

Rich Context: Affordances

stakeholders	pattern	exemplars	practices	issues
<i>researchers, analysts</i>	“knowledge graph of metadata about dataset usage”	Lyft (Amundsen), LinkedIn (Data Hub)	rec sys: "regain more time" for science – otherwise wasted by poor search and discovery	KG captures context better than federating many relational DBs in silos
<i>agencies</i>	“data as a strategic asset”	Bundesbank	"data impact factor" metric; data stewardship tools	agency resource allocation
<i>scholarly infrastructure</i>	“human-in-the-loop is a feature, not a bug”	RePEc; Researchgate	feedback loops: authors confirm inferred metadata, to update the graph and models	HITL: people are not going away; manual input is extremely valuable for successful ML projects
<i>public, policymakers, editors</i>	“domain knowledge is more important than other components of data science”	evidence based policymaking	the knowledge graph is ultimately about people	the constituency determines the value of programs in government
<i>business</i>	“building models beats crunching data”	iNaturalist	transfer learning reuses pre-trained models; scope of DL is beyond human scale, but humans make use of results.	how long do we depend on the "good graces" of tech unicorns to provide DL models?
<i>people working on KG</i>	"individuals and interactions over processes and tools”	DLA, Primer AI, etc.	validate methods and resources before automating	other HITL: collaboration surfaces key issues about curation, security, privacy, ethics