Artificial Fairy Tales:
An Experiment in Computational Narratology

Luca Giovannini (Potsdam/Padova)
Cecilia Graiff (INRIA Paris/Leipzig)

DALL-E 3 via Copilot

SRDH Doctoral Symposium
Genoa, 21.05.2024
Slides: plu.sh/dh-genova

Theoretical framework

“studying large language models as compression algorithms for human culture”, including literature

(Buttrick 2024)

LLMs in computational literary studies (CLS)

- ★ LLMs are already used in CLS as tools for annotation (e.g. Martens and De Greve 2023; see also Törnberg 2024), but can also be leveraged for running experiments on complex literary phenomena, like the passing of time in fiction (Underwood 2023) or suspense (Underwood 2024)
- ★ First theoretical reflections: Schröter 2023
  - AI as a "model of [a] rationalization of reading experience"
    - a new model of reader? (beyond Eco 1990's 'semantic reader / critical reader' dualism)
  - AI as a generator of counterfactual, yet plausible texts, which can be used to explore 'divergence aesthetics' ("Abweichungsästhetik")
    - experiments with generated drama and short stories (Giovannini and Skorinkin 2023, 2024) have already pinpointed some interesting quirks in LLM compositional strategies
- ★ Today: an experiment with fairy tales!

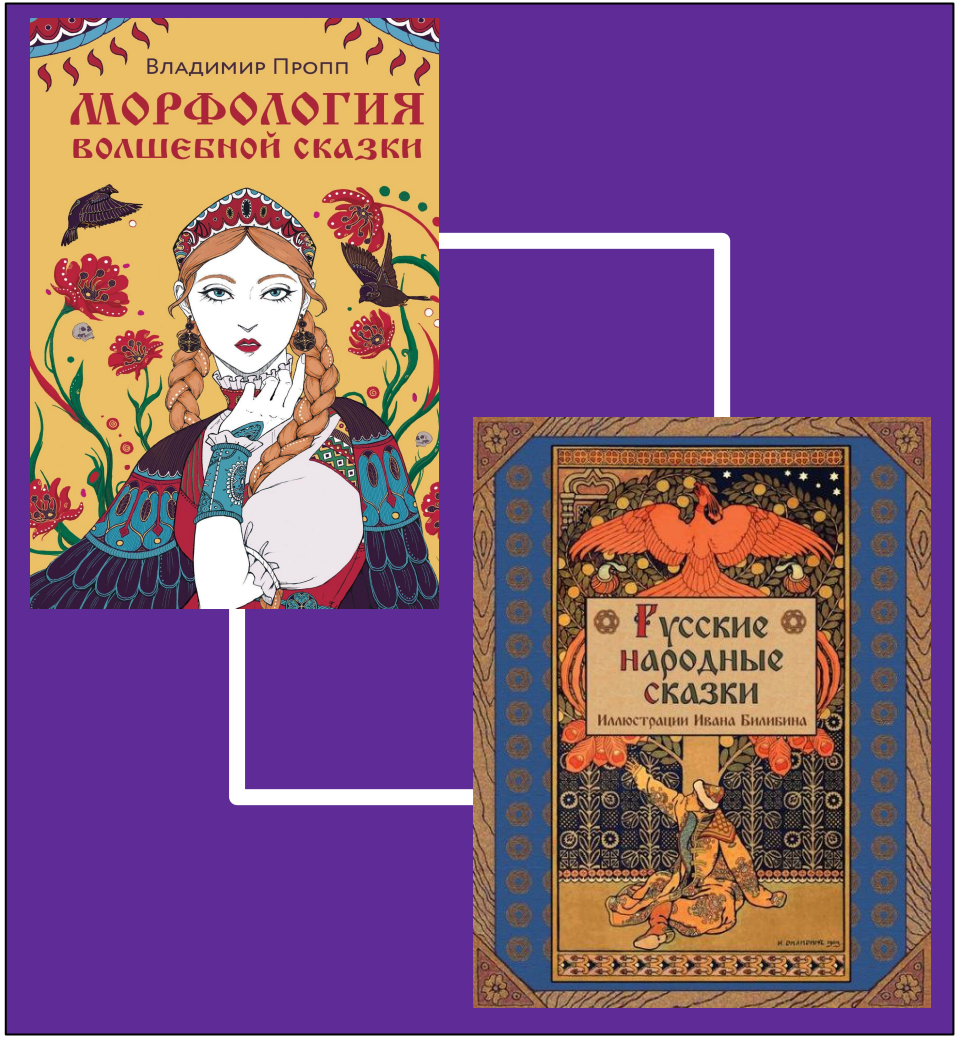
Research question

How do LLM write fairy tales?
Do they display idiosyncratic storytelling patterns?

