Enthusiast models of the Japanese visual media domain
Outline

● Japanese Visual Media Graph (JVMG) project
● Enthusiast communities and their data models
● Observations
Introducing the JVMG project

- Databases by fan/enthusiast communities have collected huge amounts of data on Japanese visual media

- **Japanese Visual Media Graph (JVMG) project proposal**

- **Project aim:** Make these databases available for large-scale quantitative research, in collaboration with the communities

- 3 year grant funded by the “e-Research Technologies” program of the German Research Foundation
Source databases

Fan/enthusiast community databases:

- **AnimeClick**: Wide interest in Japanese visual media and culture
- **The Visual Novel Database** (VNDB): Focused on visual novel games only
- **Anime Characters Database** (ACDB): Focus on one aspect of the domain

Other databases:

- **Wikidata**: Not focused on Japanese visual media
- **Media-Arts Database**: Collects information on manga, animation, games and media art from institutions, creators and publishers in Japan
Entities instead of string literals

● More possibilities for retrieval
  ○ Find all works by a creator
  ○ Find all characters that share a set of traits
  ○ Find all media that share a set of tags

● Allows for complex relationships
  ○ Role of person in media work creation
  ○ Role of character in media
Community: vndb.org
Community: animeclick.it
Entities in fan databases

- Decision of entity vs. string literal is not arbitrary
- Entities require more work
  - Database design (extra tables with fields)
  - Editorial control (avoiding duplicates and false matches)
  - Continuous development of concept hierarchies
    - New concepts
    - Usage guidelines
Common entities

- **Core entities** universal to all sources
  - Media
  - People (artists, directors, production staff, voice actors, …)
  - Visual characters

- **Additional concept hierarchies** to describe
  - Media
  - Visual characters
## Entity and concept numbers

<table>
<thead>
<tr>
<th>Enthusiast community</th>
<th>Works and media</th>
<th>Company</th>
<th>Characters</th>
<th>Work properties</th>
<th>Character properties</th>
<th>Involved people</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACDB</td>
<td>Work</td>
<td>Character</td>
<td>Work Tag</td>
<td>Character Tag</td>
<td>People</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.207</td>
<td>107.369</td>
<td>1.088</td>
<td>4.051</td>
<td>5.557</td>
<td></td>
</tr>
<tr>
<td>AnimeClick</td>
<td>Animation Work</td>
<td>Character</td>
<td></td>
<td></td>
<td>Staff</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comic Work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.491</td>
<td>11.762</td>
<td></td>
<td></td>
<td>39.604</td>
<td></td>
</tr>
<tr>
<td>VNDB</td>
<td>Visual Novel</td>
<td>Character</td>
<td>Tag</td>
<td>Trait</td>
<td>Staff</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Release Producer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28.190</td>
<td>71.349</td>
<td>10.394</td>
<td>90.077</td>
<td>2.585</td>
<td></td>
</tr>
</tbody>
</table>

**VNDB**: Visual Novel Database

**ACDB**: AnimeClick Database

**Involved people** indicates the number of people involved in the project, while **Character properties** include the properties of characters, such as tags and traits.
The object-work-franchise relationship

- **Three layers**: physical object – abstract work – franchise
- **VNDB**: physical media to abstract work (bottom two layers)
- **AnimeClick and ACDB**: Clustering media into media franchises (top two layers)
- **Media-Arts DB**: bottom up from physical objects towards franchise (franchise still work in progress)
- **Wikidata**: Very fragmented (no agreement on a common ontology)
Challenge: Representing media works

● Media works granularity is not obvious
  ○ TV series as single entry or multiple seasons?
  ○ Special episodes as part of a series, part of season or extra entry?
  ○ How about multi-part movies?

● Possible solutions
  ○ Matching similar to ontology alignment: full match, partial match
  ○ Merging information after decision on representation granularity
    ■ Aggregate data from multiple entries into one
    ■ Distribute data from one entry into one
Observations

- Enthusiast groups focus on different aspects of the Japanese Visual media domain
- Their data models are detailed, well designed and share a set of core entities
- Multiple sources can be merged into a more complete representation of the domain
- Media works remain challenging due to differing representations
Thank you for your attention!

Get in touch at: pfeffer@hdm-stuttgart.de

Visit our project website: https://jvmg.iuk.hdm-stuttgart.de/

Visit the JVMG database: https://mediagraph.link/