

SADiLaR Digital Humanities Colloquium Series, 12 April 2023

MAKING STRANGE: CO-CREATING AFRIKAANS POETRY WITH A BOUTIQUE LANGUAGE MODEL



DR. IMKE VAN HEERDEN

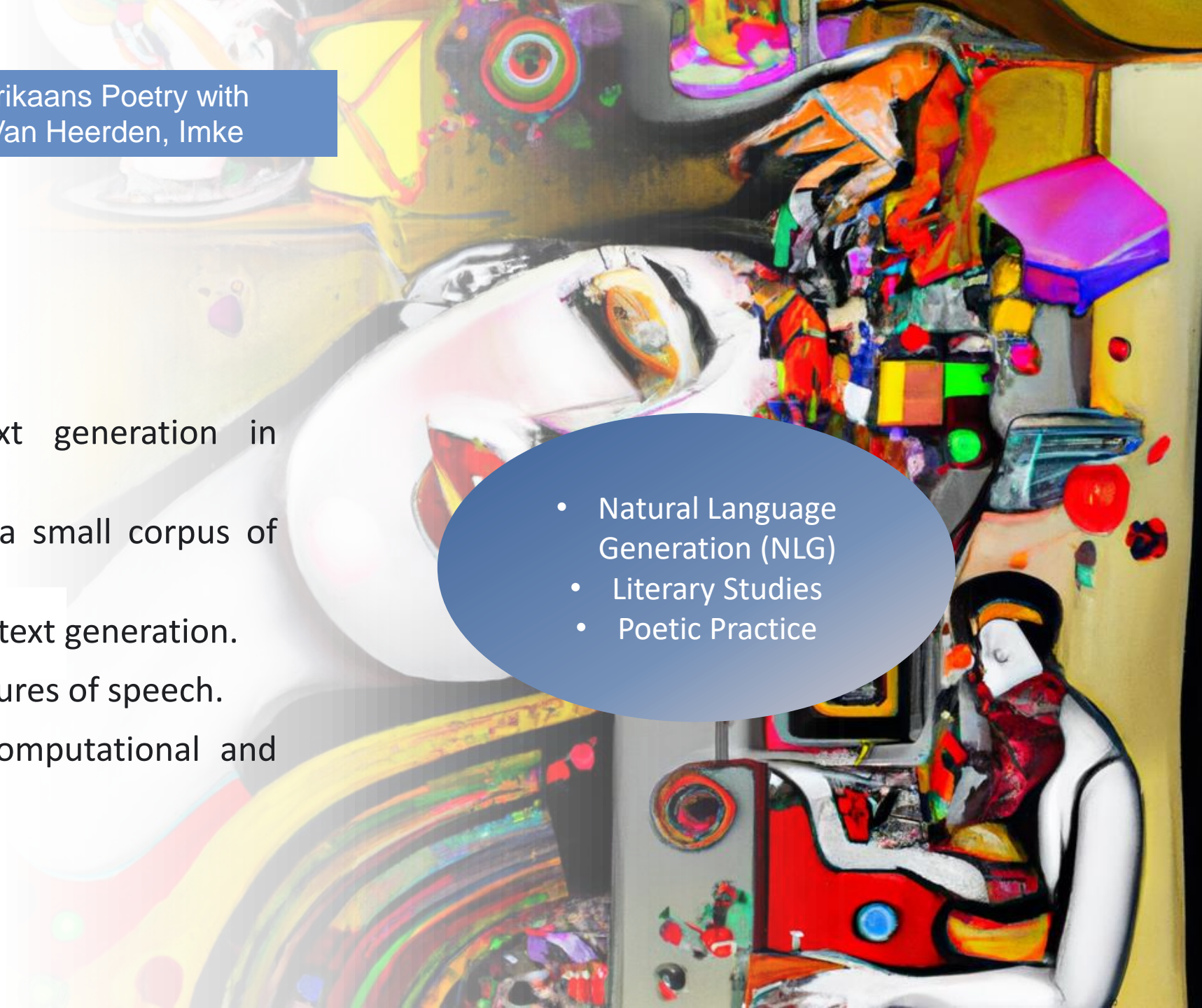
*Dept. of Comparative Literature
CSSH, Koç University, Istanbul*

✉ ivanheerden@ku.edu.tr

OVERVIEW

- First study on creative text generation in Afrikaans.
- LSTM architecture trained on a small corpus of contemporary fiction.
- Literary perspective on creative text generation.
- Results in compellingly novel figures of speech.
- Experimental integration of computational and poetic practices.