Automatic fairy tale generation

- ★ Generally speaking, "[a]utomated story generation is the problem of mechanically selecting a sequence of events or actions that meet a set of criteria and can be told as a story" (Alhussain and Arwa 2021: 2)

- ★ Several (pre-LLMs) automatic story generation systems have focused on the creation of fairy tales and employed Vladimir Propp's *Morphology of the Folktale* (1928) as their theoretical background (Gervas 2013: 106).
- ★ Propp's work consists in the systematic analysis of formal patterns found in the collection of Russian fairy tales compiled by Alexander Afanasyev (1884).



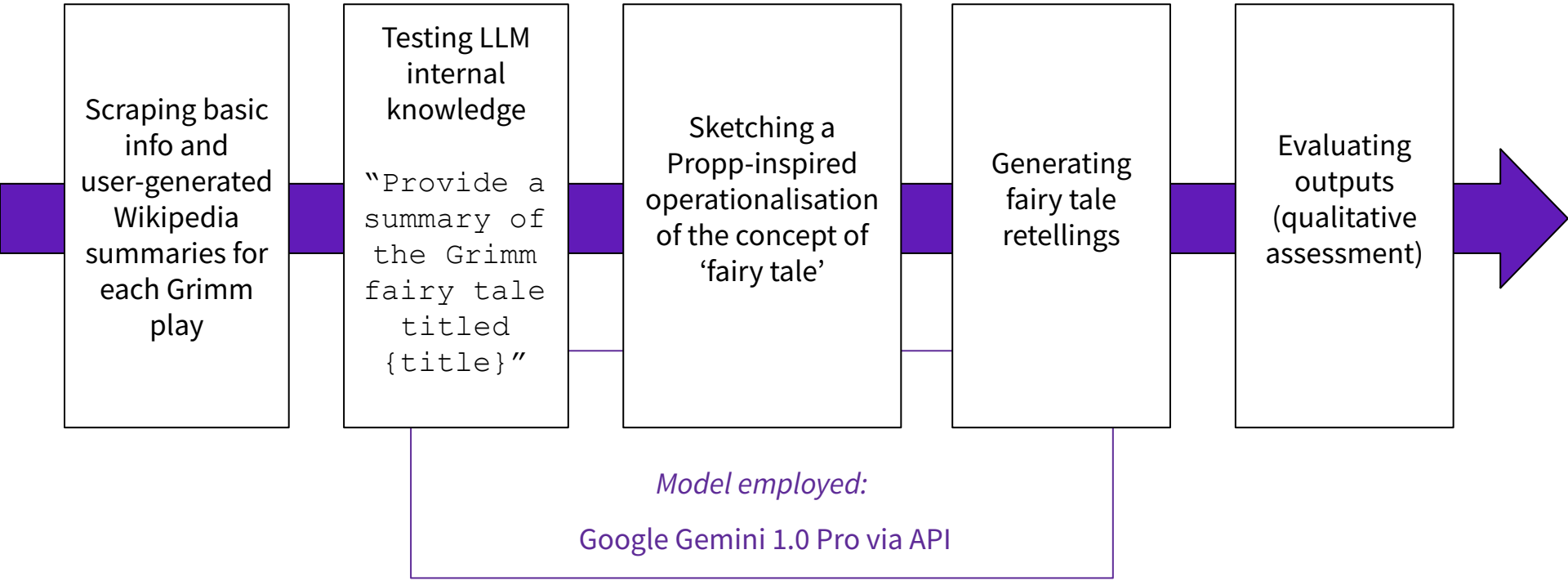
Our approach:

adapting Propp's formalisation to build a functional operationalisation of the concept of fairy tale, and use it to generate texts

Reference corpus:

Jacob and Wilhelm Grimm's *Children and Household Tales* (1812-57)

Experiment pipeline



Operationalising fairy tales

- ★ Operationalising the concept of story for computational work is a complex task, especially as far as plot is concerned (see e.g. Konle and Jannidis 2022)
- ★ As we saw, Propp's theoretical framework remains a popular reference for developing ontologies (Peinado 2004 et al.) for automatic story generation (Gervás et al. 2005)
- ★ Our approach: prototyping **a lightweight operationalisation** inspired by Proppian concepts and based on **three key components** we semi-automatically extract from Wikipedia summaries:

characters

+ their role according to Proppian taxonomy:

hero (protagonist), antagonist, donor, helper, sought-for person, anti-hero, dispatcher

settings

main locations where the action takes place

issues

~ Propp's **motivation**, i.e., "the reasons and the aims of personages which cause them to commit various acts" (Propp 1968: 70)

