

Exploring the research utility of fan-created data in the Japanese visual media domain

Senan Kiryakos and Magnus Pfeffer

`kiryakos,pfeffer@hdm-stuttgart.de`

Funded by

DFG

Deutsche
Forschungsgemeinschaft

German Research Foundation

Project context

Initial observation

- There is no central resource on Japanese visual media that is suitable for research
 - But: Enthusiast communities have been discussing and collecting information on Japanese media online for more than 20 years
- Idea: Make this data available in a central location
 - “Japanese Visual Media Graph” (JVMG) project

JVMG project: main tasks

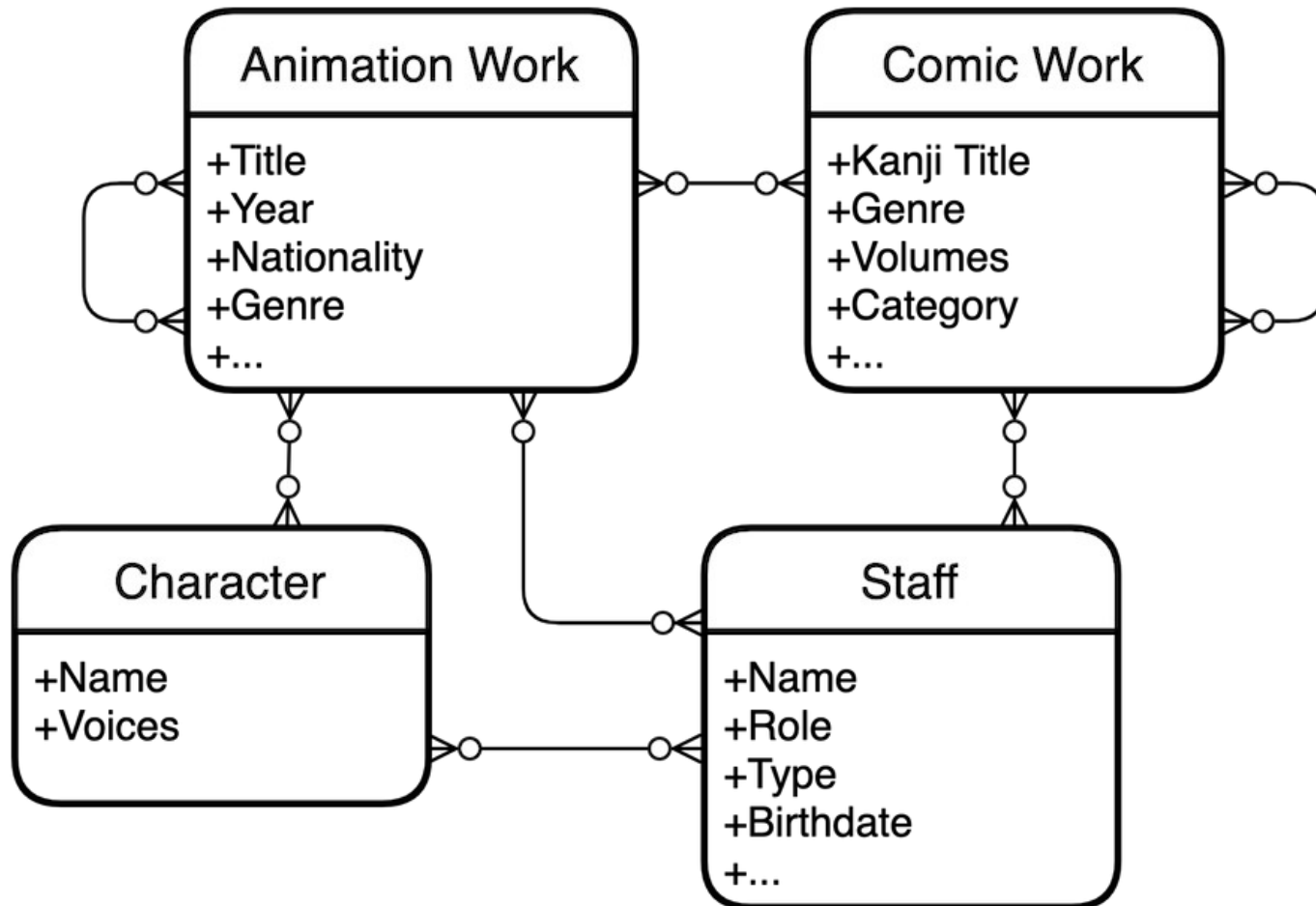
- Establish collaboration with the communities
- Create an ontology for the domain of Japanese visual media
- Build a central repository for the database
- Evaluate the utility of the fan-created data

