

jHound

Large-Scale Profiling of Open JSON Data

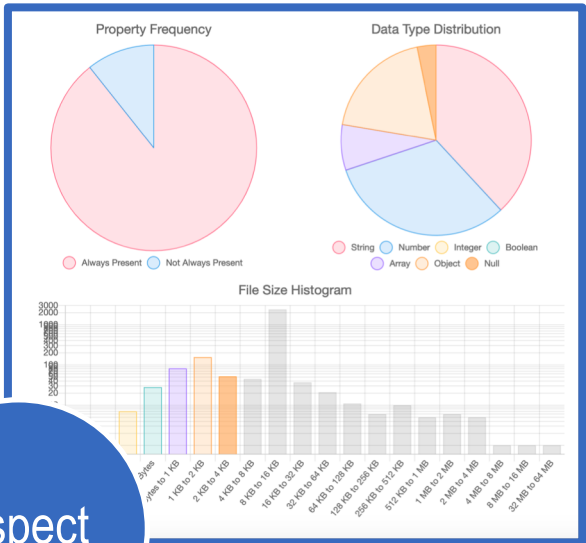
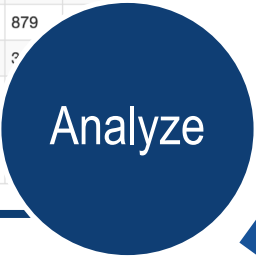
*Mark Lukas Möller, Nicolas Berton, Meike Klettke,
Stefanie Scherzinger, Uta Störl*

jHound – Large-Scale Profiling of Open JSON Data

```
{
  "problem_statement":
  [
    { "problem": "JSON can be used without a schema" },
    { "problem": "Arbitrary data can be stored" },
    { "problem": "Analyzing JSON is not trivial" }
  ],
  "tool_goal":
  [
    { "goal": "Find out how people use JSON" },
    { "goal": "Discover certain documents (e.g. CSV as JSON)" }
  ],
  "main_scope": "Open JSON-Data"
}
```

In-Depth Inspector

UUID Prefix	Req. Props.	Opt. Props.	Str	Num	Int	Bool	Arr	Obj	Null	Simple
2a6d2868	58	12	190645	56060	0	6	13954	328	60384	13937
8bfc3eaf	61	12	23260	3890	0	882	1791	340	879	
2ff63281	58	15	155558	9733	0	2007	4229	1243	?	
d6806eea	56	12	50371	4666	0	882	1821	1026		
d862df0e	62	12	187144	15776	0	6	5026	750		
50766016	58	15	22862	5847	0	6	1912	163		



```
{
  "help": "https://catalog.data.gov/api/3/action/...",
  "license_title": "Other License Specified",
  "wprdc@pitt.edu",
  "num_tags": 3,
  "id": "7a3cfcf...",
  "author": null,
  "author_email": null,
  "state": "...",
  "cache_last_updated": null,
  "package_id": "7a3...",
  "state": "active",
  "hash": "...",
  "description": "..."
}
```



```
[ [ "organization",
  "interaction_intervention_id",
  "org_intervention_id",
  "project_tags",
  "project_name", ... ], ...
[ "PROJECT C.U.R.E.",
  "59-21401",
  "11711",
  "",
  "Panama In Kind Donations ", ... ], ... ]
```

