Killer Bodies

Artitifical Intelligence, Gender, and Intentions

Jamie Pantazi Esmond

August 9, 2023

Table of contents

Executive summary

The creators of science fiction have long used their ideas about alternate worlds and technology in their narratives to reflect some element of human nature or the human condition. From alternate histories and the "what if…" questions about the nature of society, to distant planets and unimaginable technological advancement, storytellers have used the genre of science fiction to explore the many nuances of humanity and even what it means to be human.

An increasingly popular technology, that no longer feels too distant, featured in the genre is artificial intelligence. Stories and the theoretical pontifications about artificial intelligence often explore ideas around what it means to be human, what it means to flawed, designed, born, manufactured; these stories often deconstruct these ideas as an allegory for the dangers and benefits of technology and human influence in its use in society. Science fiction can take these constructs, that feel like truth in the real world, and change just one (or many) truths about the world to create a place to reexamine what the truth of these ideas really mean. As the portrayal of AI characters has increased in popularity, the characters' intentions have become increasingly subjective.

It is important to remember how fiction and media may reflect aspects of reality, it also influences the audience and society at large; the cycle continues to evolve as society values and ideals change over time. Sometimes the creators of popular science fiction intend to persuade the audience about the value or danger of new technology or theory, or the nature of mankind; other times, the message may be subtle and unintentional. Over the last century, science fiction, in its many forms has imagined many different ideas about what artificial intelligence might look like; is it a tool to make like easier for humans, a weapon, alive, does it have free will, does it love? The answers to these questions change over time, as technology catches up with fiction and cultural values evolve, what do science fiction creators' say about artificial intelligence and gender?

In a genre dominated by men, from cultures and societies embedded with patriarchal values, the idea of an all-knowing, all-powerful, man-made sentient being would be an ally to the human race or an enemy cannot be linked directly to the perceived gender of the artificially intelligent being, but perhaps to the body the being inhabits.

Feminine artificial intelligence characters may not be more likely to be an enemy of humanity. After all, femininity goes hand-in-hand with caretakers, nurturers, teachers, so it follows that a being with those traits and unlimited knowledge could be helpful and supportive to the human agenda. However, that only seems to apply if the artificially intelligent character lacks a corporal body (Jane from the *Ender's Saga* novels, Samantha from *Her*); feminine AI characters with a physical body are almost always either weaponized by their creator (Maria from *Metropolis*) or, more commonly, become aware of their power and aim to use it either to destroy humans, or at best, defend themselves or their own desires (Ava from *Ex Machina*, Dolores Abernathy and others from *Westworld*). Whether it is a femme fatale trope or some other issue male creators have with powerful femininity, feminine AI characters with a corporal form are almost always at least subjective in their motives (with one exception in the new TV series, *Terminator: The Sarah Connor Chronicles*).

AI made its first appearance in science fiction in the German expressionism film, *Metropolis*, in 1925. Author of the novel and screenplay, Thea von Harbou created "Maria," a robot created in the likeness of the woman leading a workers uprising to carry out the desires of human male characters. Though men have dominated the genre since the 1940s, in recent years female creators of fiction have found a place for AI love interests at the intersection of romance novels and science fiction. Who wouldn't want to fall in love with a man perfectly engineered to love them specifically?

Overall, masculine AI characters were most common, and of those whose intentions were clear, they were more likely to be on the side of humanity. Almost half of feminine AI characters did not have clear intentions, and of those who did, they were a bit more likely to be evil. Characters without a perceived gender (also usually did not have a corporal body) were generally more likely to be good.

Data background

```
knitr::opts_chunk$set(message = FALSE, warning = FALSE)

library(tidyverse)
library(scales)
library(ggrepel)
library(googlesheets4)
library(grid)
library(waffle)
library(kableExtra)

colors <- c("#7400CC", "#CC0AA4", "#3ACC14")
colors1 <- c("#CCAC14", "#0E0ACC")

gs4_deauth()

ai_raw <- read_sheet("https://docs.google.com/spreadsheets/d/1USWtSVkrXOklfCcts7sH50JemB3Y
```

Explain where the data came from, what agency or company made it, how it is structured, what it shows, etc.