Data structures and vocabularies

- Localized in Italy, main language is Italian
- Community with a **broad** focus on Japan and Japanese media
 - Manga and anime
 - TV drama
 - Culture, Food, Language
- Core services
 - Structured information on media and topics
 - Discussion boards
 - Presenting fan works (art, cosplay)

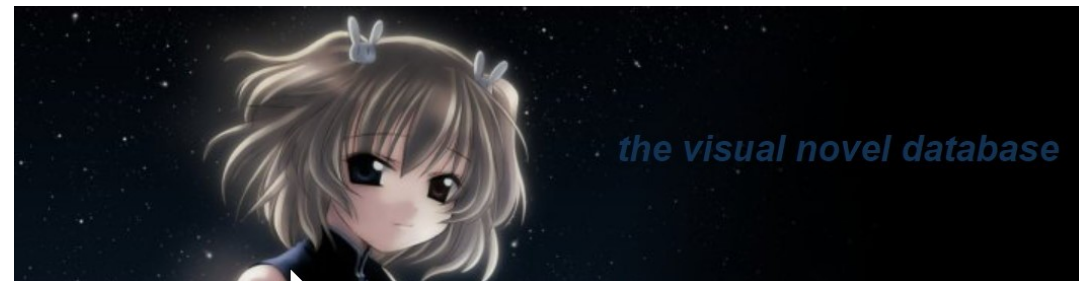


AnimeClick

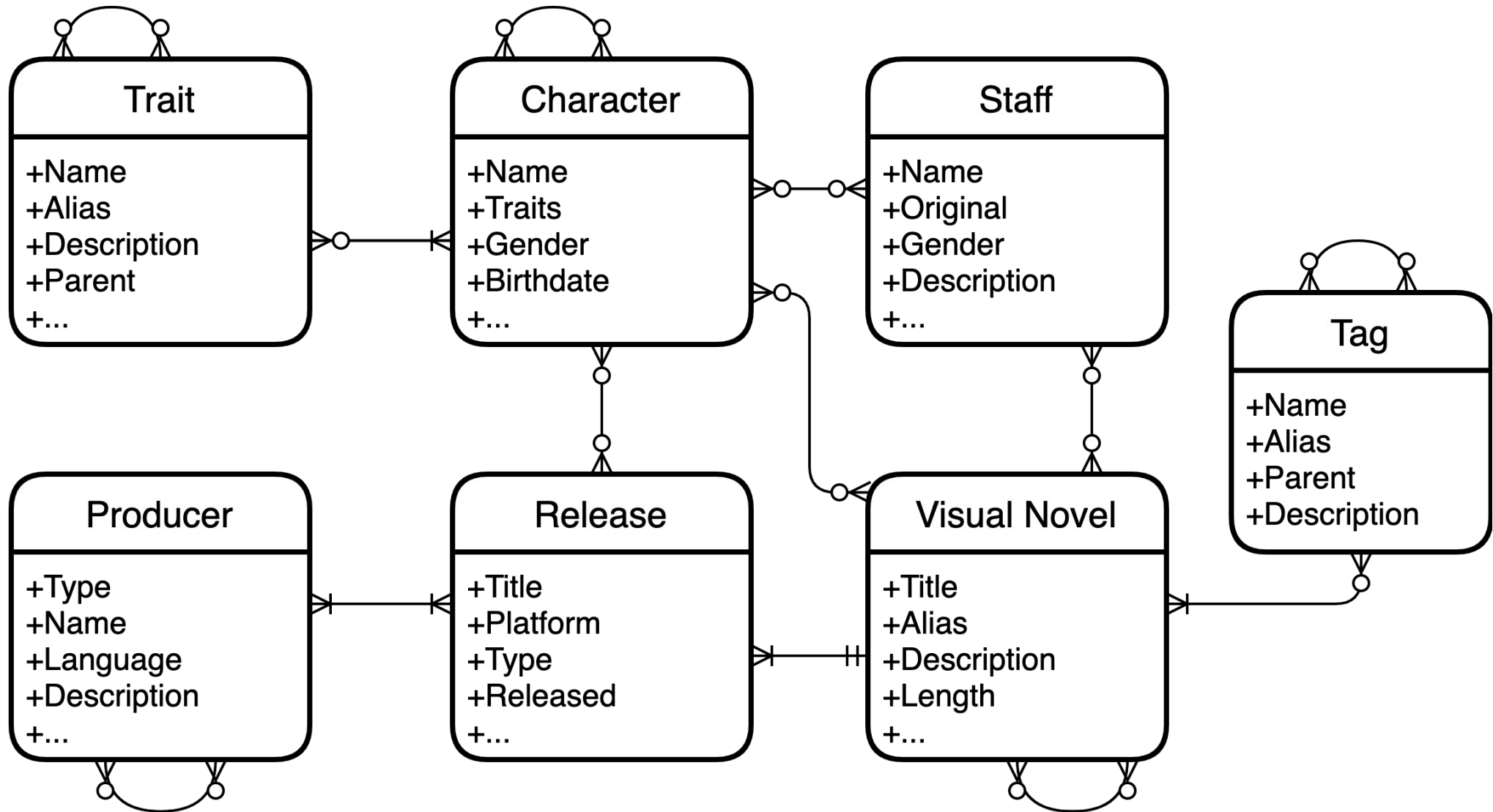


The Visual Novel Database

- International community, English language
- Community with a very **narrow** focus on a single computer game genre
- Core service: Games database
 - Detailed information
 - Tag hierarchy for search and discovery
 - Discussion forums