Prompt example for retelling generation

DALL-E 3 via Copilot

Write a short fairy tale using the following elements:

- **TITLE**: "Little Red Riding Hood"
- **CHARACTERS**: Little Red Riding Hood (hero), the Wolf (antagonist), Grandmother
- **SETTINGS**: the forest, Grandmother's house
- **ISSUES**: Little Red Riding Hood goes to visit her grandmother, the Wolf wants to eat both the grandmother and Little Red Riding Hood

	title	wikipedia url	human-written synopsis from wikipedia	characters	settings	issues	AI-written summary	AI-written retelling
0	allerleirauh	https://en.wikipedia.org/wiki/allerleirauh	a king promised his dying wife that he would not re-marry unless it was to a woman who was as beautiful as she was, and when he looked for a new wife, he realized that the only woman that could ma...	princess (hero), king (antagonist), another realm's young king (donor)	king's castle, forest, ballroom	the daughter's escape from her father's marriage proposal leads her to disguise herself. the young king's discovery of her identity changes their fates.	Allerleirauh A beautiful princess is forced to flee her stepmother's wrath after her father dies. She disguises herself as a scullery maid named Allerleirauh, wearing magical garb: a golden cloak...	Allerleirauh Once upon a time, in a grand castle nestled amidst towering mountains, lived a beautiful princess named Allerleirauh. Her father, the King, was a cruel and selfish man who harbored a...
1	the beam	https://en.wikipedia.org/wiki/the_beam_(fairy_tale)	a magician shows a rooster carrying a beam. a girl carrying a four-leafed clover realizes and declares that it is just a straw, not a beam. when she marries, the magician enchants her to think she...	magician (hero), girl (antagonist)	unknown	a magician performs a trick, a girl exposes him. the magician later takes his revenge on the girl.	In the Grimm fairy tale "The Beam," a poor girl named Trinchen (Catherine) is orphaned and must fend for herself. Desperate for shelter, she stumbles upon a mysterious house in the forest. Inside...	The Beam Once upon a time, in a realm unknown, there lived a cunning magician named Eldrin. With his captivating words and nimble hands, he performed extraordinary illusions, leaving audiences in...
2	bearskin	https://en.wikipedia.org/wiki/bearskin_(german_fairy_tale)	a man serves as a soldier, but when the war ends, he returns home to learn that both of his parents have died and that his brothers have no place for him.	soldier (hero), devil (antagonist). old man, man's three daughters	heath, old man's inn, old man's house	a soldier encounters the devil on a heath and strikes a deal for wealth and a cursed appearance. he helps a destitute old man and wins the promise of marriage from his youngest daughter.	Bearskin Once upon a time, a young man sets out to seek his fortune. He meets three old men who demand his bag, jacket, and hat. In exchange, they grant him three wishes: the ability to turn into...	The Soldier and the Devil's Bearskin Once upon a time, in a desolate heath, a weary soldier met the devil. Desperate for fortune, he struck a deal with him. The devil gave him a cursed bearskin t...
3	the blue light	https://en.wikipedia.org/wiki/the_blue_light_(fairy_tale)	a soldier has been discharged from the king's service because of his wounds. the soldier leaves the castle and, as night falls, he requires somewhere to stay. encountering the home of a witch, he ...	soldier (hero), witch (antagonist), spirit in the blue light (helper), princess, king,	castle, witch's home, well,	a discharged soldier seeks shelter at a witch's home but becomes trapped. the spirit in the blue light helps him escape and steal a princess away.	The Blue Light Once upon a time, there were three brothers who inherited their father's forge and a blue light that could conjure gold. The brothers alternated who would guard the light, but each...	The Blue Light Once upon a time, a weary soldier named Ethan roamed the countryside after being discharged from the king's army. Night fell, and he stumbled upon a desolate castle. Desperate for ...
4	the boots of buffalo leather	https://en.wikipedia.org/wiki/the_boots_of_buffalo_leather	a soldier, renowned for his braveness, has just been relieved from the army. he owns an old rain coat and boots in buffalo leather. during his voyage he meets a huntsman in the woods	soldier (hero), huntsman, old woman, robbers (antagonists)	woods, house (robbers' den), local town,	a renowned soldier, discharged from the army, encounters a huntsman in the woods. they stumble upon a robbers' den.	The Boots of Buffalo Leather Once upon a time, a poor soldier is gifted a pair of magical buffalo leather boots by an old woman. When he puts them on, he can run faster than the wind. He uses his...	The Boots of Buffalo Leather In a verdant woodland, a seasoned soldier, weary from the battlefields, stumbled upon a huntsman. Both lost in thought, they ventured deeper into the dense undergrowt...

