

A Description of the Visual Novel Database Dataset and Ontology Used in the Japanese Visual Media Graph Project

Senan Kiryakos¹ and Magnus Pfeffer²

^{1,2} Stuttgart Media University, Stuttgart, Nobelstraße 10, 70569 Stuttgart, Germany

¹kiryakos@hdm-stuttgart.de

²pfeffer@hdm-stuttgart.de

1. Intro

This document serves to describe the ontology created by the Japanese Visual Media Graph (JVMG) project for the Visual Novel Database dataset. Following this introduction, the contents of this document are as follows: The Visual Novel Database website will be briefly introduced, followed by a table of the dataset, which includes its primary entities, relationships, and descriptive properties. In a future version, the document will also describe the RDF ontology created based on the source data for use with the JVMG knowledge graph.

In addition to this explanatory document, the Zenodo archive includes the current VNDB RDF ontology created by the JVMG project in the Turtle serialization, as well as an html document which catalogues the ontology.

2. Introduction to the Visual Novel Database

According to their homepage, the Visual Novel Database (VNDB) “strives to be a comprehensive database for information about visual novels. This website is built as a wiki, meaning that anyone can freely add and contribute information to the database, allowing us to create the largest, most accurate and most up-to-date visual novel database on the web.”¹ In other words, the VNDB is a crowdsourced, editable database for information on the visual novel game genre.

3. Entities, Relationships, and Properties in the VNDB Dataset

This section serves to list the exact entities, relationships, and properties featured in the VNDB dataset; these are listed in Table 1. Notes regarding this table:

- I. Certain properties that are only used for internal reference or hold no semantic value outside of the context of the actual VNDB website, such as internal IDs or an integer indication spoiler status, have been omitted.
- II. Certain tags marked with an * have been consolidated into a single property (e.g. the ‘Resolution’ property under the ‘Release’ entity is technically two distinct properties, horizontal resolution and vertical resolution).
- III. Due to the nature of the dataset featuring distinct lookup tables, properties that have a relationship to another entity could logically be attributed to that entity instead; to limit redundancy, these relationship properties have only been attributed to one entity, but all are symmetric.

¹ From <https://vndb.org/>

Table 1. Summarization of the original VNDB dataset

Entity	Property	Value Description	Relationship
Visual Novel	Title	VN title	
	Original	VN title in its release original language	
	Alias	Alias titles	
	Length	Estimated completion time	
	Description	Short description	
	Tags	Descriptive genre and content tags	Tag
	Relations	Related visual novels	Visual Novel
	External Links*	Links to various websites, e.g. homepages, digital stores	
Release	Credits	Individual contributors	Staff
	Title	Release title	
	Original	Release title in its original language	
	Producer	Release producer and/or developer	Producer
	Type	Release type (i.e. complete or partial)	
	Catalog	Barcode identifier (e.g. JAN, GTIN)	
	Released	Release date	
	Subtype*	Indicates if the release is a patch, freeware, or doujin	
	Voiced	Indicates if the content is voiced	
	Animated Story	Indicates if the release features animated scenes	
	Engine	Game engine used	
	Platform	PC or console platform for the release	
	Resolution*	Release screen resolution	
	Character	Name	Character name
Original		Character name in their original language	
Alias		Character name aliases	
Voiced by		Voice artist for the character, if applicable	Staff
Other Instances		Other instances of the same character	Character
Description		Character description	
Traits		Descriptive character traits	Trait
Gender		Character gender	
Measurements*		Character body measurements	
Birthdate*		Character birthdate	
Blood Type		Character blood type	
Age		Character age	
VN Appearance		Visual novel the character appears in	Visual Novel
Release Appearance		Visual novel release the character appears in	Release
Role		Character role in the VN or release	
Producer	Type	Producer type (e.g. publisher, developer)	
	Name	Producer name	
	Original	Producer name in its original language	
	Alias	Producer name aliases	
	Relations	Other related producers, e.g. parents, subsidiaries	Producer
	Language	Producer's primary language	
	Description	Producer description	
Staff	Name	Staff name	
	Original	Staff name in their original language	
	Gender	Staff gender	
	Language	Staff's primary language	
	Description	Staff description	
	Character	If a voice artist, name of the character voiced	Character