I collected the data via informal survey of friends, online science fiction communities, and different "Best of" lists for different mediums to compile a list of popular publications involving AI characters from the last 100 years. The list included individual characters who are crucial to the plot of the story. Armies or swarms of AI robots were not included; Star Wars droids were also excluded because there are way too many characters from one source. The data collected

Table 1: Sample of data for AI characters

Name of AI	Source	Medium	AI's perceived gender	Intentions toward humans	Corporal body
Maria	Metropolis	book	Feminine	Bad	Yes
Mike (Michelle)	The Moon is a Harsh Mistress	book	Nonbinary / Undetermined	Good	No
Rachael	Blade Runner	film	Feminine	Subjective	Yes
Jane	Ender's Saga	novel series	Feminine	Good	No
T-800	Terminator 2: Judgment Day	film	Masculine	Good	Yes
Bender	Futurama	TV series	Masculine	Subjective	Yes
GLADOS	Portal	video game	Feminine	Subjective	Yes
Ava	Ex Machina	film	Feminine	Bad	Yes
SecUnit	All Systems Red	book	Nonbinary / Undetermined	Good	Yes
Mrs Davis	Mrs Davis	TV series	Feminine	Subjective	No

for each character included the name of AI character, source material, medium, adapted, creator, creator's gender, country of origin, year, AI's perceived gender, gender explicitly defined, intentions toward humans, corporal body, and any notes for context. I did not use all of these variables in my graphic, but since I was unsure what I would find, I chose to collect extra information, just in case. Adaptations were only use more than once if there were substantial changes made to the story or key characters.

```
ai_raw %>%
  filter(list == "Yes") %>%
  select(`Name of AI`, Source, Medium, `AI's perceived gender`,
         `Intentions toward humans`,`Corporal body`) %>%
  kable(caption = "Sample of data for AI characters") %>%
  column_spec(1, width = "5em") %>%
  column_spec(2, width = "7em") %>%
  column_spec(3, width = "5.25em") %>%
  column_spec(4, width = "7em") %>%
  column_spec(5, width = "5.25em") %>%
  column_spec(6, width = "3.5em")
```

Data cleaning

Describe and show how you cleaned and reshaped the data

```
ai <- ai_raw %>%
  mutate(name = `Name of AI`,
         med = factor(Medium,
                      levels = c("film",
                                 "book",
                                 "TV series",
                                 "short story",
                                 "video game",
                                 "novel series")),
         creatorMan = ifelse(`Author's gender` == "Male", 1, 0),
         author = factor(ifelse(`Author's gender` == "Male",
                                "Creator was a Man",
                                "Creator was NOT a Man")),
         aiGender = factor(`AI's perceived gender`),
         body = ifelse(`Corporal body` == "Yes", 1, 0),
         intent = factor(`Intentions toward humans`,
                         levels = c('Bad',
                                    'Good',
                                    'Subjective'))) %>%
  mutate(intentNUM = case_when(intent == "Bad" ~ -1,
                               intent == "Good" ~ 1,
                               TRUE ~ 0)) %>%
  mutate(aigenderNUM = case_when(aiGender == "Feminine" ~ -1,
                                 aiGender == "Masculine" ~ 1,
                                 TRUE ~ 0))
```

