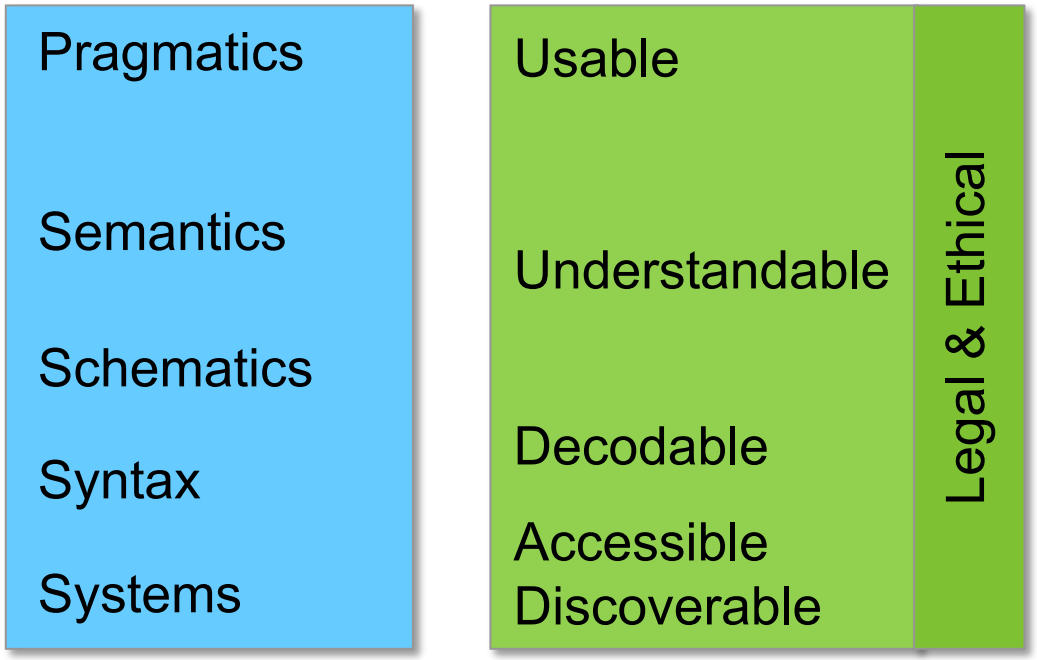
Interoperability vs. Heterogeneity

Semantic Web



(Berners-Lee, 2000)