- Natural Language Generation (NLG)
- Literary Studies
- Poetic Practice



OVERVIEW

- First study on creative text generation in Afrikaans.
- LSTM architecture trained on a small corpus of contemporary fiction.
- Literary perspective on creative text generation.
- Results in compellingly novel figures of speech.
- Experimental integration of computational and poetic practices.

*Die lewe in hierdie nuwe hande
waar ek algoritmies kuier*

*The life in these new hands
where I socialise algorithmically*

CONTENTS

1. AI AS AUTHOR
2. AFRIKI
3. FIGURATIVE LANGUAGE
4. MODEL
5. RESULTS
6. TOWARDS EVALUATION
7. MACHINE-IN-THE-LOOP
8. CREATING A POETRY COLLECTION



MAIN SOURCES

Van Heerden, Imke, and Anil Bas. “AI as Author – Bridging the Gap Between Machine Learning and Literary Theory.” *Journal of Artificial Intelligence Research*, 71, pp. 175–189, 2021.

---. “Towards Figurative Language Generation in Afrikaans.” Abstract. In *Proc. Workshop on Research in Computational Typology and Multilingual NLP, Annual Conference of the North American Chapter of the Association for Computational Linguistics*, p. 10, 2021.

---. “AfriKI: Machine-in-the-Loop Afrikaans Poetry Generation.” In *Proc. Workshop on Bridging Human–Computer Interaction and NLP, 16th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 74–80, 2021.

Van Heerden, Imke with Anil Bas and Etienne van Heerden. *Silwerwit in die soontoe: Afrikaans se eerste KI-gedigte*. Cape Town: Naledi Publishers, 2023.



Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

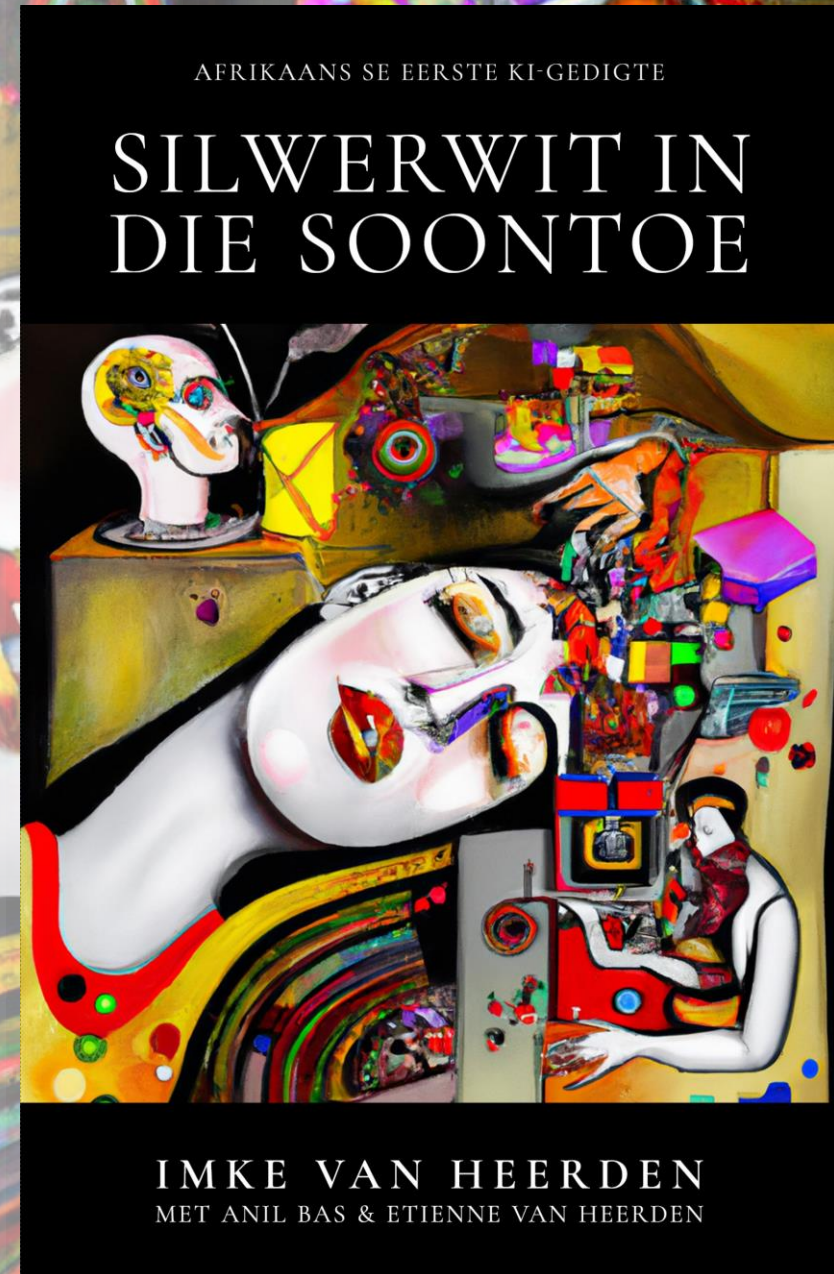
MAIN SOURCES

Van Heerden, Imke, and Anil Bas. “AI as Author – Bridging the Gap Between Machine Learning and Literary Theory.” *Journal of Artificial Intelligence Research*, 71, pp. 175–189, 2021.

---. “Towards Figurative Language Generation in Afrikaans.” Abstract. In *Proc. Workshop on Research in Computational Typology and Multilingual NLP, Annual Conference of the North American Chapter of the Association for Computational Linguistics*, p. 10, 2021.

---. “AfriKI: Machine-in-the-Loop Afrikaans Poetry Generation.” In *Proc. Workshop on Bridging Human–Computer Interaction and NLP, 16th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 74–80, 2021.

Van Heerden, Imke with Anil Bas and Etienne van Heerden. *Silwerwit in die soontoe: Afrikaans se eerste KI-gedigte*. Cape Town: Naledi Publishers, 2023.





PART 1

INTRODUCING AI AS AUTHOR

Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

AI AS AUTHOR

LITERATURE + MACHINE LEARNING

- *AI as Author: Bridging the Gap Between Computational Creativity Techniques and Literary Theory*
- Four-year project funded by the Scientific and Technological Research Council of Türkiye (2020–2023)
- Brings literary expertise to bear on text generation problems (e.g. Van Heerden & Bas, 2021a)
- Seeks to create a shared language between these disciplines
- **Website: authorai.ku.edu.tr**

