Challenges in working with fanfiction and fanfiction metadata for large-scale analysis

Research funded by:

Deutsche Forschungsgemeinschaft
German Research Foundation

Saskia Dreßler
Mechademia 2025
Kyoto, 7-9 June 2025

Outline of the presentation

HOCHSCHULE DER MEDIEN

1. Fanfiction from the perspective of metadata and textual analysis
2. Case study 1: Collaboration between the GOLEM and JVMG projects
3. Case study 2: Analyzing fanfiction from Fanfiktion.de
4. Summary: Challenges and possibilities of working with fanfiction and fanfiction metadata

Fanfiction from the perspective of metadata and textual analysis

Fanfiction as source data for research

HOCHSCHULE DER MEDIEN

- Metadata
  - Basic information (e.g. publication date, word count, authorship)
  - Content-related data (e.g. genre, age rating)
  - Fan tagging for more detailed classification
- Text
  - Analyzing about narratives and tropes
- Paratext
  - Author notes or comments

HOCHSCHULE DER MEDIEN

Case study 1: Collaboration between the GOLEM and JVMG projects

Introduction of the Case Study

- GOLEM project:
  - Graphs and Ontologies for Literary Evolution Models (GOLEM)
  - Create models for fan fiction to explore thematic relationships across linguistic boundaries
  - 6 different languages and different fan fiction platforms
- Central aspect of the research:
  - Combine both data sets and explore correlations between patterns in fanfiction production and timelines
- Presented at FanLIS 2024
- Published as: Kacsuk et al. (2024) "Utilizing Metadata from Heterogeneous Sources within the Framework of the JVMG and GOLEM Projects to Identify Patterns in Anime-based Fandoms on AO3," *Proceedings from the Document Academy*: Vol. 11 : Iss. 1 , Article 6. DOI: https://doi.org/10.35492/docam/11/1/6

Challenges of combining metadata

HOCHSCHULE DER MEDIEN

- Broad AO3 tag “Anime & Manga”
  - Can also be manhwa/manhua or in some cases video games
  - How can fandoms be clearly defined, and tags distinguished between specific anime/manga use and broader meanings?
- AO3 tags (related to source materials) have a vary wide level of aggregation
  - Supra-work level:
    - Multimedia, transmedia, media franchise
    - Media mix
    - Superwork

Supra-work - Example

HOCHSCHULE DER MEDIEN

- AO3 fandom tag = single franchise
  - e.g. Naruto
- AO3 Tag “... & Related Fandoms” = not a single work, name of the franchise itself
  - e.g. Gundam & Related Fandoms
- AO3 Tag = content based
  - e.g. Universal Century Gundam -> particular timeline of Gundam
- AO3 Tag = specific subset of a given franchise
  - e.g. Attack on Titan (Movies)
  - e.g. Saint Seiya (Manga & Anime 1986)
- AO3 Tag = work level fandoms
  - e.g. Yu-Gi-Oh! Sevens (Anime)

機動戦士ガンダム | Mobile Suit Gundam (TV 1979) (585)

Gundam 00 (946)

機動新世紀ガンダムX | After War Gundam X (23)

BB戦士 三国伝 | BB Senshi Sangokuden (11)

Gundam Build Fighters (171)

機動戦士ガンダム 逆襲のシャア | Gundam: Char's Counterattack (247)

Gundam F91 (28)

G Gundam (163)

機動戦士ガンダム 閃光のハサウェイ | Mobile Suit Gundam: Hathaway's Flash (Movies) (74)

機動戦士ガンダム MS igLoo | Mobile Suit Gundam MS IGLOO (7)

ガンダム Gのレコンギスタ | Gundam Reconguista in G (6)

Gundam SEED Astray (123)

Gundam SEED Stargazer (5)

機動戦士ガンダム サンダーボルト | Gundam Thunderbolt (11)

機動戦士ガンダム TWILIGHT AXIS | Gundam Twilight Axis (4)

Universal Century Gundam (1690)

Gundam & Related Fandoms (14551)

機動戦士ガンダム0083 STARDUST MEMORY | Mobile Suit Gundam 0083: Stardust Memory (33)

機動戦士ガンダムAGE | Gundam AGE (16)

Gundam Build Divers (57)

Gundam Build Fighters Try (24)

機動戦士クロスボーン・ガンダム | Crossbone Gundam Series (5)

SD Gundam Force (55)

ＳＤガンダム外伝 | SD Gundam Gaiden (Anime) (4)

機動戦士ガンダム 鉄血のオルフェンズ | Mobile Suit Gundam: Iron-Blooded Orphans (470)

機動戦士ガンダムＮＴ | Mobile Suit Gundam Narrative (2018) (20)

Gundam SEED (1170)

Gundam SEED Destiny (546)

機動戦士ガンダム 第08MS小隊 | Mobile Suit Gundam: The 08th MS Team (19)

Turn A Gundam (25)