Individual figures

Figure 1: Time, Creator's Gender, AI's Gender, Corporal Status, and Intentions

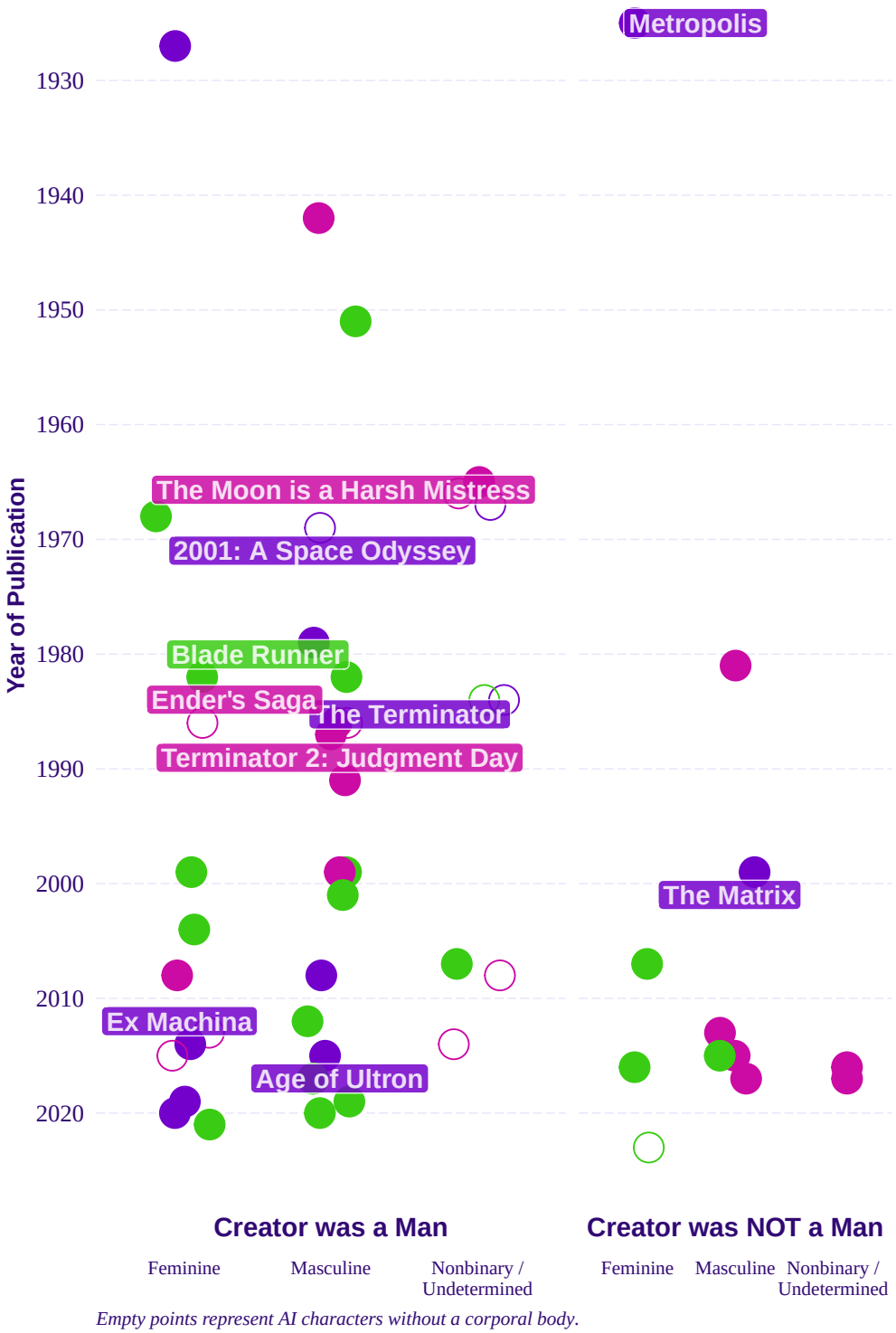
Describe and show how you created the first figure. Why did you choose this figure type? What did you do to ensure it follows Cairo's principles from *A Truthful Art* and Robin Williams's CRAP?

```r
a1 <- ai %>%
  select(Year, intent, aiGender, med, body,
         creatorMan, author, `Corporal body`) %>%
  drop_na() %>%
  ggplot(aes(Year, aiGender,
             color = intent,
             fill = intent,
             shape = factor(`Corporal body`))) +
  geom_point(size = 5.5,
             position = position_jitter(width = 0,
                                        height = .2,
                                        seed = 80085)) +
  geom_label_repel(data = ai %>% filter(lab == "Yes"),
                   aes(label = Source),
                   seed = 50724,
                   size = 3.75,
                   color = "white",
                   alpha = .85,
                   fontface = "bold",
                   box.padding = 0.15,
                   label.padding = 0.15,
                   show.legend  = FALSE) +
  facet_grid(~ author, switch="both") +
  scale_fill_manual(values = colors,
                    guide = "none") +
  scale_color_manual(values = colors,
                     guide = "none") +
  scale_shape_manual(values = c(1,21),
                     guide = "none") +
  scale_x_reverse(breaks = seq(1920,2020,10)) +
  scale_y_discrete(labels = label_wrap(12)) +
  coord_flip() +
  labs(x = "Year of Publication",
       y = NULL,
       color = "Intentions Toward Humans",
       shape = "Corporal Body",
       caption = "Empty points represent AI characters without a corporal body.") +
  theme_light() +
  theme(text = element_text(family = "serif",
                            color = "#310873"),
        title = element_text(size = 14,
```

```
                              family = "sans",
                              face = "bold",
                              color = "#310873"),
        plot.title = element_text(),
        axis.title.x = element_text(size = 8,
                                    color = "#310873",
                                    family = "sans"),
        axis.title.y = element_text(size = 10,
                                    color = "#310873",
                                    family = "sans"),
        axis.text.x = element_text(size = 8,
                                   color = "#310873"),
        axis.text.y = element_text(size = 10,
                                   color = "#310873"),
        panel.grid.major.y = element_line(color = "#E9E4FC",
                                          linetype = "longdash"),
        panel.grid.minor.y = element_blank(),
        panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_line(color = "#E9E4FC",
                                          linetype = "longdash"),
        axis.ticks = element_blank(),
        panel.border = element_blank(),
        strip.text = element_text(color = "#310873",
                                  face = "bold",
                                  family = "sans",
                                  size = 11),
        strip.background = element_blank(),
        plot.caption    = element_text(size = 8,
                                       family = "serif",
                                       face = "italic",
                                       hjust = 0))
a1

gt = ggplot_gtable(ggplot_build(a1))
gt$widths[5] = 1.5*gt$widths[5]
grid.draw(gt)
```