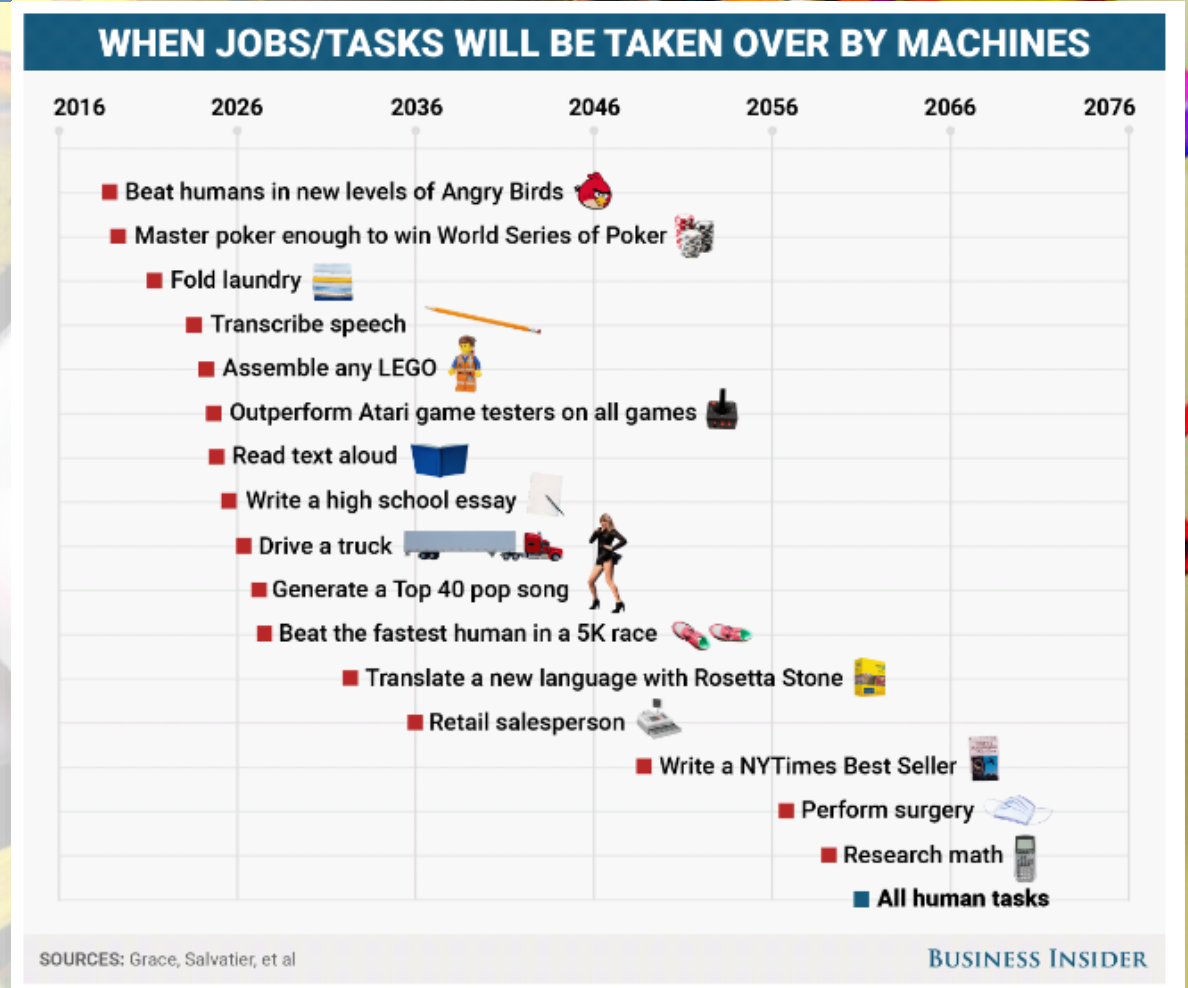


AI as Author Project Logo
authorai.ku.edu.tr

AI AS AUTHOR

MOTIVATION (VAN HEERDEN & BAS, 2021A)

- World Economic Forum: AI will write a New York Times best seller by the year 2049 (Hall, 2018)
- “When will AI exceed human performance?” (Grace et al., 2018)
- AI writing technologies are progressing rapidly
- Generation of jokes, metaphors, poems, essays, short stories, novels, screenplays, theatrical scripts, etc.
- Literary scholars?
- Can literary scholars help steer creative text generation research and, if so, how?



“Here’s when robots will start beating humans at every task”
(Weller and Gould, 2017; based on data from Grace et al., 2018)

