

A PAGES Floods WG core project:



UNI
FREIBURG



PAGES OSM 2017
5th Open Science Meeting

Collaborative Flood Database for Multiple Archive Types

Albert-Ludwigs-Universität Freiburg

11-05-2017

OSM Pages, Zaragoza

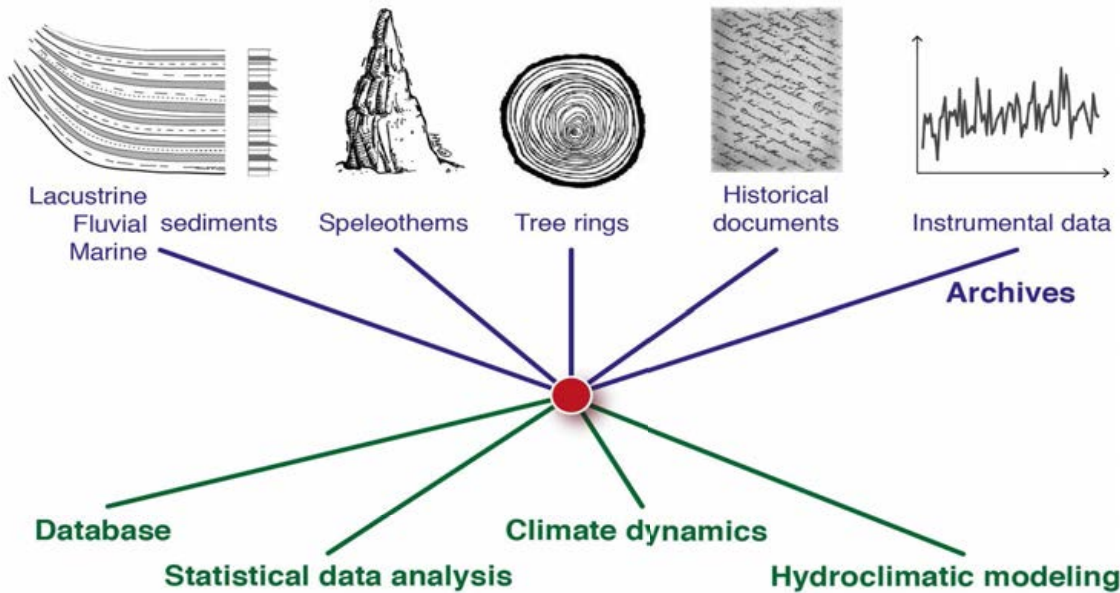
Michael Kahle (1), Rüdiger Glaser (1), Pierre Francus (2), Pages Flood WG (3)
University Freiburg (1), Institut National de la Recherche Scientifique (2),
<http://www.pages-igbp.org/ini/wg/floods/people> (3)

- From the Floods WG page:
- “Compile published data sets on floods for **open-access archiving** in order to facilitate the **visibility** of existing data and their **inter-comparison**”
- (<http://www.pages-igbp.org/ini/wg/floods/scientific-goals> -> iii)
- Necessary: Common Data Structure across Multiple Proxy Types

If possible, use Sensor-Archive-Observation concept

- (Evans et al 2013 : <https://doi.org/10.1016/j.quascirev.2013.05.024>)