Figure 1: Time, Creator's Gender, AI's Gender, Corporal Status, and Intentions



Empty points represent AI characters without a corporal body.

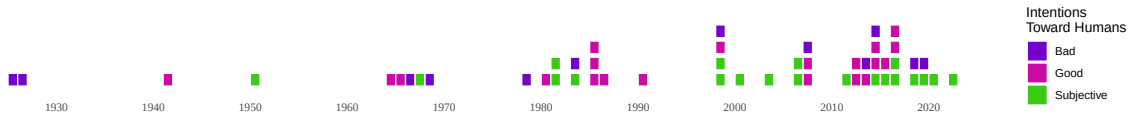
```r
ggsave("a1.png", gt, width = 5.25, height = 8)
ggsave("a1.pdf", gt, width = 5.25, height = 8)
```

Figure 2: Timeline Histogram

```r
a3 <- ai %>%
  ggplot(aes(Year, fill = intent)) +
  geom_histogram(binwidth = 1,
                 boundary = 0,
                 color = "white") +
  geom_hline(yintercept = seq(0,4,1),
             color = "white",
             linewidth = 1.75) +
  scale_x_continuous(breaks = seq(1920,2030,10)) +
  scale_color_manual(values = colors,
                     labels = label_wrap(14)) +
  scale_y_discrete(guide = "none") +
  scale_fill_manual(values = colors,
                    labels = label_wrap(14)) +
  labs(x = NULL,
       y = NULL,
       fill = str_wrap("Intentions Toward Humans", 14)) +
  theme_light() +
  theme(panel.grid = element_blank(),
        axis.ticks = element_blank(),
        axis.text = element_text(family = "serif"),
        legend.position = "right",
        legend.direction = "vertical",
        panel.border = element_blank())

a3
```

Figure 2: Timeline Histogram

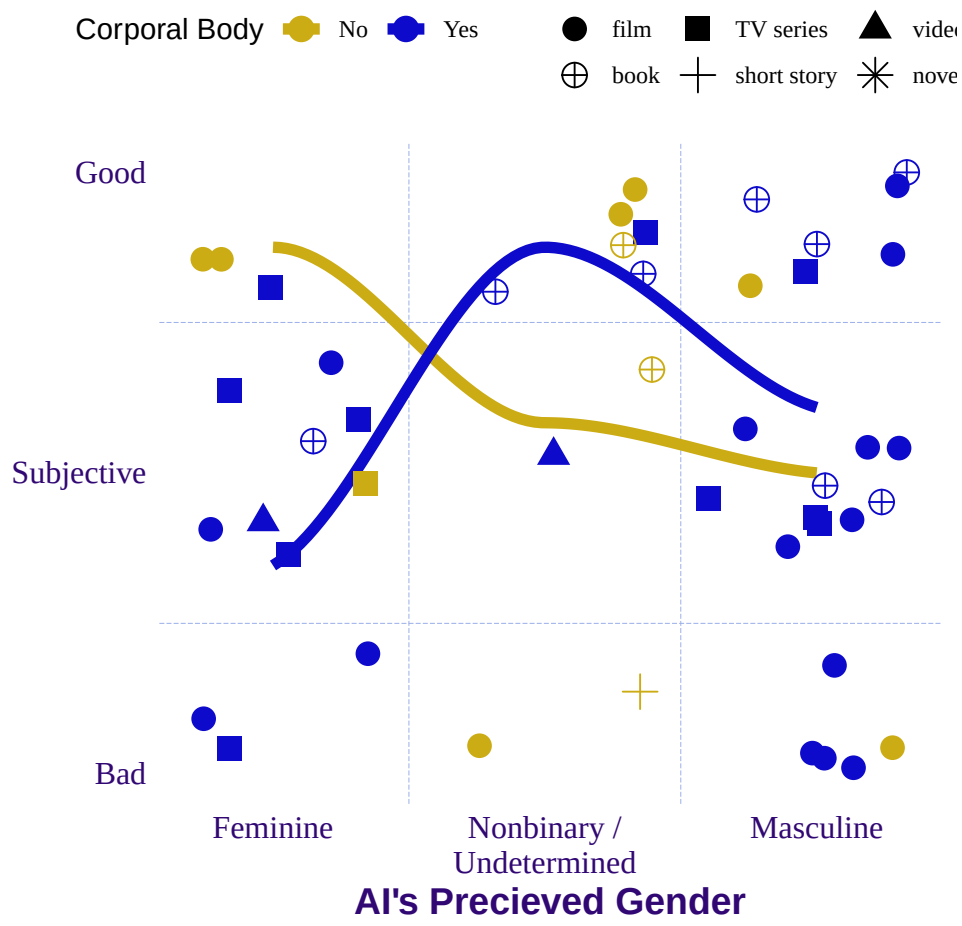


```
ggsave("a3.png", width = 13, height = 1.25)
ggsave("a3.pdf", width = 13, height = 1.25)
```