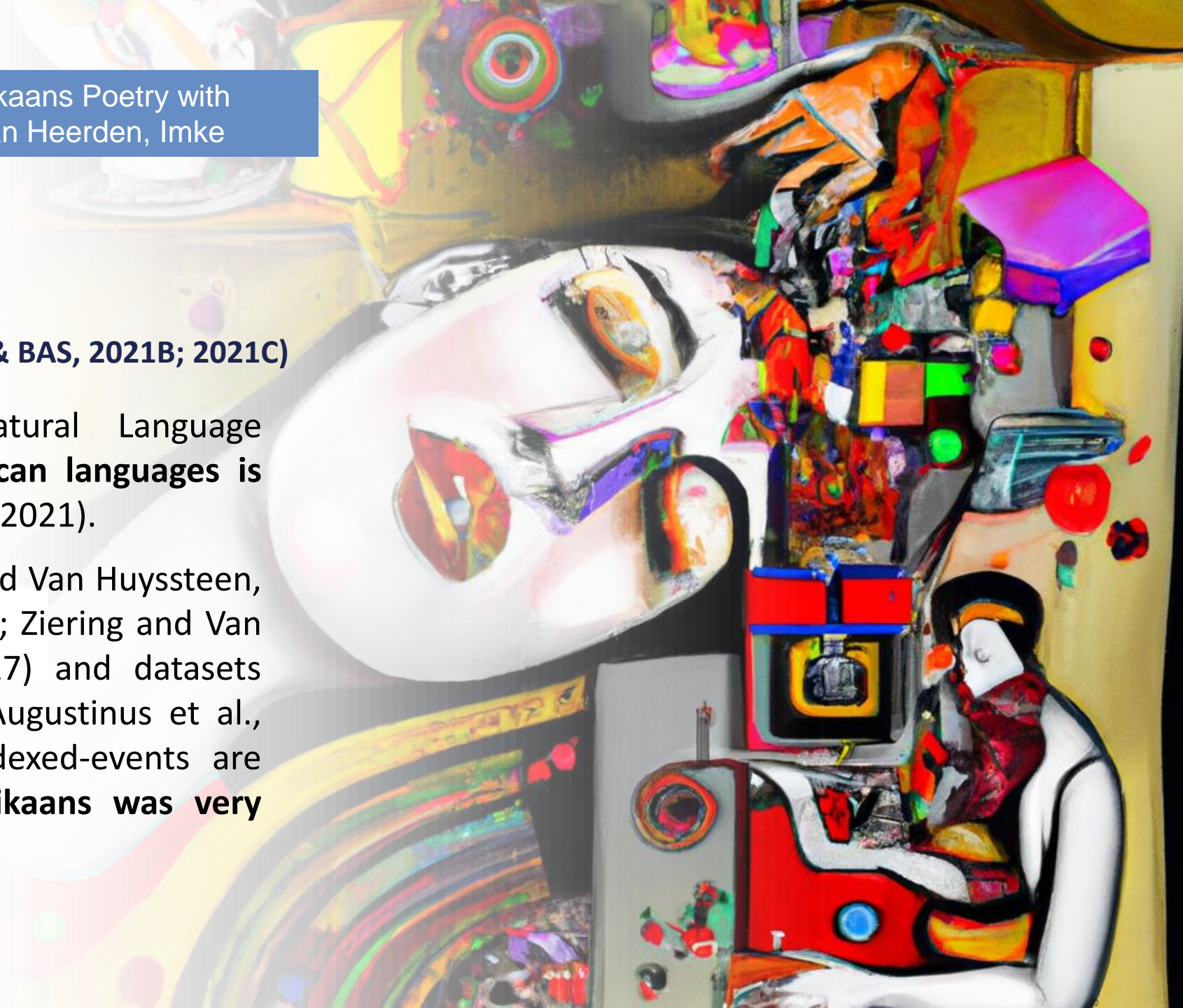


PART 2 *AFRIKI – AFRIKAANSE KUNSMATIGE INTELLIGENSIE*

AFRIKI

NOVELTY: LANGUAGE (VAN HEERDEN & BAS, 2021B; 2021C)

- Masakhane shows that Natural Language Processing (NLP) **research in African languages is under-represented** (2020a; 2020b; 2021).
- Although studies (Van Zaanen and Van Huyssteen, 2003; Sanby, Todd, and Keet, 2016; Ziering and Van der Plas, 2016; Dirix et al., 2017) and datasets (Eiselen and Puttkammer, 2014; Augustinus et al., 2016; Roux, 2016) from ACL indexed-events are available, **text generation in Afrikaans was very limited at the time.**



AFRIKI

LOW-RESOURCE LANGUAGES (VAN HEERDEN & BAS, 2021B; 2021C)

- ▣ Afrikaans is considered a **low-resource language** in the taxonomy of Joshi et al. (2020).
- ▣ NLP systems are currently dominated by a handful of languages (Joshi et al. 2020), and Afrikaans is one of many across the world **presently unable to match their progress**.
- ▣ **ACL 2022 Theme Track: *Language Diversity: From Low-Resource to Endangered Languages*** (ACL, 2022).
- ▣ Importance of representing **all languages**.





PART 3

FIGURATIVE LANGUAGE

Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

FIGURATIVE LANGUAGE

INTRODUCTION (VAN HEERDEN & BAS, 2021B)

- Definition: “Language that is more expressive and/or poetic than referential in its linguistic function” (Chandler and Munday, 2011)
- Metaphor, simile, personification, etc.



FIGURATIVE LANGUAGE

METAPHOR IN CREATIVE WRITING (VAN HEERDEN & BAS, 2021B)

- Vital resource (Baldick, 1996)
- Associated with **originality**
- Facilitates **connection** between reader and text (Cohen, 1978)
- **Unusual** metaphors stand out (Kovecses, 2010)
- Unexpected metaphors in literature **capture attention** (Steen, 1994)
- Original metaphors convey **greater emotional intensity** (Gibbs et al., 2002)
- Original metaphors are related to **creativity** and emotion (Fainsilber and Ortony, 1987; Fussell and Moss, 1998)



FIGURATIVE LANGUAGE

IMPLICATIONS

■ **Generated Poetry: Areas of Improvement**

(Van Heerden and Bas, 2021a)

- Emotional Impact
- Depth
- Originality

■ **Future Work** (Van Heerden and Bas, 2021b)

Does the inclusion of original metaphor enhance the perception of originality, creativity and emotion?