Data Interoperability

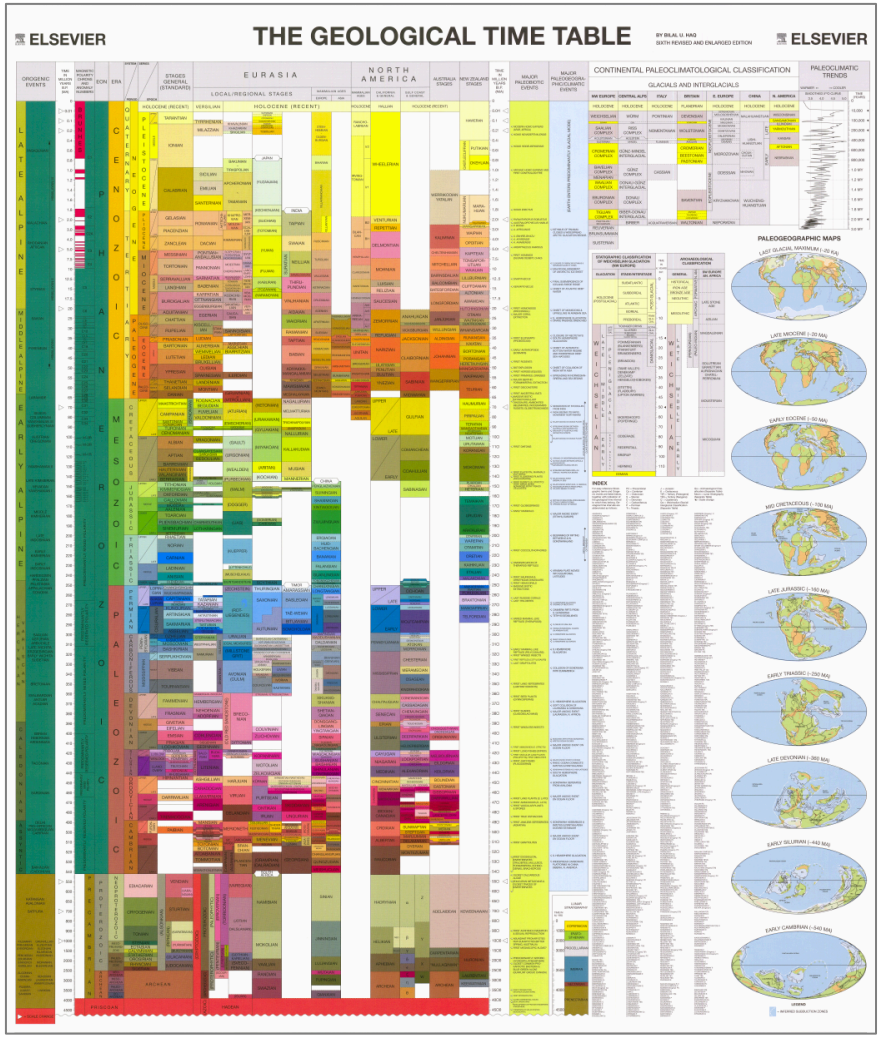


(Bishr, 1998; Sheth, 1999; Ludascher et al., 2003; Brodaric, 2007)

(Wood et al., 2010; Ma et al., 2011; 2014)

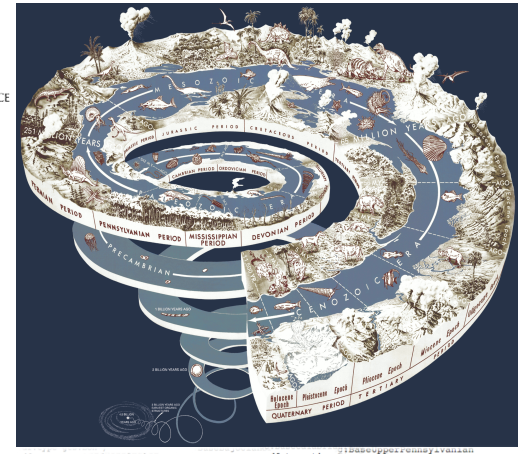
Our focus: A knowledge base of deep time to facilitate data harmonization

- NSF funded, #1835717, 12/2018 to 11/2021
- Machine-readable: Semantic modeling and encoding
- Interoperability: Connect heterogeneous time concepts among disciplines and standards
- Data Revolution: Leverage existing data resources and facilitate reproducible workflows



(Haq, 2007)

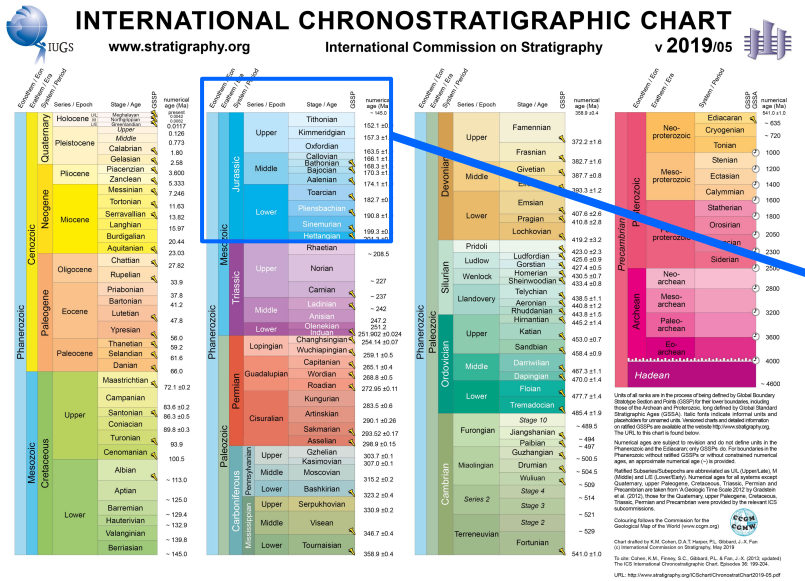
- Collaborating communities
- Potential data sources for case studies

