

Coding gender for characters in video games

Introduction

The aim is to code the (Western) player community's interpretation of the gender of each character. Any label can be used as a gender category. Letters should be lowercase. Some conventions help the unification of the data later on. If you assign a character to the male category, use the label "male" (rather than "man", "boy" etc.). If you assign a character to the female category use the label "female" (rather than "woman", "lady", "girl" etc.). Characters can be coded as having no gender (e.g. "genderless"), but only with positive evidence.

Characters can belong to multiple groups. For example, a character might belong to both the groups "female" and "trans".

You might assign characters to the "neutral" category if:

- They are supernatural or cosmic beings, or very different species, or inanimate object for which interpreting gender using the indicators below is very difficult.
- The name refers to a group (e.g. "people") or general species (e.g. "Chocobo") which might include more than one gender.
- They have a generic name (e.g. "Guard", "Villager") AND their gender isn't readily identifiable from the indicators below. This includes where it isn't possible to identify the specific character. They are an invisible third person narrator, especially in text-based adventures. (Although sometimes the gender of the narrator can be determined and should be coded accordingly, for example where they are a named character – such as Varric in Dragon Age 2 – or where there is sufficient indication of their gender from their voice/pronouns used etc.)
- The dialogue comes from the game system not tied to a character, e.g. in tutorials.

Indicators

There are various indicators used to code the gender of the game characters. Some are more authoritative than others, but any individual indicator is defeasible. Usually several indicators are required to make a decision.

1. Category defined by community wiki

Some games have community-written wikis or discussion boards with pages for each character. Some of these define the gender of the character (e.g. https://finalfantasy.fandom.com/wiki/Cloud_Strife under "Gender" in the info bar on the right). The community wiki categorisation should not necessarily be taken as the gold standard. Firstly, the category can be assigned erroneously (as appears to be the case for some minor characters), debatable, or change according to the language/release of the game. For example, in Mass Effect, the Asari species is, according to the game lore, a mono-gender species. However, all Asari characters have visual cues, pronouns, and voice actors which suggest players will confer female gender to them.

Secondly, coding on wikis are likely based on a set of gender indicators such as the ones listed below. So the information in wikis can be seen as 'shorthand' for a gender coding that takes into account many factors. However, in most cases, these will agree with the coding scheme presented here.

2. Character name

The gender of some characters is revealed in their name, e.g. "King Graham". However, the names in the corpus data are sometimes given by the transcribers rather than the game, so other indicators may be needed when coding characters like "Adolescent Male #1".

3. Appearance

Visual appearance may provide clear cues to gender, including visual signifiers (e.g. pink bows or makeup are often restricted to female characters) or secondary sex characteristics (e.g. breasts). This is not to suggest that all and only females wear makeup, or have specific secondary sex characteristics. Their use in a game, however, is frequently a choice by developers in order to indicate that such characters are not male. Some wiki pages have higher-resolution fan art for characters (though these may not be reliable).

4. Claims that the character themselves make in dialogue Characters rarely announce their gender directly, but still might bring it up in conversation. E.g. in King's Quest Chapter 1:

Amaya: "I'm a woman of action, if you couldn't tell."

Or in Stardew valley:

"Abigail": "Oh, it's because I'm a girl... isn't it? Ugh..."

5. How other game characters refer to them

Including pronouns, but this can be unreliable and switch between translations. E.g. https://finalfantasy.fandom.com/wiki/Quina_Quen

6. Pronouns used in community wikis

In a wiki page about a character, they may be referred to by a specific pronoun that might give you a clue.

7. Gender of the voice actor

If the voice actor is known, this might provide a clue to the gender, but may be quite unreliable.

Edge cases

For the majority of characters, coding is straightforward, as there are multiple indicators of gender present and in agreement. Any coding decisions in edge cases should be documented in the "README.md" files for each game. Difficult cases tend to fall into one of five categories:

1. Internal ambiguity

In a small range of cases, the gender of some characters is treated inconsistently within a game. Quina Quen from Final Fantasy IX falls into this category: Quina is referred to by characters in the English translation of the game as 's/he' but using the possessive pronoun 'his'. Narratively, Quina is a member of a genderless race called the Qu, but the game mechanics are inconsistent in their treatment of Quina, for instance, certain spells that only affect male characters affect Quina, and certain abilities that only affect female characters do not. This inconsistency affects players' experience of a character's gender (see e.g. discussions of Quina, [here](#) or [here](#)).

2. External disagreement

We are aiming to code a typical player's experience of characters' gender, but there are cases in which players disagree. For example, dialogue in Final Fantasy VII suggests that the character 'Big Bro' (also known as 'Beautiful Bro') is a cross-dressing man, but some players read Big Bro as a trans woman. In the Final Fantasy VII remake, Big Bro is replaced with a character called 'Jules' who the wiki describes as an effeminate man, but some players have interpreted as a trans man (e.g. Big Bro in Final Fantasy VII: [e.g. here](#)).

For cases of type (1) and (2) we looked for agreement between as many indicators as possible, and documented decisions in the readme.md files for each game.

3. Counting

Determining whether two or more character appearances are instances of the same character, or different characters, can be challenging. These cases tend to fall into one of two varieties. Firstly, there are generics: NPCs that share a character model and cannot be distinguished by name, voice acting, appearance or contextual clues. Soldiers and guards often fall into this category, but townsfolk and other background characters can as well. Where it is possible to consistently differentiate such NPCs we have done so using aliases (e.g. if two identical soldiers are guarding a gate, but one is on the left and the other on the right). Where it is not possible to do so, we have noted that the character is a generic rather than an unique individual in a game's readme.md file. The second variety is where characters appear in disguise

or have undergone a transformation changing their visual appearance. Depending on contextual clues, and in particular, what the player can be expected to know, we either consider the different versions of a character a single person, or multiple people. For instance, in King's Quest VI, a genie appears to the player in different forms that vary by age and gender. These interactions occur before the player has met the genie in their true form, and the player is unlikely to realise as they experience these disguises that they belong to a male genie. In this case we coded each disguise of the genie as a separate character. By contrast, a character who cross-dresses in a scene but is easily recognisable (such as Cloud in Final Fantasy VII) is not recorded as a separate character. Decisions and their rationale are documented in the readme.md file for each game.

4. Gendered objects

Some video game characters do not have a physical form usually associated with having a gender, such as talking books or vending machines. However, a game may nonetheless treat such a character as gendered, and players confer gender on them, based on their voice acting, the pronouns other characters use to describe them, and so forth. Such cases are documented in the readme.md file for each game.

5. Insufficient evidence

Sometimes there is insufficient evidence available to categorize the gender of a character. This most often occurs for very minor characters, that appear off-screen or in the distance, and are unvoiced. However there are exceptions, particularly with characters from fantastic races or with unusual physical features. For instance, the 'bangaa' species from Final Fantasy XII are not sexually dimorphic, and nothing about their physical appearance is indicative of their gender, including their clothes and accessories. Some bangaa are indicated to be male or female in gender through the use of pronouns and voice acting, but most of the bangaa NPCs are unvoiced. Where insufficient evidence was available, characters were coded as 'neutral'.