Utilizing Metadata Analytics for Research on Manga, Anime and Video Games

Introducing the Japanese Visual Media Graph

Research funded by:
Outline of the presentation

1. Introducing **metadata analytics** and the **JVMG project**

2. A sample **research question**: Testing one of Hiroki Azuma’s points from *Otaku: Japan’s Database Animals*
   i. Hypothesis
   ii. Data sources
   iii. Descriptive statistics
   iv. Regression analysis

3. Implications of the results for theory, and lessons learned
Introducing metadata analytics and the JVMG project
Data-driven methodologies

From digital humanities to **computational humanities**

- Quantitative **textual analysis** - distant reading, culturomics
- **Web native** approaches - ‘digital methods’, webometrics
- **Algorithmic data extraction** - cultural analytics
- **Metadata** analytics
Metadata analytics

- Roots in **bibliometrics** and **scientometrics**
- Based on **descriptive metadata** in **library and information science**
- Large metadata databases built by **fan/enthusiast communities** online*
- Understanding cultural objects and processes **at scale**
Introducing the JVMG project

- Databases by fan/enthusiast communities are the **go to resource for checking information**

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

- **Project aim:** Make these databases available for **large-scale quantitative research**

- Funded by the **German Research Foundation**’s (Deutsche Forschungsgemeinschaft) e-Research Technologies program
### Goku

Resource: [http://mediagraph.link/acdb/character/15533](http://mediagraph.link/acdb/character/15533)

#### Graph: acdb

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>label</td>
<td>Goku</td>
</tr>
<tr>
<td>type</td>
<td>Character</td>
</tr>
<tr>
<td>Age</td>
<td>Adult</td>
</tr>
<tr>
<td>Animal Ears</td>
<td>No</td>
</tr>
<tr>
<td>Appears in</td>
<td>15</td>
</tr>
<tr>
<td>Dragon Ball (Series)</td>
<td>• Dragon Ball GT</td>
</tr>
<tr>
<td>Dragon Ball GT</td>
<td>• Dragon Ball Super</td>
</tr>
<tr>
<td>Dragon Ball Super</td>
<td>• Dragon Ball Z</td>
</tr>
<tr>
<td>Dragon Ball Z</td>
<td>• Dragon Ball Z: Battle of Gods</td>
</tr>
<tr>
<td>Dragon Ball Z: Battle of Gods</td>
<td>• Dragon Ball Z: Bojack Unbound</td>
</tr>
<tr>
<td>Dragon Ball Z: Broly - The Legendary Super Saiyan</td>
<td>• Dragon Ball Z: Broly - The Legendary Super Saiyan</td>
</tr>
<tr>
<td>Dragon Ball Z: Cooler's Revenge</td>
<td>• Dragon Ball Z: Cooler's Revenge</td>
</tr>
<tr>
<td>Dragon Ball Z: Dead Zone</td>
<td>• Dragon Ball Z: Dead Zone</td>
</tr>
<tr>
<td>Dragon Ball Z: Lord Slug</td>
<td>• Dragon Ball Z: Lord Slug</td>
</tr>
<tr>
<td>Dragon Ball Z: Revival of ‘F’</td>
<td>• Dragon Ball Z: Revival of ‘F’</td>
</tr>
<tr>
<td>Dragon Ball Z: Super Android 13!</td>
<td>• Dragon Ball Z: Super Android 13!</td>
</tr>
<tr>
<td>Dragon Ball Z: The Tree of Might</td>
<td>• Dragon Ball Z: The Tree of Might</td>
</tr>
<tr>
<td>Dragon Ball Z: The World’s Strongest</td>
<td>• Dragon Ball Z: The World’s Strongest</td>
</tr>
<tr>
<td>Dragon Ball Z: Wrath of the Dragon</td>
<td>• Dragon Ball Z: Wrath of the Dragon</td>
</tr>
<tr>
<td>Character Role</td>
<td>Protagonist</td>
</tr>
<tr>
<td>Character Type</td>
<td>Normal</td>
</tr>
<tr>
<td>Character tag</td>
<td>• arm guards</td>
</tr>
<tr>
<td>karate</td>
<td>• karate</td>
</tr>
<tr>
<td>is Clone of</td>
<td>• Super Saiyan 3 Goku</td>
</tr>
<tr>
<td>Super Saiyan Blue Goku</td>
<td>• Super Saiyan Blue Goku</td>
</tr>
<tr>
<td>Super Saiyan 4 Goku</td>
<td>• Super Saiyan God Goku</td>
</tr>
<tr>
<td>spiky hair</td>
<td>• spiky hair</td>
</tr>
<tr>
<td>Extra details</td>
<td>Birthday: April 16</td>
</tr>
<tr>
<td>Height</td>
<td>175cm</td>
</tr>
<tr>
<td>Race</td>
<td>Saiyan</td>
</tr>
</tbody>
</table>
A sample research question:
Testing one of Hiroki Azuma’s points from *Otaku: Japan’s Database Animals*
Find out more