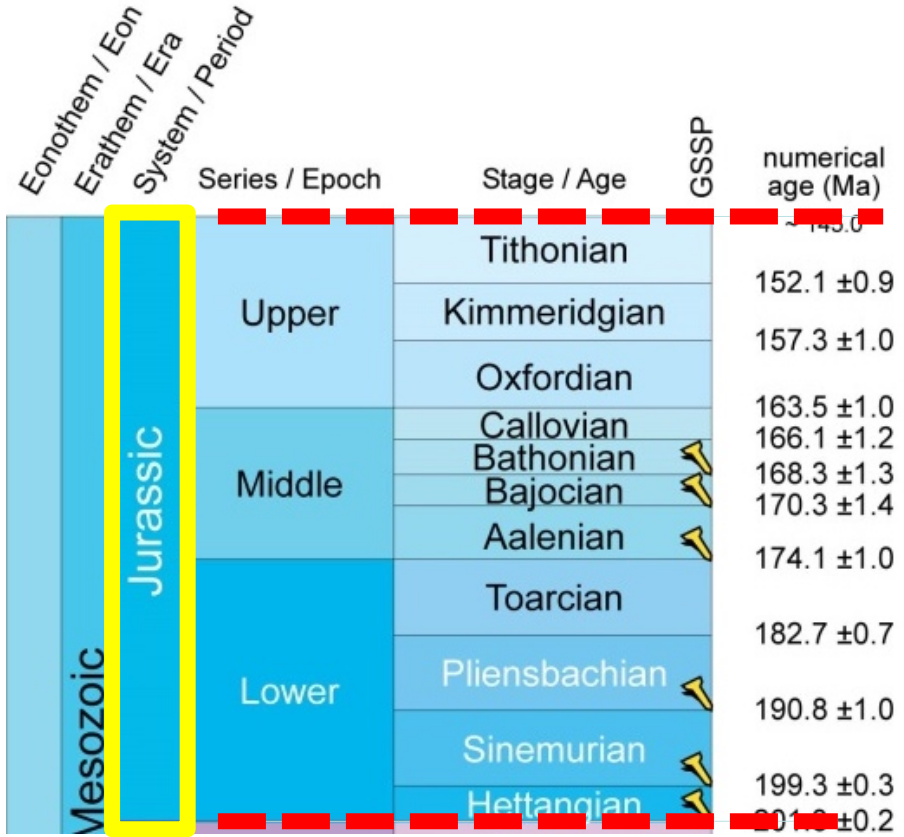


(Background image from: USGS)

How computer scientists see the geologic time scale

- Two key concepts
  - **Interval (Unit)**: a period of time between two events
  - **Instant (Boundary)**: a particular point in time





We can see Jurassic as an **Interval** and its start and end time (base and top boundary) each as an **Instant**

