Internet Yellow Pages
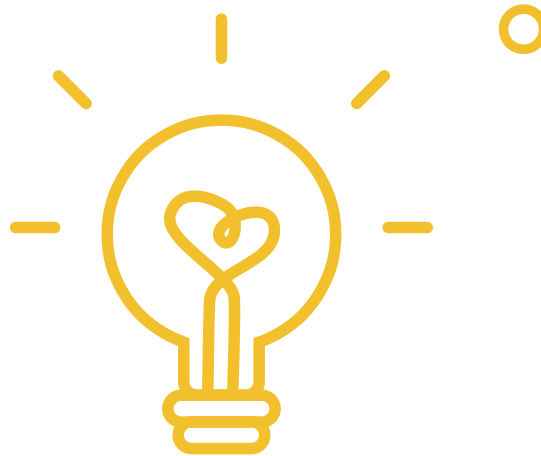
Romain Fontugne et al.

IIJ

Internet Initiative Japan

How many Internet measurement datasets have you used this year?

Internet measurements data

The good, the bad, and the ugly

Tons of open data.

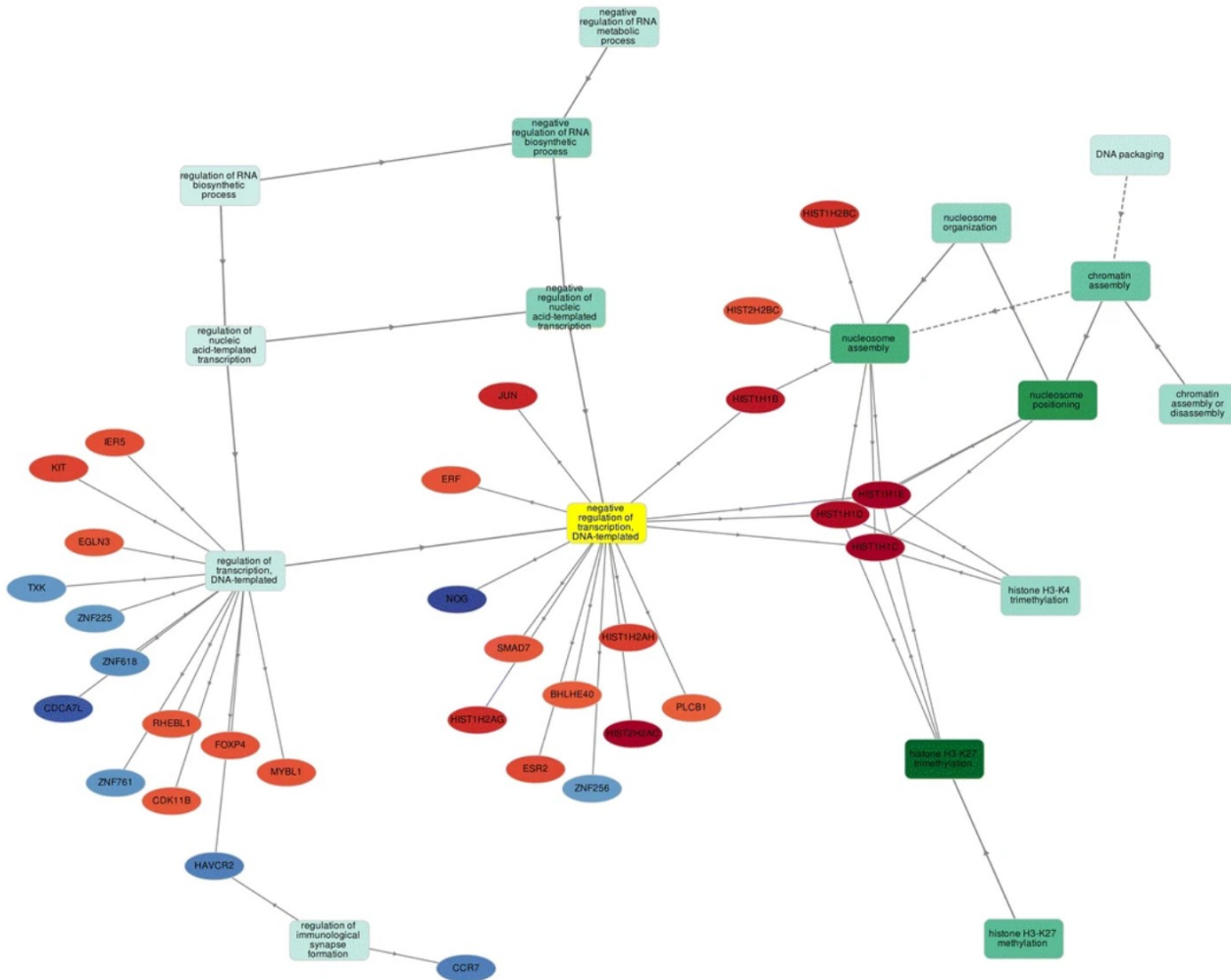Internet measurement community produces numerous open access datasets.
This is great!

