

API vs. Data Format

- Not exclusive
- Users and use cases
- Few big players vs. many small players
- Privacy easier using API?
- APIs as easier to get started with, RDF has learning curve (e.g. tooling)
- APIs not extendible
- ACTION: what formats, what they contain, who is using them

Characteristics of common format

- Small minimal core, extendible
 - e.g. optional/required/recommended
 - Not the ontology of all things social
- Syntax-free, if possible
 - Syntax should be easily parsable, embeddable into other formats without breaking them etc.
- Preferably not a new format
 - Extensions, or even just mappings between existing formats
- Nice to have:
 - Privacy or at least provenance
 - Tracking usage of data

Current options

- Where should data live?
 - In page, linked to page, in silo
 - Data in logs
- FOAF vs. hCard
- Documentation, tooling important

Modeling domain-specific information

- Needed: use cases
 - What are the typical queries, transformations, mining tasks?
 - Existing combinations of social media information
 - Zones of information: personal, social, professional
- Think for the user: he/she may not realize what can be deduced

Privacy handled by data format?

- Thorny, esp. in conjunction with data aggregation
- Nice to have:
 - Rights to data
 - Owner of data
 - Expiry of data
 - Retraction of data
- Complex issue: ethical, legal challenges globally
- Is this DRM for social networks?
- Keep it simple:
 - Robots.txt
 - Tagging for privacy