Raw dataset:
- RRUFF MED
- IMA mineral name list

Under development: Paleogeographic map of mineral occurrence

ages: min 252.2

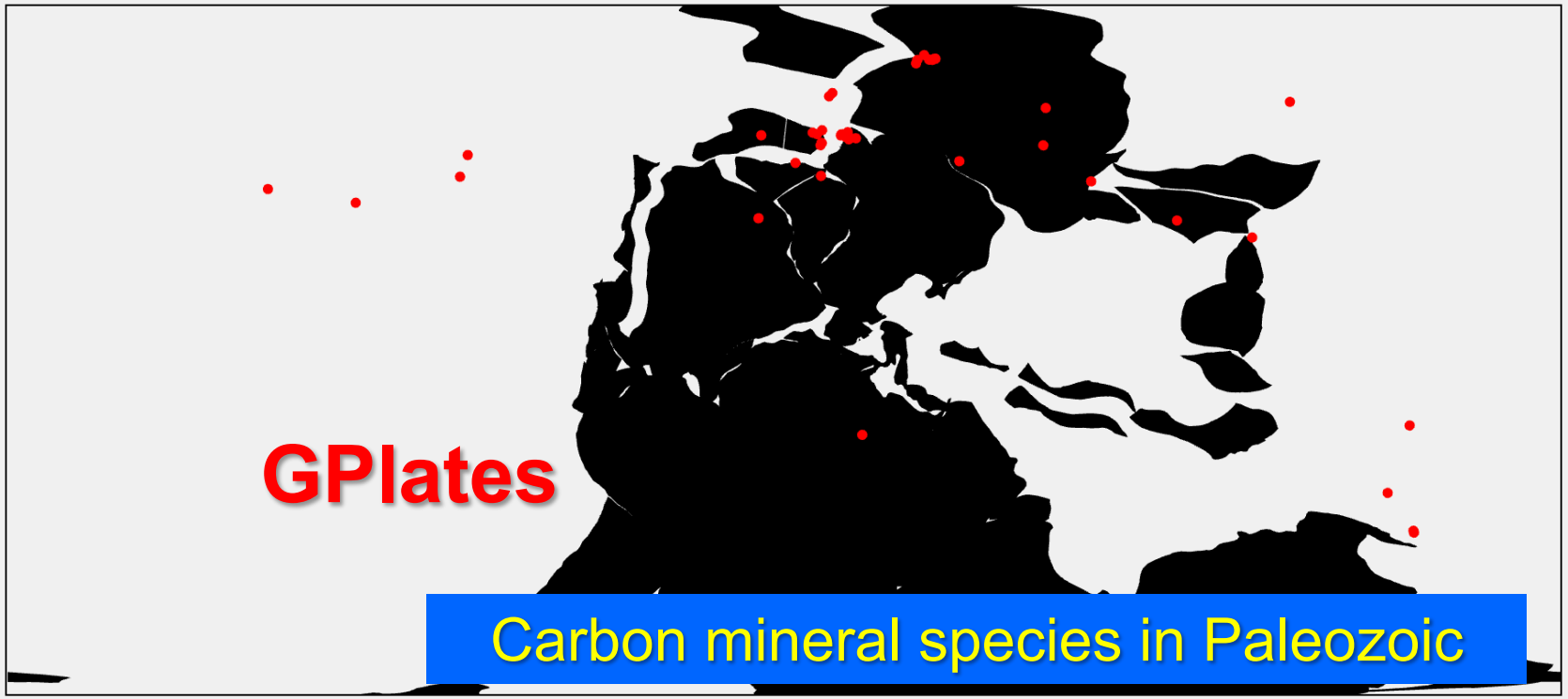
ages: max 541

RRUFF

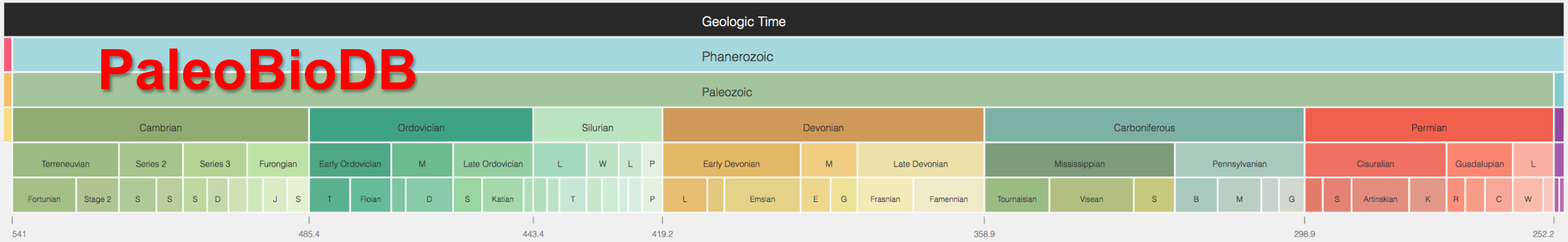
H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	REE	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	**	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	Uuq	UuP	Uuh	Uus	Uuo
			La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
			Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

Selected: C

Submit



PaleoBioDB



Conclusions

- **Why we need a machine-readable knowledge base of deep time**
  - Wider data connections across disciplines
  - Faster relationship analysis and inference
  - Help data integration and exploration

- **How to build it**
  - Semantic modeling and encoding
  - Machine-readable knowledge graph
  - A service to support data assimilation

Acknowledgements




Chao Luke Jenn Xiang

Our collaborators at many institutions

Thank You

More Information

deeptimekb.org