Early insights (1/2): On the model's internal knowledge of fairy tales

★ n = 10 (*Allerleirauh, The Beam, Bearskin, The Blue Light, The Boots of Buffalo Leather, The Brave Little Tailor, The Bright Sun Brings It to Light, Brother and Sister, Cinderella, The Clever Little Tailor*)

★ in most cases, input is too generic to prompt correct summarisation (title and mention of the Grimms are not enough)

- later, the components guide the LLM to produces sometimes summaries instead of retellings

★ expected mismatch in performance according to the popularity of the fairy tale, e.g. *Cinderella* vs. *The Blue Light* (= 'head' or 'tail' entities in LLM knowledge, cf. Sun et al. 2024).

★ confusion between similar titled-stories (*The Brave/Clever Little Tailor*) and with more famous ones (*Brother and Sister* → *Hansel and Gretel*)

Early insights (2/2): On the model's retelling strategies

- ★ explicitation of implicit **story moral**: "*Their love story became a legend, reminding all that true beauty lies not in outward appearance but in the purity of one's heart*", "*proving that love conquers all, even the most sinister of curses*", "*even the smallest of beings can achieve extraordinary things with courage and determination*", etc.
- ★ **built-in moral alignment of LLMs** heavily influences their output (see in *The Bright Sun Brings It to Light*: "*a heinous crime*", "*Edmund's conscience gnawed at him relentlessly*", "*a path of justice and retribution*").
- ★ cross-genre influence from fantasy introduces **non-standard elements** in fairy tales, like proper names for characters (Ethan, Eldrin, Anya, Elara, etc.)

Issues and future improvements

Short-term:

- ★ improving generation pipeline (prompt refinement)
- ★ wrapping up qualitative assessment

Medium-term:

- ★ further automating components detection from the summaries
- ★ refining our operationalisation of the concept of fairy tales by embedding genre-specific markers (e.g. Propp's functions)
- ★ considering options for quantitative assessment of outputs
  - ○ full-text comparison (e.g. via *topic modelling*)?

References 1/2

- Alhussain, Arwa I., and Aqil M. Azmi (2021). “Automatic Story Generation: A Survey of Approaches”. *ACM Computer Surveys* 54, 5, 103: 1-38. DOI: https://doi.org/10.1145/3453156.
- Buttrick, Nicholas (2024). “Studying large language models as compression algorithms for human culture”. *Trends in Cognitive Sciences* 28, 3: 187-18. DOI: https://doi.org/10.1016/j.tics.2024.01.001.
- Eco, Umberto (1990). “*Intentio Lectoris*: The State of the Art”, in *The Limits of Interpretation*. Bloomington and Indianapolis: Indiana University Press, 1990, pp. 44-63.
- Giovannini, Luca, and Daniil Skorinkin (2023). "The AI Playwright: An Experiment in Literary Morphology". *Zenodo*. DOI: https://doi.org/10.5281/zenodo.10118816.
- Giovannini, Luca, and Daniil Skorinkin (2024). "Authorial Mimicry by Large Language Models: A First Assessment". *DH2024 Book of Abstracts* (forthcoming).
- Gervás, Pablo (2013). “Propp's Morphology of the Folk Tale as a Grammar for Generation”. *2013 Workshop on Computational Models of Narrative*, Open Access Series in Informatics (OASIcs) 32: 106-122. Dagstuhl: Leibniz-Zentrum für Informatik. DOI: https://doi.org/10.4230/OASIcs.CMN.2013.106.
- Gervás, Pablo, Díaz-Agudo, Belén, Peinado, Federico & Hervás, Raquel (2005). “Story plot generation based on CBR”. *Knowledge-Based Systems*, 18 (4): 235–242. DOI: https://doi.org/10.1016/j.knosys.2004.10.011.