Not easy to deal with.

Each dataset using a different format, terminologies, identifiers, etc... and each researcher repeating the same data formating tasks.

Collective knowledge?

- Currently hard to answer simple questions.
- Datasets are merged only if there is expected benefits.
- Penalize innovation and data exploration.

Gene Ontology

Drug Bank

Internet Yellow Pages

A knowledge graph of Internet resources:
- integrate various Internet datasets
- provide quick and homogeneous access to data
- collective knowledge of the community

Building IYP

Datasets — **Knowledge Extraction** — **Fusion & Refinement**

PeeringDB AS to Org.

asn	org_id
AS65536	123

BGPKIT pfx2asn

asn	prefix
65536	2001:0db8::/64
65536	2001:0db8::/32

IHR ROV

prefix	RPKI
2001:DB8::/32	Valid

OpenINTEL Tranco

hostname	ip
example.com	2001:DB8::1

AS — Managed by — Org.

AS — Originate — Pfx.
AS — Originate — Pfx.

Pfx. — Categorized — RPKI Valid

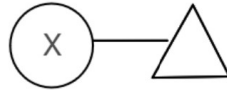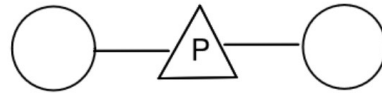Hostname — Resolves to — IP

Part of
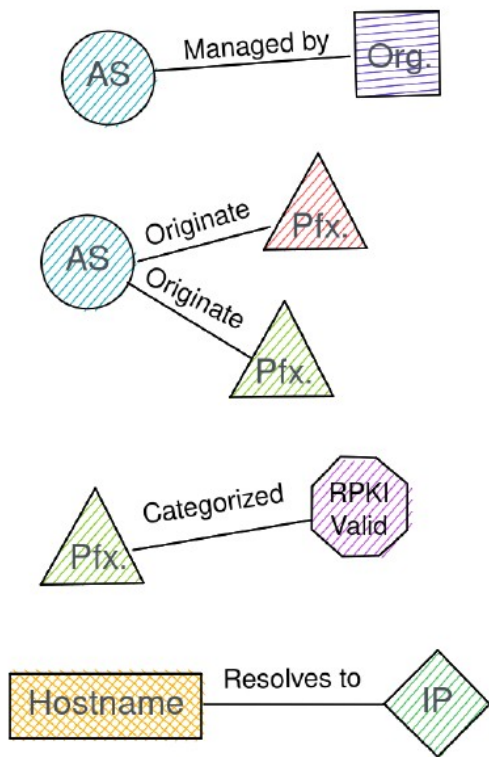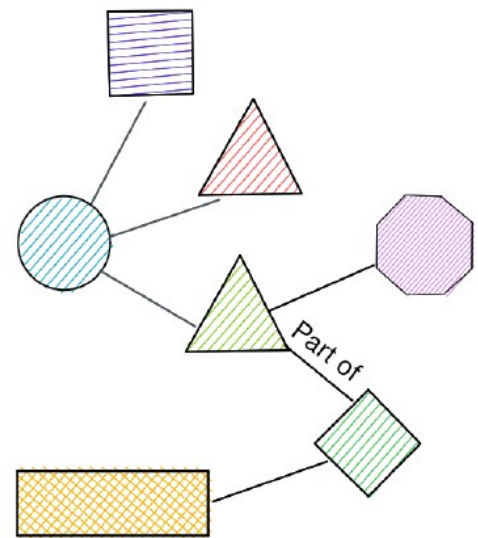
IYP in numbers

- 1 database (neo4j)
- 46 datasets from xx organizations
- 10 million nodes
- 68 million relationships