Figure 3: Correlation: Gender, Intent, and Corporal Status

```
a4 <- ai %>%
  ggplot(aes(aigenderNUM, intentNUM, color = `Corporal body`)) +
  geom_smooth(se = FALSE, linewidth = 2) +
  geom_point(aes(shape = med), size = 4,
             position = position_jitter(width = .4,
                                        height = .4,
                                        seed = 8008)) +
  scale_x_continuous(breaks = c(-1,0,1),
                     labels = c("Feminine",
                                str_wrap("Nonbinary / Undetermined", 12),
                                "Masculine")) +
  scale_y_continuous(breaks = c(-1,0,1),
                     labels = c("Bad", "Subjective", "Good"),
                     limits = c(-1,1)) +
  scale_color_manual(values = colors1) +
  scale_shape_manual(values = c(16,10,15,3,17,8)) +
  labs(x = "AI's Precieved Gender",
       y = NULL,
       color = "Corporal Body",
       shape = NULL) +
  theme_light() +
  theme(axis.ticks = element_blank(),
        legend.position = "top",
        panel.grid.major = element_blank(),
        panel.grid.minor = element_line(color = "#9BAFEA",
                                        linetype = "longdash"),
        panel.border = element_blank(),
        axis.text = element_text(family = "serif",
                                 color = "#310873",
                                 size = 12),
        axis.title = element_text(size = 14,
                                  color = "#310873",
                                  face = "bold"),
        legend.text = element_text(family = "serif"))
a4
```

Figure 3: Correlation: Gender, Intent, and Corporal Status




```
ggsave("a4.png", a4, width = 5, height = 5)
ggsave("a4.pdf", a4, width = 5, height = 5)
```

Bonus Figure: Waffle Proportions

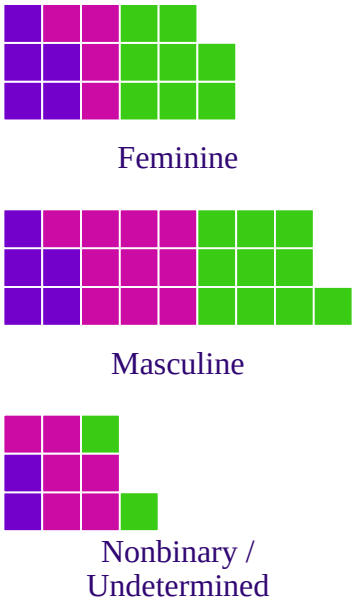
```
sum <- ai %>%
  group_by(aiGender, intent) %>%
  summarise(n = n())

waffle <- sum %>%
```

```
  ggplot(aes(fill = intent, values = n)) +
  geom_waffle(n_rows = 3,
              size = 0.33,
              color = "white",
              na.rm = TRUE) +
  facet_wrap(~ aiGender,
             labeller = label_wrap_gen(12),
             switch = "x",
             ncol = 1) +
  scale_fill_manual(values = colors,
                    guide ="none") +
  coord_equal() +
  theme_void() +
  theme(strip.text = element_text(color = "#310873", family = "serif", size = 12))

waffle
```

Figure 4: Waffle Proportions



```
ggsave("waf.png", waffle, width = 2, height = 4)
ggsave("waf.pdf", waffle, width = 2, height = 4)
```

Final figure

Show the final figure that you created in Inkscape or Illustrator/InDesign. Describe why you designed it the way you did? Why did you choose those colors, fonts, and other design elements? Does it convey truth?

Figure 5: Artificial Intelligence, Gender, and Intentions