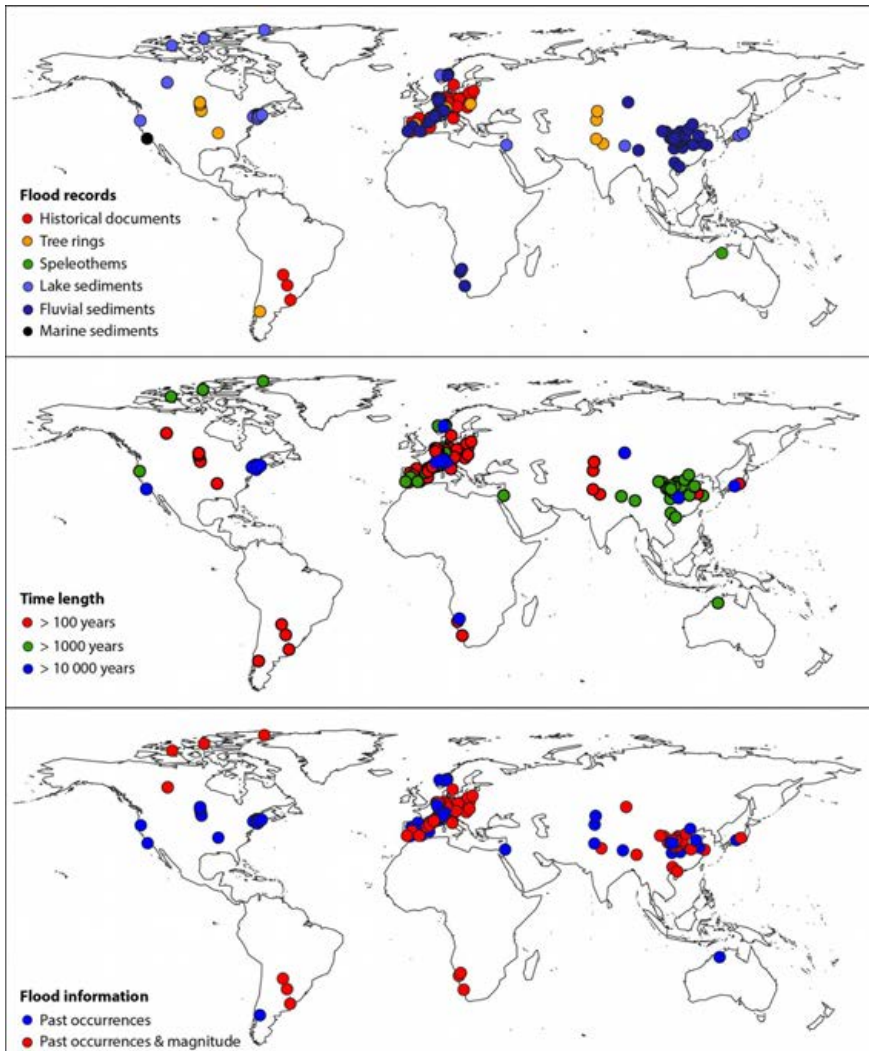
Multiple Proxy Types



- Sediments
 - Lacustrine
 - Fluvial
 - Marine
- Speleothems
- Tree rings
- Historical documents
- Instrumental data
- ...

Source: <http://www.pages-igbp.org/ini/wg/floods/intro>

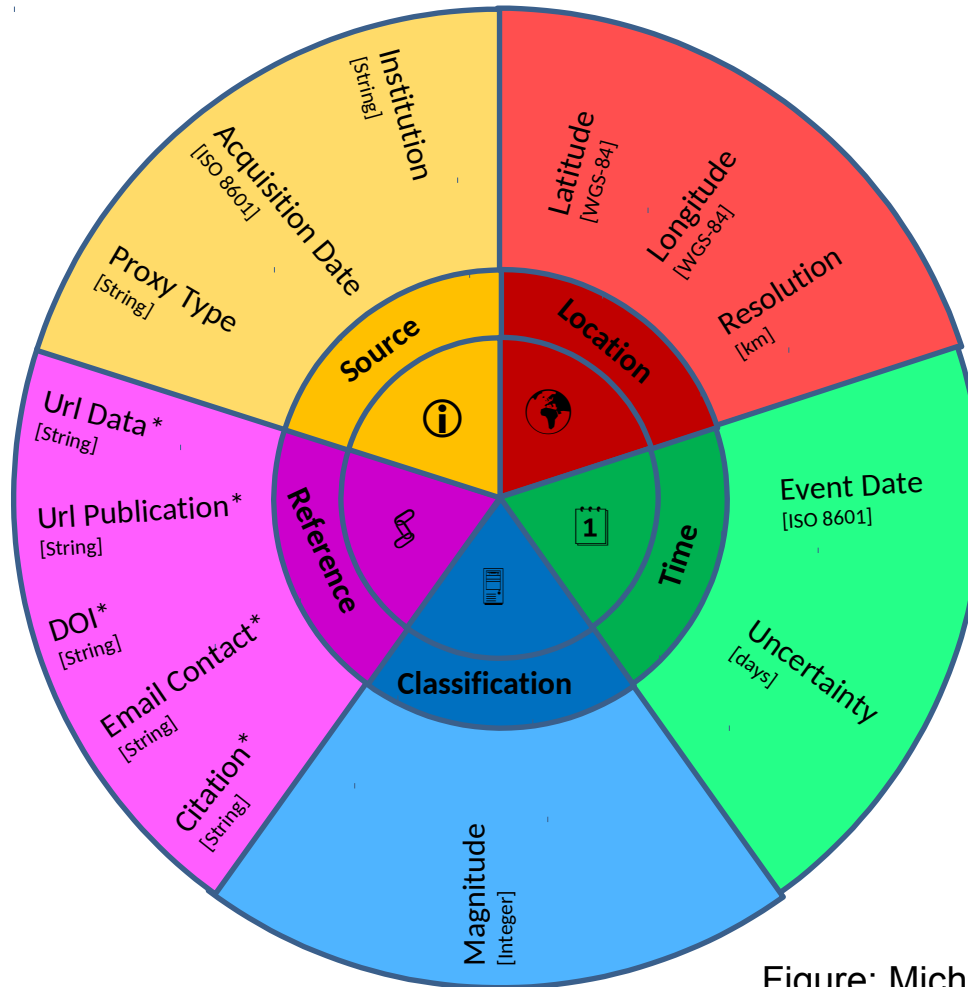
Meta Data Collection



- Lots of candidates for a common flood database across all types of archives, countries and time ranges
- Thanks to all contributors

Figure: Bruno Wilhelm, 2017

Minimal Data Structure



Clusters






-  : Source
-  : Location
-  : Time
-  : Classification
-  : Reference

Figure: Michael Kahle 2017

Multiple Proxy Types

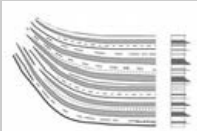



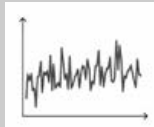


Differ by

- Time range covered (years back, seasons)
- Temporal resolution
- Location types covered
- Sensibility to floods (by extension or duration)
- Negative signals (droughts)
- Noise
- Additional information (causes, impacts, temperature, ...)