The Visual Novel Database

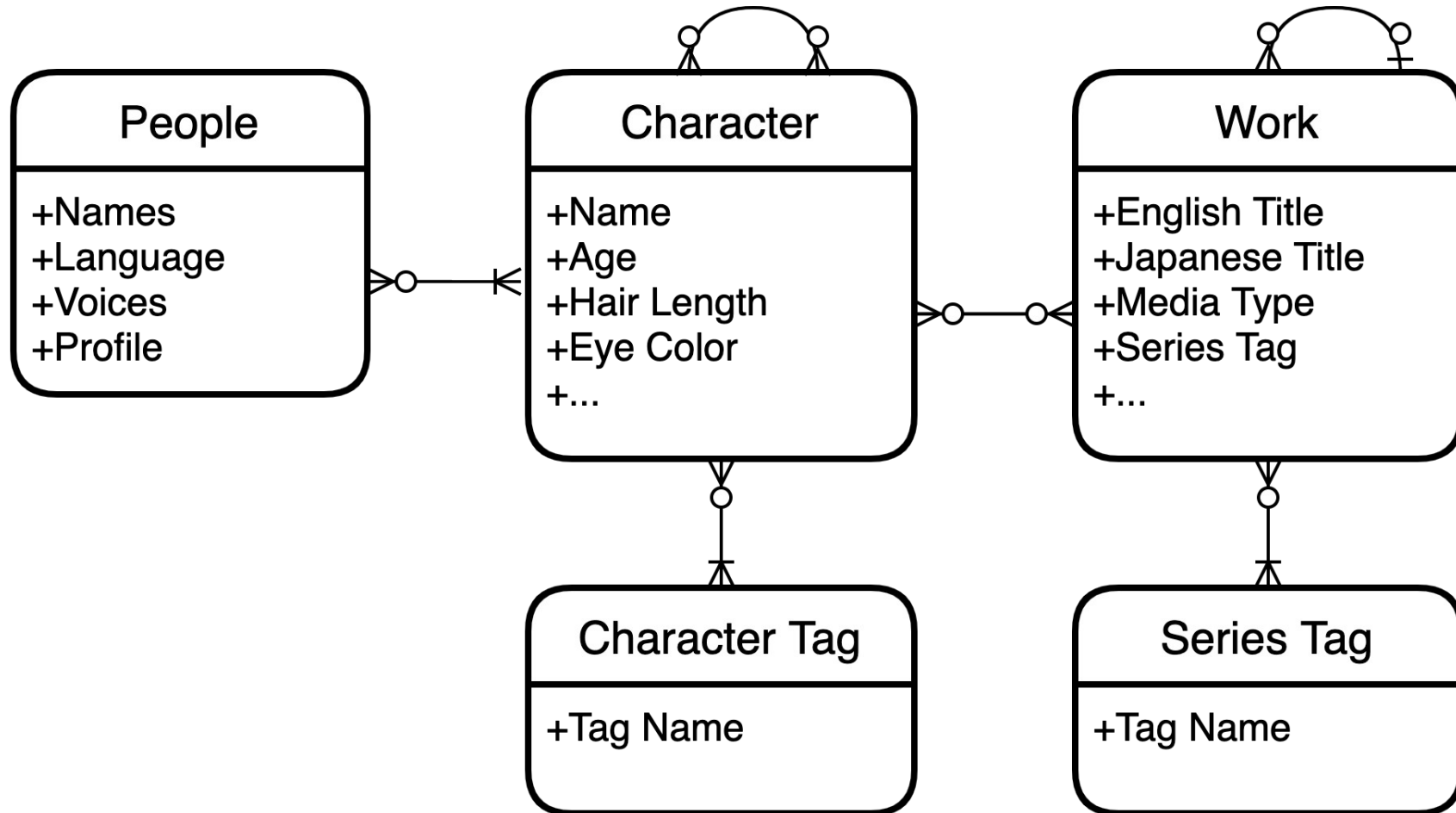


Anime Characters Database

- International community, English language
- Community with a focus on a **specific aspect** of visual media
- Core service: Visual search engine
 - Appearance and clothing of characters
 - Games and other fun interactions with the site



Anime Characters
Database



Results from working with community data

Impact on ontology development

- Very different granularity of the data
 - More focussed → more details
- Data from different sources can be connected
 - Enough overlap between data
- Requirement of “media mix” or “media franchise” as a type of entity
 - Connects different media forms
 - Point of entry for queries and browsing
 - c.f. Wikipedia articles

Original research

- Method: “Tiny Use Case” (TUC)
 - Similar to agile development methods
 - Mixed team of researchers from media studies and information science
 - Basis: A research question of limited scale that can be answered given the right data
 - Acts as a bridge between disciplines
 - Discussion on a concrete matter instead on principles

Original research

- TUC: Co-occurrence of character traits
 - Discovery of recurring patterns (character archetypes)
 - Tracking of trends over time
- Highly detailed data on characters available from VNDB
 - Almost all characters are described using multiple traits
 - Traits are ordered in a hierarchy with usage notes
 - Very consistent use of the individual traits

- Coverage
 - International communities might not have access to rare media
 - Use of authoritative data from Japan to supplement the database

- Quality
 - Community members might have trouble with Japanese titles
 - Randomized study (CI 95%+/-5%)
 - Typographical differences in 5.6%-28.2% of cases
 - True semantic errors only in 0-2.5% of samples

Outlook

Current state and developments

- Database available as RDF with custom web frontend
 - Separate data for each community
 - Additional data
 - Wikidata
 - Media Arts Database
 - AniDB (titles only)
 - Web interface for browsing
 - Plug-in architecture for data analysis web services
- Multiple “Tiny Use Case” research scenarios
 - See project blog for details

Thank you for your attention

Project blog:
<https://jvmg.iuk.hdm-stuttgart.de/>

Main database page:
<https://mediagraph.link/>