- Ontology:
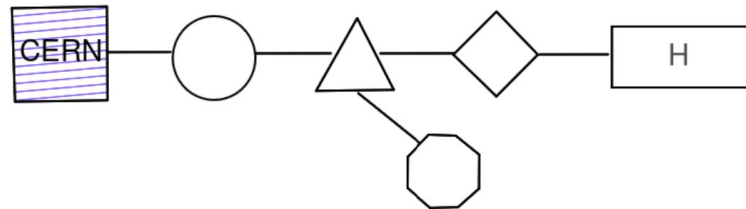   - 24 node types
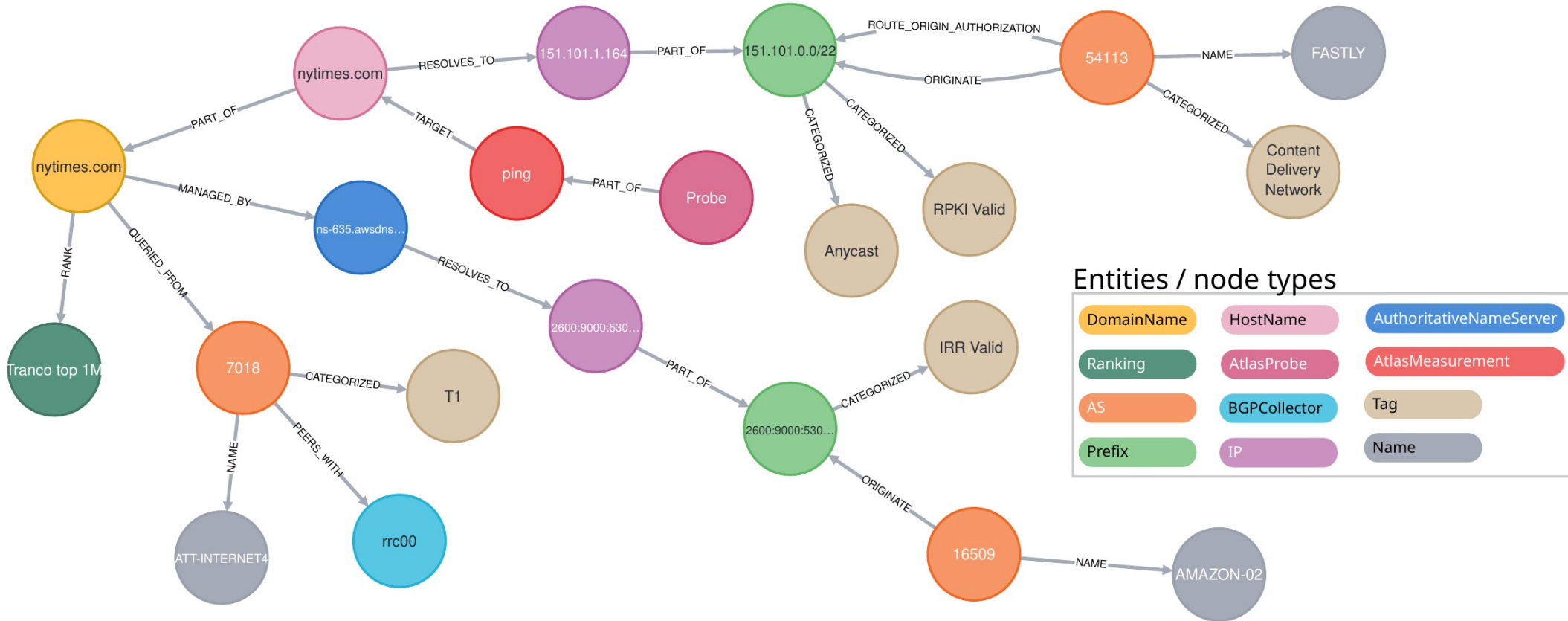   - 24 relationship types

Semantic search

Data exploration

In practice

Entities / node types

- DomainName
- HostName
- AuthoritativeNameServer
- Ranking
- AtlasProbe
- AtlasMeasurement
- AS
- BGPCollector
- Tag
- Prefix
- IP
- Name

01 Sharing queries

- Sharing queries = Sharing datasets
- Jupyter netbook = On demand reproducibility

02 Augmented datasets

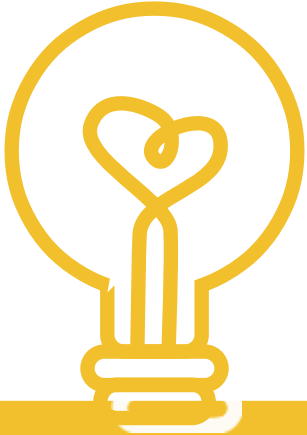- Integrate your data to IYP = connect your data to all IYP datasets

03 Local instance

- Play locally with IYP
- Integrate confidential data

04 Community engagement

- Maintain ontology
- Integrate datasets
- Entirely open source
- https://iyp.iijlab.net

THANK YOU

romain@iij.ad.jp
https://iyp.iijlab.net