Multiple Proxy Types

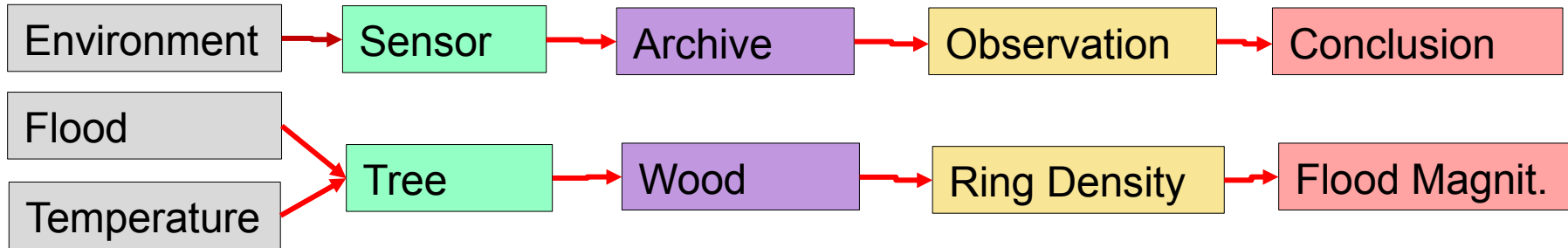


Proxy		Location	Time range	Time resol.	Misc
Sediments		River, Lake	10000s of years	1 year	
Speleothems		Carst, Cave	1000s of years	1 year	Hydrol. Data, e.g. Discharge
Tree rings		Woods	500-1000 years	1 year, seasons	Precipitation
Historical documents		City, Village	500-1000 years	1 day, hour	Causes, Impacts
Instrumental measurements		Civilization	100s of years (1000s)	1 day, hour	Precise Magnitude

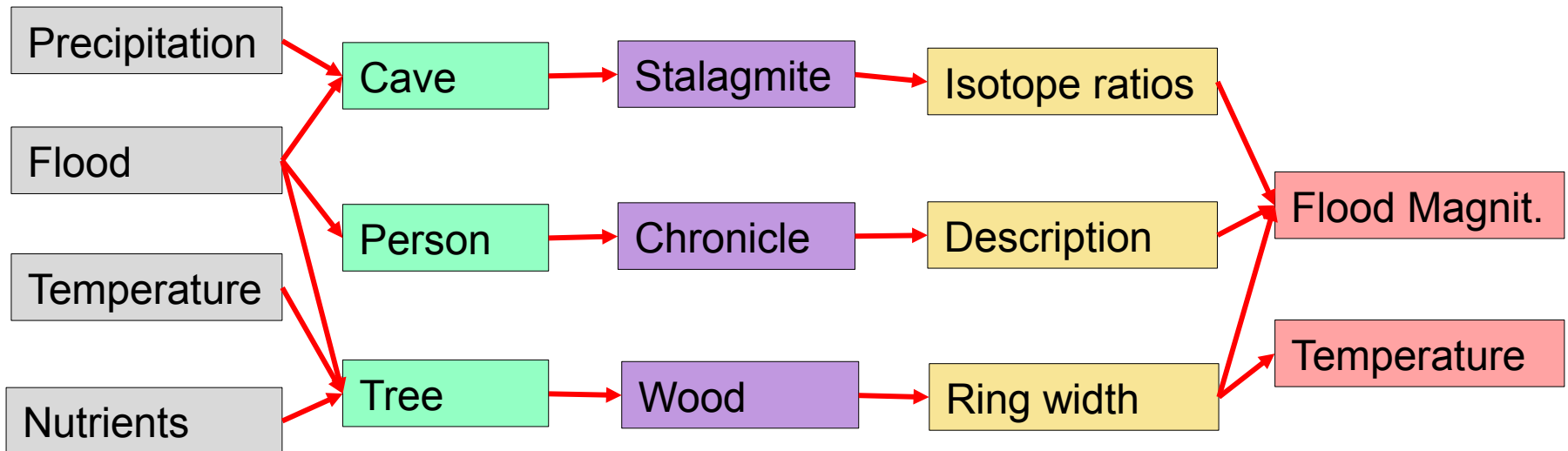
Sensor, Archive, Observation



Single archive approach (by [Evans et al 2013](#)):



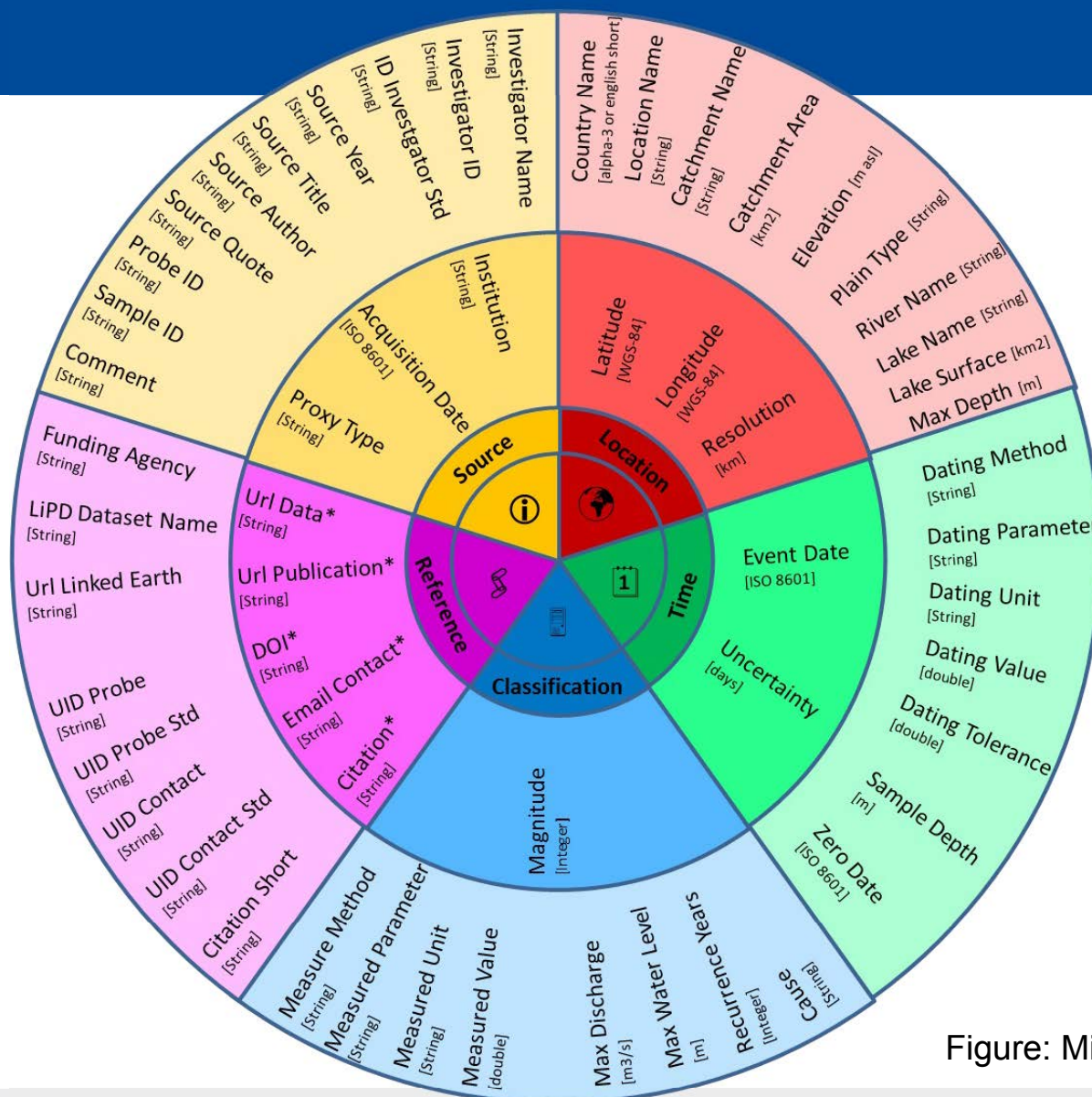
Multiple archive approach:



- To use Sensor-Archive-Observation model more data is necessary
- Also documentation of whole process leading to flood data is desirable

- Example data was provided by members of the Floods WG
- → More (partially optional) data fields are needed.

Common Data Structure

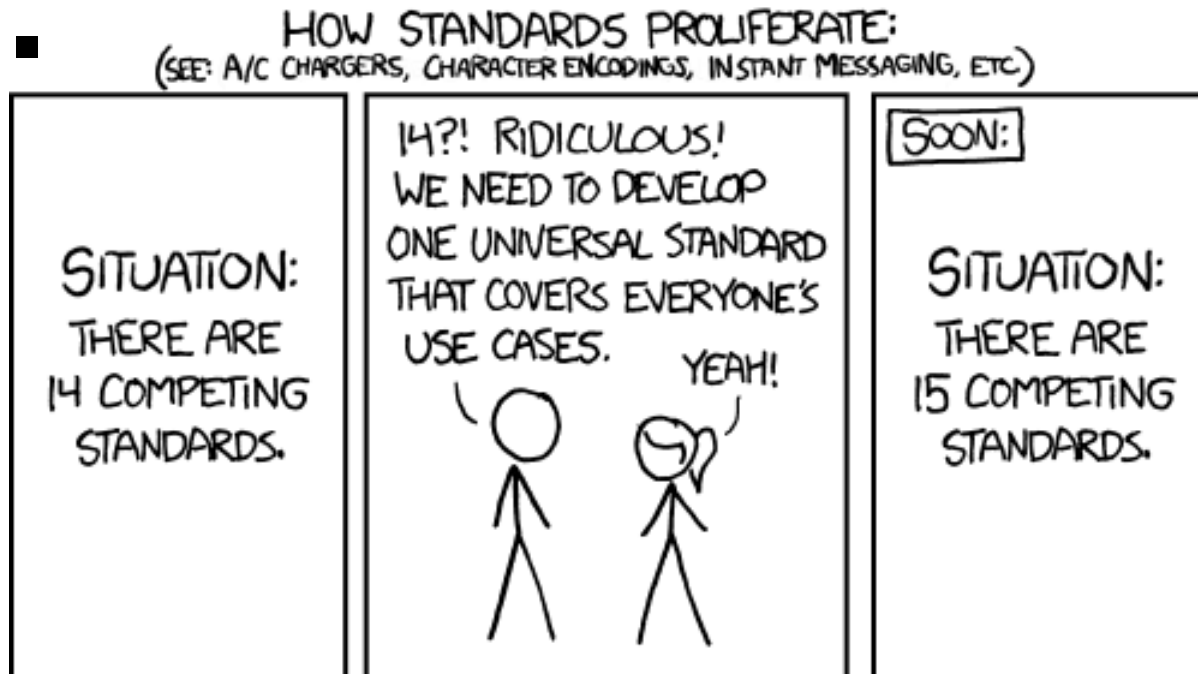


Clusters

- : Source
- : Location
- : Time
- : Classification
- : Reference

Figure: Michael Kahle 2017

- Human readable
 - Easy to create and edit
- Machine readable
 - Import and export to and from software tools
- Check for completeness and consistency
- Changes traceable
- → Avoid binary format