Tag	Name	Tag name	
	Alias	Tag name aliases	
	Parent	Hierarchical parent tags	Tag
	Description	Tag description	
Trait	Name	Trait name	
	Alias	Trait name alises	
	Description	Trait description	
	Parent	Hierarchical parent traits	Trait
	Sexual	Indicates a sexual tag for filtering purposes	

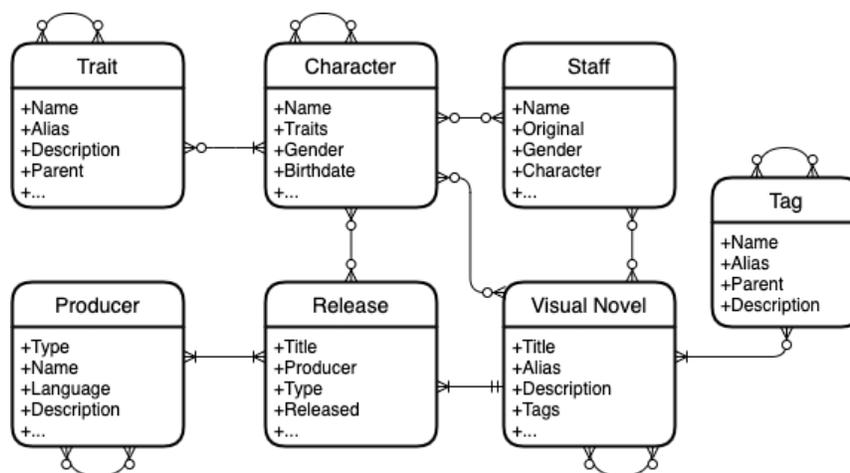


Fig 1. Entities, relationships, and a sample set of properties for the Visual Novel Database dataset.

Visual Novel. Visual novels (*VNs*) are the central entity type in the VNDB. Generally, these represent the conceptual level of the game, similar to the Work entity in the FRBR model[1], distinct from individual releases or versions. These entities also act as a hub connecting most other entities; *VN* pages feature descriptive data for the visual novel itself, character and staff rosters, and a list of releases. These connections are reflected in the entity relationships, which include *Staff* who worked on the *VN*, descriptive content *Tags*, individual {Releases}, *Character* appearances, and other related *VNs*.

Release. *Release* entities represent individual, distinct releases of a *VN*. Every *VN*, even if still unreleased, has at least one *Release*, which can be physical or digital. Releases contain release-specific information, such as a bar code and release date, connections to responsible agents (*Producers*), and technical details such as screen resolution and game engine used. Translation or other patches are also listed as a distinct *Release*. *Releases* feature relationships between the *Producers*, the *VN* to which the *Release* belongs to, and included *Characters*.

Character. These entities contain data for characters featured in *VNs*. Primary contents are physical and personality descriptions, and relationships to other entities, such as descriptive *Traits*, *VNs* or *Releases* the *Character* appears in, other instances of the *Character*, and *Staff* that may have voiced the *Character*.

Producer. *Producer* entities are for corporate bodies or groups responsible for some part of the creation or publication of a *VN*. Multilingual aliases, a variety of relevant external links, and connections to related *Producers*, and *VNs* / *Releases* which a *Producer* has contributed to, are found here.

Staff. Individual staff members are described using this entity. This also contains connections to *VNs* they have contributed to and their role (e.g. character artist, scenario writer) in that contribution, as well as *Characters* they have voiced, if the individual is a voice artist.

Tag. *Tag* entities hold descriptive data for the content of *VNs*. In the VNDB, there is a robust tag hierarchy, with several broad groups (e.g. theme, style, setting), each of which contains a number of individual tags and child tags. A single *Tag* entity will list a small description, occasionally an external link to a more descriptive resource, normally Wikipedia, and aliases. The entity is connected to hierarchical child and parent *Tag* entities, and any *VNs* which have that *Tag* applied.

Trait. *Traits* are much like the *Tag* entities, but they act as descriptors for a *Character* rather than a *VN*. These descriptors are fairly granular, and identify the physical and personality traits of a given *Character*. Each entity page contains data similar to those found in the *Tag* entity pages, such as descriptions, aliases, and connections to hierarchically related *Traits*, and any *Character* which has been identified as having that *Trait*.

[1] IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional Requirements for Bibliographic Records, <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>. Last accessed 8 June 2021