Martin Roth, Hiroshi Yoshida and Martin Picard (Eds.) (2021)

Japan's Contemporary Media Culture between Local and Global

https://hasp.ub.uni-heidelberg.de/catalog/book/971
Otaku: Japan’s database animals

- 東 浩紀, 2001: 動物化するポストモダン—オタクから見た日本社会（講談社現代新書）

- Hiroki Azuma, 2009: Otaku: Japan’s database animals (University Of Minnesota Press)

- On the production and consumption paradigm defining Japanese anime, manga, light novels and games in late modernity.
Formulating a hypothesis

- The **metaphor of the database** captures the space of imagination shared by creators and consumers.

- "As a result, many of the otaku characters created in recent years [the late nineties to two thousand] are connected to many characters across individual works, rather than emerging from a single author or a work." (Azuma, 2009: 49)

- The **hypothesis** to test: The **portion of new characters with shared traits** should increase over time.
Data sources

1. **The Visual Novel Database** (VNDB) [vndb.org](http://vndb.org)
   - only focuses on visual novels
   - very rich and rigorously structured ontology of traits
   - ~79,000 number of characters at the time of the study

2. **Anime Characters Database** (ACDB) [www.animecharactersdatabase.com](http://www.animecharactersdatabase.com)
   - collects data on a wide range of characters from various media (although predominantly focusing on anime and visual novels)
   - hybrid system for describing characters:
     - a closed ontology for eight flagship traits
     - opportunity for free-form tagging of characters with user created labels
   - ~100,000 number of characters at the time of the study
Operationalization of our concepts

- **minimum number of five shared traits** needed to consider two characters to have shared traits
- **one year window** is the time frame for which we compare characters
- to allow for a better comparison between the VNDB and ACDB data, the **ACDB data was separated into two datasets**:
  - visual novels
  - other works (the majority is anime)
Descriptive statistics
Number of characters in the ACDB data by year and type

- Anime, manga, etc
- Visual novels
Average number of traits per character in the ACDB data by year and type

- Green line: anime, manga, etc
- Purple line: visual novels
Average number of characters traits are shared with in the VNDB data by year
Average number of characters traits are shared with in the ACDB data by year and type

- anime, manga, etc
- visual novels
Regression analysis
(ACDB only)
Regression analysis

- **Dependent variable:**
  - *average number of characters traits are shared with*

- **Independent variables:**
  - *number of characters*
  - *average number of traits*
  - *year*

- Also included in the model building and selection process:
  - **Squared terms** for independent variables
  - **Interaction terms** between independent variables
Regression results for ACDB data for visual novel characters with first appearance 1998-2019

No. of characters traits are shared with (avg)

Adj. R-squared: 0.982

No. of characters X no. of traits (avg)
Regression results for ACDB data for anime & other non-VN characters with first appearance 1975-2019

No. of characters traits are shared with (avg)

No. of characters X no. of traits (avg)

Adj. R-squared: 0.961
Controlling for the number of traits

- Examining the regression models for subpopulations of characters with given number of traits
Regression results for ACDB data for visual novel characters with first appearance 2001-2020 (only characters with number of traits equal to seven)

Adj. R-squared: 0.991
Summary of regression results

The **hypothesis** (the portion of new characters with shared traits should increase over time) **was not substantiated** by our regression analyses:

- In the **ACDB visual novel characters** dataset **no connection** was found with the temporal variable

- In the **ACDB other characters** dataset **no connection** was found with the temporal variable

- In the **VNDB dataset** the results are a bit more ambiguous, but **most likely** there is no connection
Implications and lessons learned
Implications of the results for theory

• Cannot fully rule out that the hypothesis might still be correct, however, based on the results it is time to seriously consider the alternative

• Without this whole analysis it would not have occurred to me that there is maybe no paradigm shift going on on the production side
What if there was no shift on the production side?

- **Going back to Osamu Tezuka**
  - In manga the **“star system”** (e.g. Shunsaku Ban “Higeoyaji”)
  - In anime the **cel bank**

- If the production side has followed this model of relying on the “database” of character elements, then maybe it is only later that the consumption side has caught up to it

- We would need a **stronger acknowledgment of Toshio Okada’s work** highlighting the way otaku have from the start been engaged with the database aspect of the production side of anime

- **Thomas Lamarre**, has already pointed out this connection between Okada and Azuma in *The Anime Machine*
Lessons learned

- The **datasets are not complete**

- It is **still possible** to use the available data **to test research questions**, thanks to:
  - The **large number of data points** in the datasets
  - The availability of **multiple sources of data allows for cross-checking results**, and helps circumvent potential biases that might be present in the datasets
Thank you for your attention!

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

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

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