<https://xkcd.com/927/>

Data Format: CSV



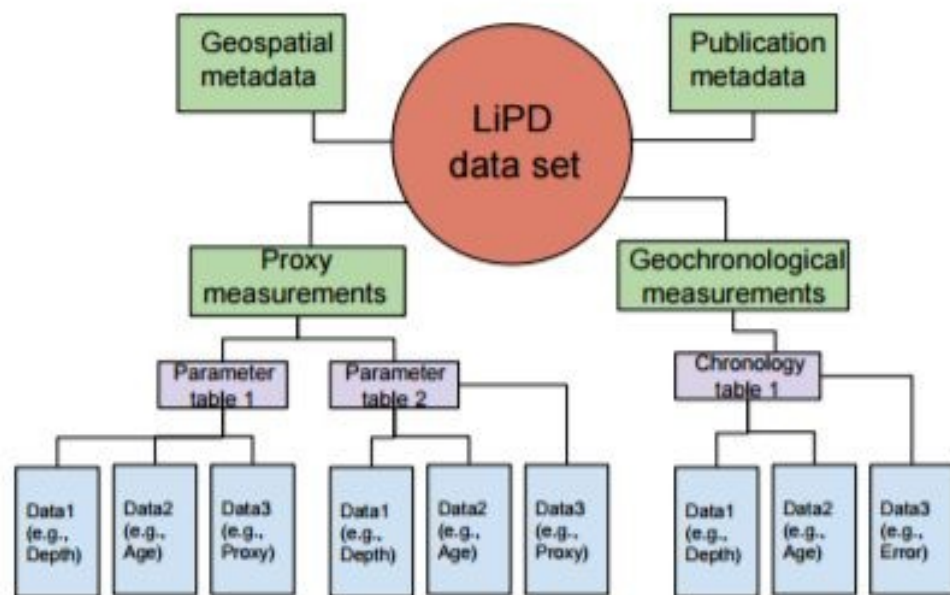
Mandatory fields are compact

A	B	C	D	E	F	G	H	I	J	K	L	M	N
Institution	Proxy Type	Acquisition Date	Latitude	Longitude	Resolution	Event Date	Uncertainty	Magnitude	Url Data	Url Publicatio DOI	Email Contac	Citation	
University,	fluvial	2009-02-24	51,8619	6,1186	100	1926	180	3			w.h.j.toonen@	Toonen, W.H.	

BUT – optional data is huge

O	P	Q	R	S	T	U	V	W	X
Investigator N	Investigator L	UID Investigat	Source Year	Source Title	Source Autho	Source Quote	Probe ID	Sample ID	Comment
Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
Country Name	Location Name	Catchment Name	Catchment Area	Elevation	Plain Type	River Name	Lake Name	Lake Surface	Max Depth
AI	AJ	AK	AL	AM	AN	AO			
Dating Metho	Dating Param	Dating Unit	Dating Value	Dating Tolera	Sample Depth	Zero Date			
AP	AQ	AR	AS	AT	AU	AV	AW	AX	
Measure Met	Measured Par	Measured Uni	Measured Val	Measured Tol	Max Discharge	Max Water Le	Recurrence Yr	Cause	
AY	AZ	BA	BB	BC	BD	BE	BF		
Laser Diffract	Coar	Funding Agenc	LPD Dataset I	Url Linked Ear	UID Probe	UID Probe Std	UID Contact	UID Contact S	Citation Short

Mixture of json and csv files in zip



<https://doi.org/10.5194/cp-12-1093-2016>

N. P. McKay and J. Emile-Geay

Advantage

- Covers: Sediments, Speleothems, Tree rings, Measurements on fixed points, ...
- Well established file format with tools available
- Fine structured

BUT - Enhancements needed for

- Historical documents
- Standardize column names & Units

Proxy Archives vs Documents



Proxy Archives	Historical Documents
One Probe on fixed location	One source often covers multiple locations, multiple observers
Mostly annually, periodically	Daily or even hourly resolution, randomly scattered samples
Time derived (counting, isotopes)	Time absolutely (calendar, clock)
Continuously recorded	Often only extremes recorded
May interfere with circumstances	Can distinguish circumstances
Causes and Impacts to guess	Causes & Impacts often mentioned

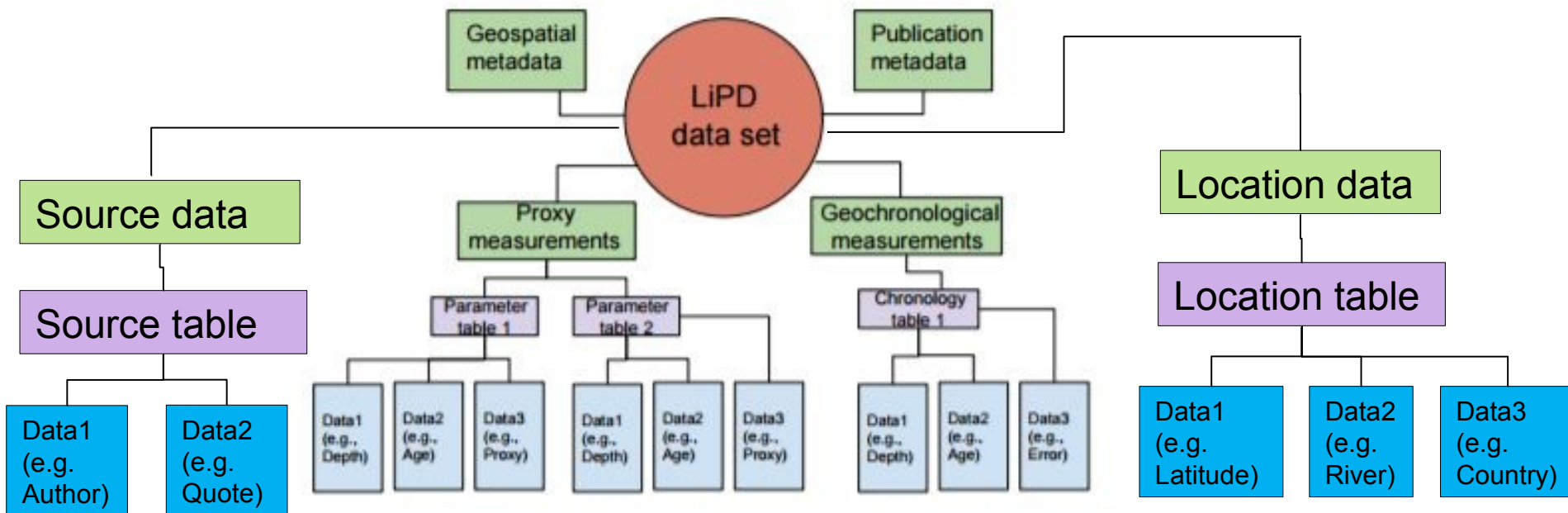
LiPD – Documents Enhancements (1)