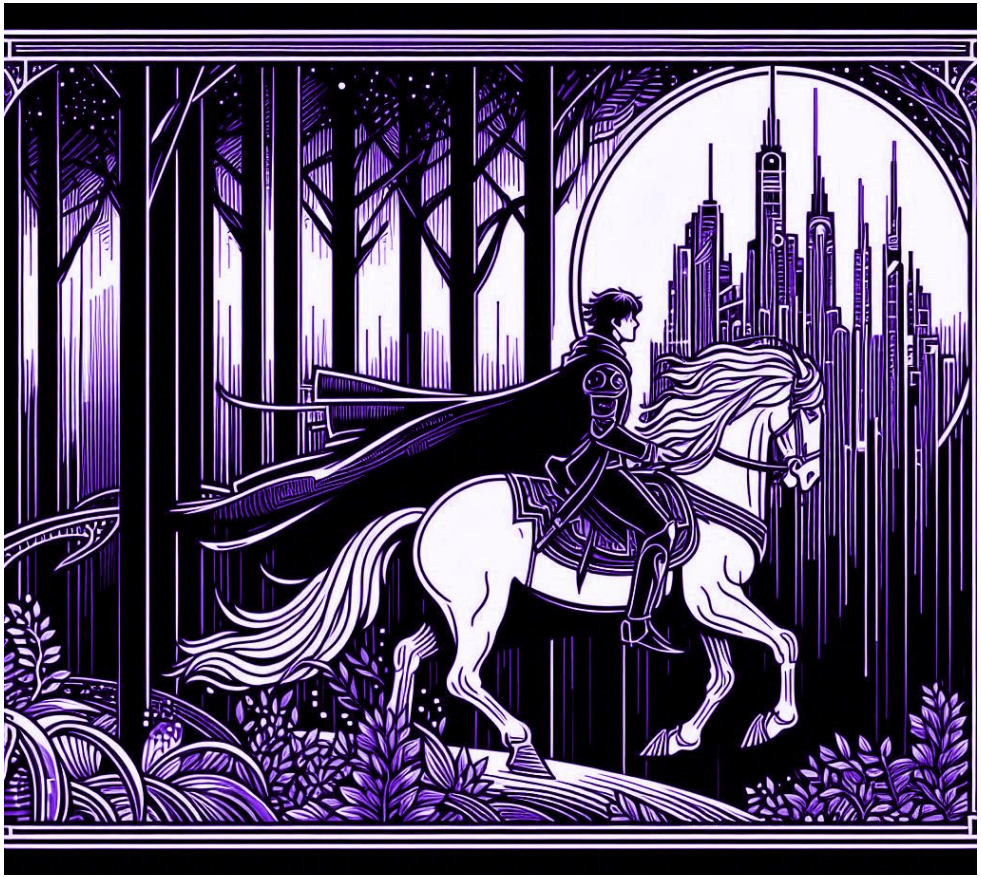
References 2/2

- Konle, Leonard, and Fotis Jannidis (2022). “Modeling Plots of Narrative Texts as Temporal Graphs”. *Proceedings of the Computational Humanities Research Conference 2023*, 3290: 318-336. URL: https://ceur-ws.org/Vol-3290/long_paper2313.pdf.
- Martens, Gunther, and Lore De Greve. “Annotation via LLM (Large Language Models) in Digital Literary Studies : Beating (about) the Bots?”. *DH Benelux 2023 Book of Abstracts* (forthcoming).
- Peinado, Federico, Pablo Gervás, and Belén Díaz-Agudo (2004). "A description logic ontology for fairy tale generation." *Proceedings of the Workshop on Language Resources for Linguistic Creativity*, 4: 56-61. URL: http://lrec.elra.info/proceedings/lrec2004/ws/ws13.pdf.
- Propp, Vladimir (1968). *Morphology of the Folktale*. Austin: University of Texas Press.
- Schröter, Julian (2023). “Generative Sprachmodelle und literaturwissenschaftliche Arbeitspraktiken”. *29. Workshop des Instituts für Allgemeine und Vergleichende Literaturwissenschaft, LMU München*. DOI: https://doi.org/10.5281/zenodo.8199796.
- Sun, Kai, Xu, Yifan Ethan, Zha, Hanwen, Liu, Yue, and Xin Luna Dong (2023). “Head-to-Tail: How Knowledgeable are Large Language Models (LLM)? A.K.A. Will LLMs Replace Knowledge Graphs?”. *arXiv*. DOI: https://doi.org/10.48550/arXiv.2308.10168.
- Petter Törnberg, Best Practices for Text Annotation with Large Language Models. *arXiv*. DOI: https://doi.org/10.48550/arXiv.2402.05129
- Underwood, Ted (2023). “Using GPT-4 to measure the passage of time in fiction”. *The Stone and the Shell*. URL: https://tedunderwood.com/2023/03/19/using-gpt-4-to-measure-the-passage-of-time-in-fiction.
- Underwood, Ted (2024). “Can language models predict the next twist in a story?”. *The Stone and the Shell*. URL: https://tedunderwood.com/2024/01/05/can-language-models-predict-the-next-twist-in-a-story.

Thanks for your attention!

✉ giovannini@uni-potsdam.de
✉ cecilia.graiff@inria.fr

↻plu.sh/dh-genova



DALL-E 3 via Copilot