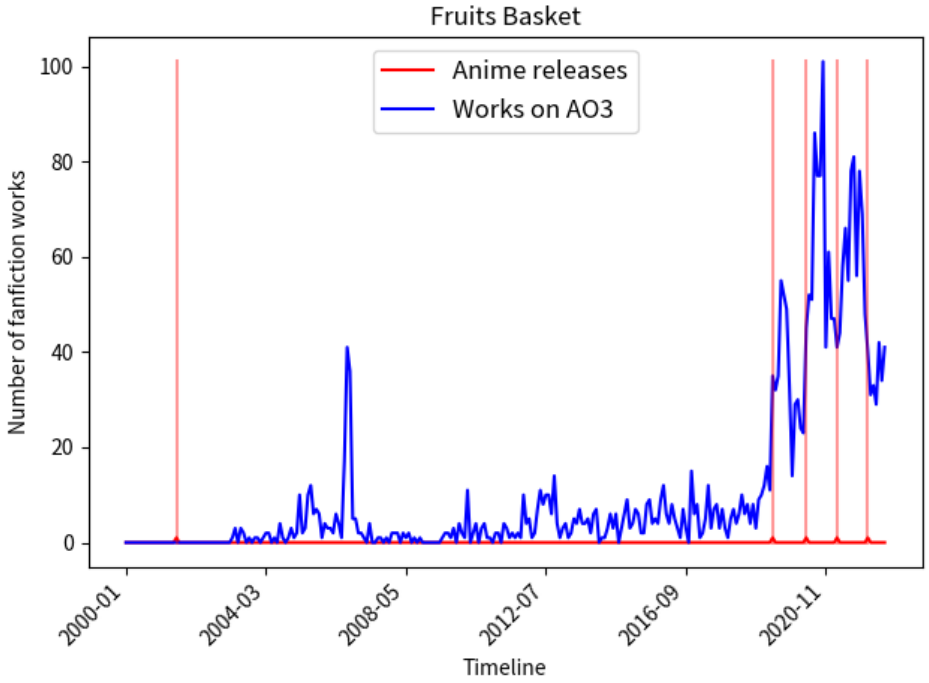
Gundam Unicorn (109)

機動戦士Vガンダム | Victory Gundam (12)

Example of supra-work in the Gundam fandom listed on Ao3

Example from research: Fruits Basket

- Original series aired before AO3's launch in 2009, but later adaptations triggered renewed fan engagement
  - Maybe driven by nostalgia and new audiences
- Additional spike in 2006
  - Maybe due to key plot developments or anticipation of the manga's ending



Graph 4: Fruits Basket from Kacsuk et al. (2024).

Case study 2:
Analyzing fanfiction from Fanfiktion.de

Introduction of the Case Study

HOCHSCHULE DER MEDIEN

- General question:
  - Are there heteronormative elements in German-language fan fiction with homosexual pairing?
  - Own Master's thesis
- Platform: German-speaking fan fiction platform "Fanfiktion.de"
- Method: Qualitativ content analysis

Challenges in working with Metadata

HOCHSCHULE DER MEDIEN

- Limited availability of metadata (in comparison to AO3)
  - Basic data like word count, author
  - Less content-related data (only genre, pairing)
  - No individual fan tagging
- Lack of possibility for ordering the fan fiction
  - Solution: Downloading of the metadata from the website and manually ordering
- Incorrect use of the tag "MaleSlash"

Challenges in working with text

	Naruto-Fandom		Sailor Moon-Fandom	
	MaleSlash	FemSlash	MaleSlash	FemSlash
Number of chapters (average)	22,5	2,1	43,8	1,1
Number of words (average)	117.307	4.097	101.866	2.677
Recommendations (total number)	1.182	15	280	60
Reviews (total number)	2.774	123	277	58
Publication rage	2012 - 2021	2007 - 2022	2014 - 2018	2015 - 2018

Table: Created from own material.

Example from research: MaleSlash-Tagging

HOCHSCHULE DER MEDIEN

- Incorrect use of “MaleSlash” tag in the Sailor Moon fandom:
- No gay protagonists in all stories -> mostly a heterosexual pairing
- Two possible explanations:
  1. Technical error, though this seems unlikely since other Shōjo fandoms did not exhibit the same problem in random sampling.
  2. Indication of internalized heteronormativity among fans

HOCHSCHULE DER MEDIEN

Summary: Challenges and possibilities of working with fanfiction and fanfiction metadata

Possibilities of working with fan fiction

HOCHSCHULE DER MEDIEN

- Broad spectrum of data
  - Metadata
  - Text
  - Paratext
- Allows drawing conclusions about different fan cultures and the mechanisms of fan practices
  - Cross-cultural comparisons are possible
- Both qualitative and quantitative methodologies are applicable, opening up a wide range of analytical tools and approaches

Challenges of working with fan fiction

HOCHSCHULE DER MEDIEN

- Metadata can be incomplete or lack sufficient detail
- Tagging systems may be inconsistent or imprecise
- User-generated tags can be misleading or incorrect
- Different story length
- Available material does not align well with the research question
- From a practical perspective:
  - Data cannot simply be taken at face value

Thank you for your attention!

Get in touch with us:

Email: s.dressler@uni-tuebingen.de

Website: **https://mediagraph.link/**