1st Possibility:

Extra table for

- Source data
- Location data (refines Geospatial)



Based on N. P. McKay and J. Emile-Geay

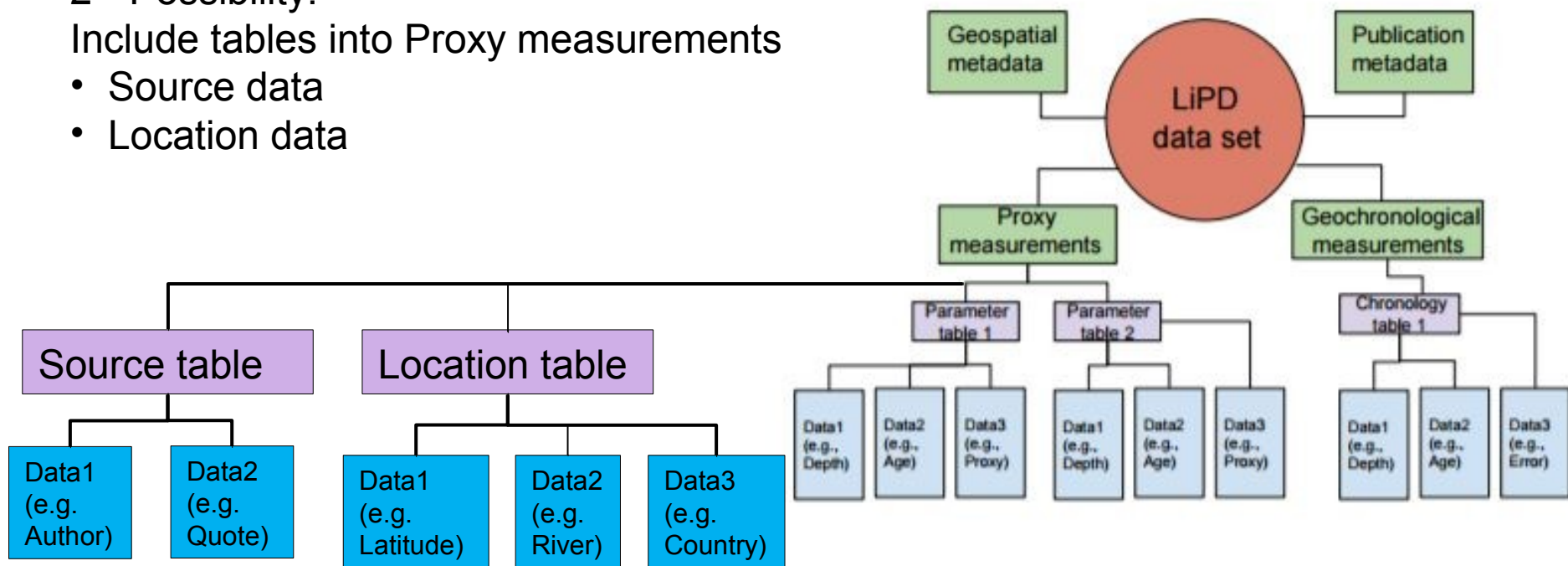
LiPD – Documents Enhancements (2)



2nd Possibility:

Include tables into Proxy measurements

- Source data
- Location data



Based on N. P. McKay and J. Emile-Geay

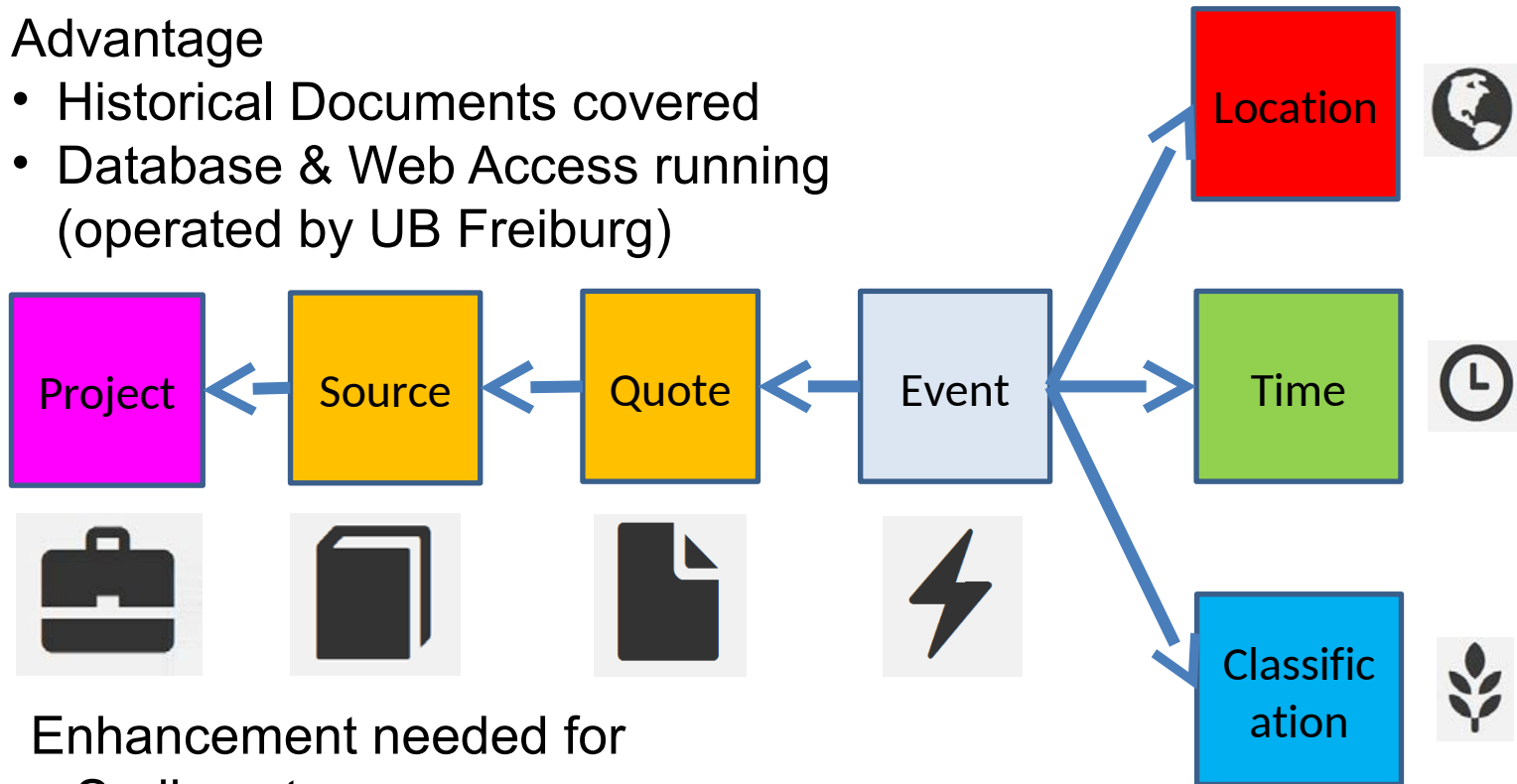
Additional wish: Provide a version with json only (no csv files)



- The final goal is to provide an online platform to store the flood data into hosted DB
- Entries should be easy to add (e.g. by uploading csv or LiPD files)
- Search function provided to search floods by archive type, time frame, location area and/or magnitude
- Access data by downloading files or via API
- Sustainable long-term availability desired

