

## NEXT 275: Sentient Beings from Folk Tales to AI

*Course premise:* After a Google engineer proclaimed its large language model LaMDA to be “sentient,” based on chat interactions, he was promptly fired. But the idea that generative AI may become capable of independent cognition, some form of consciousness, and therefore a degree of autonomy has led leading AI scientists, including “AI godfather” Geoffrey Hinton, to warn that advanced AI represents an “existential” threat to humanity. The idea of runaway AI eliminating humans is familiar from sci-fi through killer robots. But the idea of sentient nonhuman entities did not arrive with AI: an enchanted world was a fixture of 19th century romanticism, with inanimate objects able to speak and interact with human characters. The Grimms’ *Fairy Tales* feature personified, talking natural objects, human to animal transformations, and nonhuman helpers or antagonists. In the 19<sup>th</sup> and 20<sup>th</sup> centuries, sentient beings often played a critical role in children’s literature, horror stories, and sci-fi. The course explores narratives (in fiction, film, AI) portraying material, natural, or digital objects as sentient, able to interact with humans.

*Course info:* No cost; online asynchronous; instructor: Robert Godwin-Jones, rgjones@vcu.edu

.....

*Questions we will address (among others):*

- In German romanticism, frogs, trees, and rocks can talk, how do they differ from AI chatbots?
  - In fairy tales, non-human beings/objects often help human protagonists. In sci-fi, when AI turns sentient, it is almost always threatening to humanity. Why?
  - What are the parallels between folk tales and AI? For example: 1) unknowable sources/obscure transmission methods (anonymous authors, machine learning); 2) that both are based on patterns (AI’s training data/folk tale typologies); 3) or use of identity masking/impersonations/deep fakes
  - The world’s most famous linguist, Noam Chomsky, called ChatGPT “a lumbering statistical engine for pattern matching,” characterized by “amorality, faux science and linguistic incompetence.” Why is Chomsky so negative on AI’s language abilities?
  - Only carbon-based life forms are alive. Isn’t that blatant discrimination?
  - If ChatGPT is already someone’s “lover”, what happens when we have life-like AI robots?
- .....

*Likely readings include (novels will be excerpted):*

### 19<sup>TH</sup>-CENTURY

The **Brothers Grimm**, *Children's and Household Tales* (1812):

From unknown stories, “The Mouse, the Bird, and the Bratwurst” and “Straw, Coal, and Bean” to classics: “Rumpelstilzchen”, “Snow White”, and “Cinderella”

**German romanticism:** Tieck: “Blond Eckbert” (1797); Hoffman: “The Sandman” (1816)

Mary Shelley, **Frankenstein** or The Modern Prometheus (1818); Butler, **Erewhon** (excerpts; 1872)

### 20<sup>TH</sup> CENTURY

**Fantastic transformations:** Kafka, “Metamorphosis” (1915); Gustav Meyrink, “The Golem” (1913)

**Socialist fairy tales:** zur Mühlen, “Peter's Friends” (1921); Marshak, “Yesterday & Today” (1930)

**Science fiction:** R. Bradbury, “There Will Come Soft Rains” (1950); H. Ellison, “I Have No Mouth, and I Must Scream” (1967); P. K. Dick, *Do Androids Dream of Electric Sheep?* (1968)

**Cyberpunk:** William Gibson: *Neuromancer* (1984); Neal Sephenson, *The Diamond Age* (1995)

### 21<sup>ST</sup> CENTURY

**Contemporary:** Chiang, “The Lifecycle of Software Objects” (2010); Ishiguro, *Klara & the Sun* (2021)

*Possible film viewing/discussion:* *The Golem* (1920); *Metropolis* (1927); *2001: A Space Odyssey* (1968); *Blade Runner* (1982); *Terminator* (1984); *Wall-E* (2008); *Her* (2013); *Ex Machina* (2014)

.....