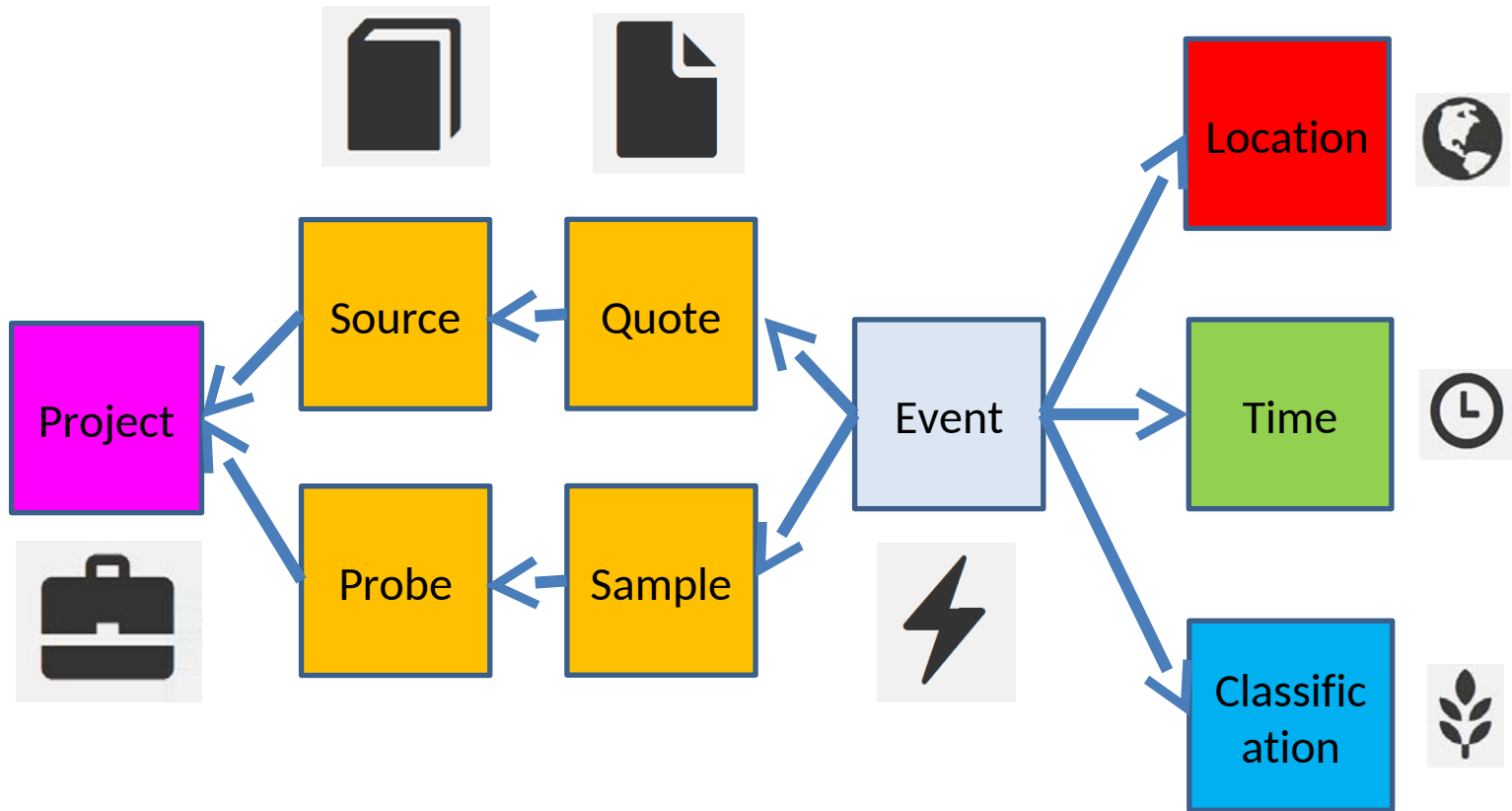
Advantage

- Historical Documents covered
- Database & Web Access running (operated by UB Freiburg)



Enhancement needed for

- Sediments
- Speleothems
- Tree rings
- Measurements on fixed points



- Maybe other organizations or institutions can host the flood database
- Other Pages projects or WG may face the same issue
- LiPD on <http://linked.earth> would be very promising
- Long-term operation is crucial
-

Next steps



- Enhance LiPD to cover historical data as well
 - Add individual locations
 - Add individual sources and quotes
 - Standardize column names and units
- Agree on online platform and DB
 - Where to host
 - Find sustainable long-term strategy
- Agree on implementation
 - Who can join the work
 - How to finance development (Projects, etc.)

Search Events

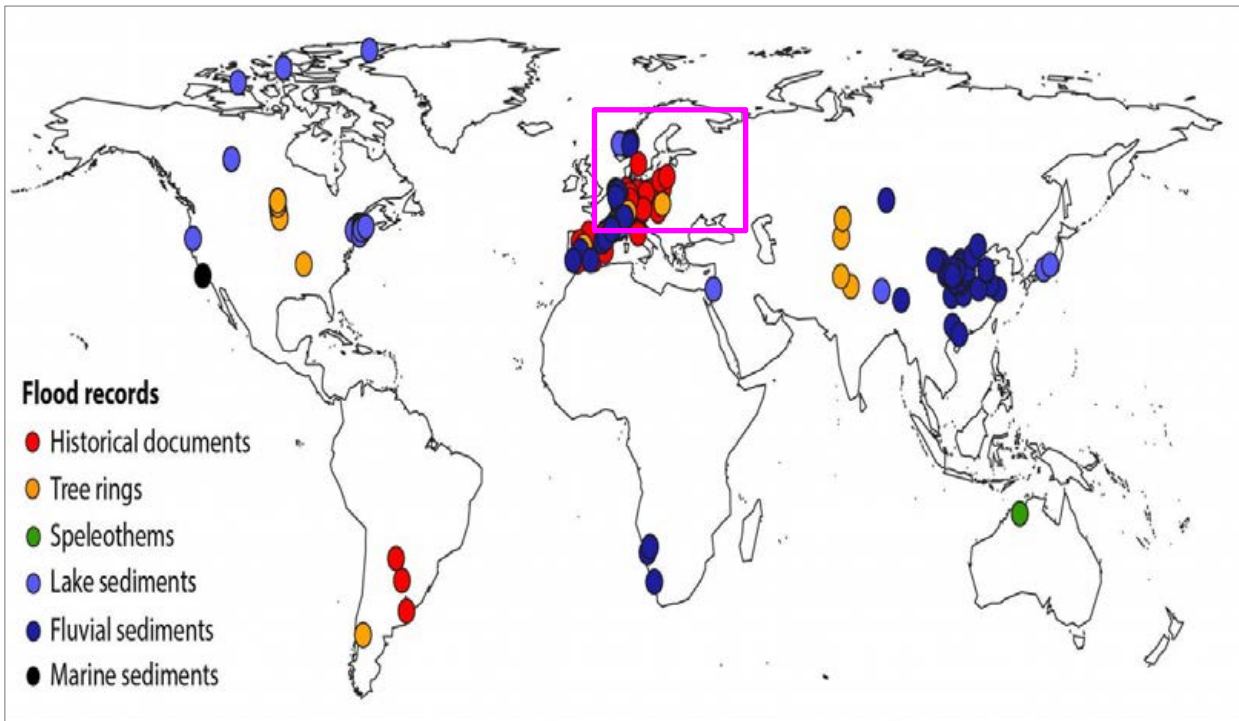
Q Query Text ✕

🕒 By Year ✕

🌐 By Country ✕

🌊 By River ✕

🔗 Sort by



Institution	Proxy Type	Acquisition	Latitude	Longitude	Resol	Event Date	Uncert	Magni	DOI	Email	Citation
Uni Freiburg	Document	2017-05-05	8,4578	42,1478	1	1852-11-23	1	3		michael.kahle@geog1	
Uni Freiburg	Tree rings	2016-10-04	8,1234	43,457	2,5	1852	365	2		michael.kahle@geog1	
Uni Freiburg	Spelethems	2017-01-14	7,9874	42,879	1,5	1847	365	3	10.6094/tambora.org/2016/c15		
Uni Freiburg	Sediments	2016-08-04	8,1234	42,573	2,5	1122	365	3		michael.kahle@geog1	



Thank you