■ **Multifaceted**





PART 4

MODEL

Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

Architecture

LSTM layer with 50 units

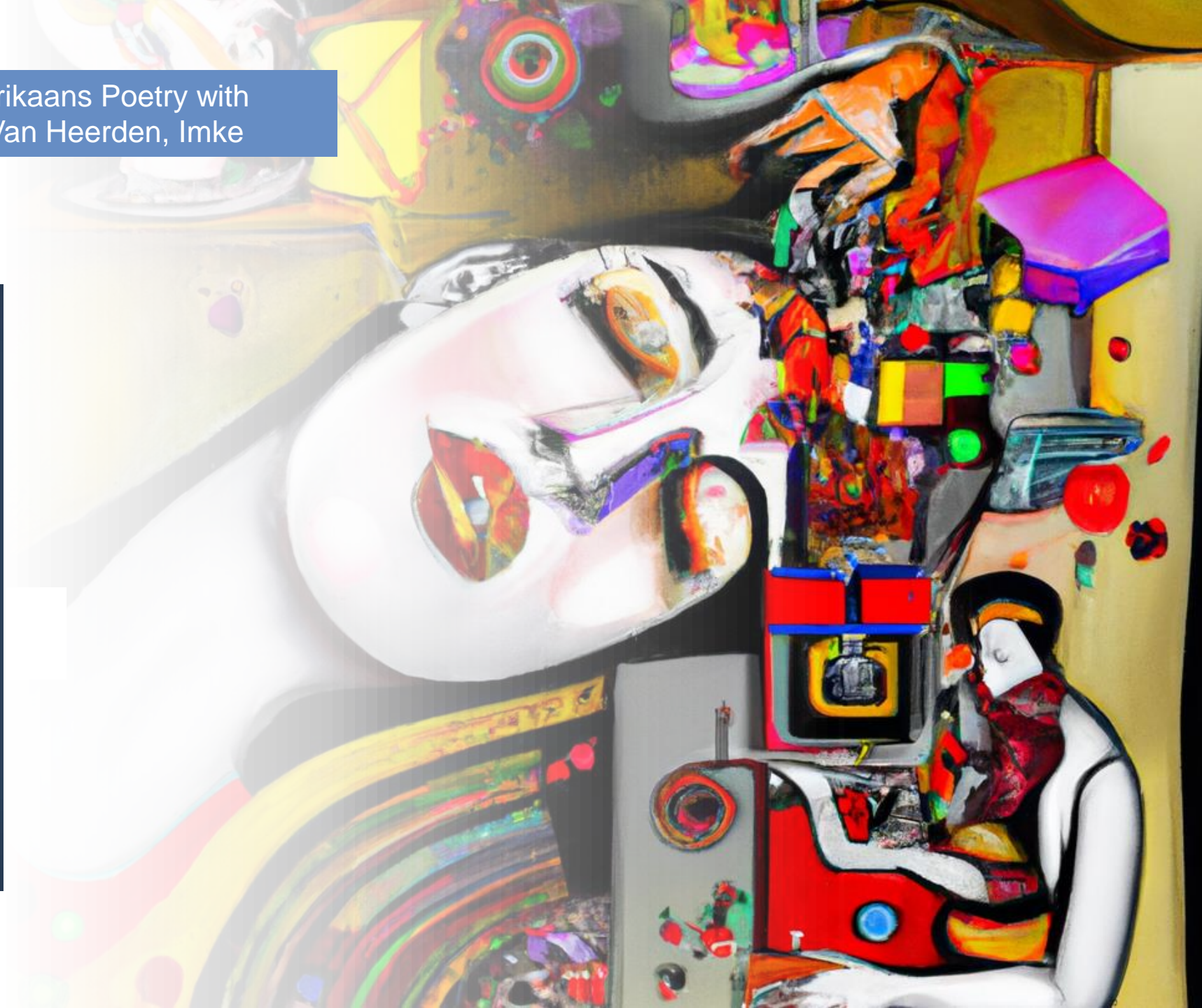
Dropout layer

LSTM layer with 50 units

Dropout layer

Fully connected layer

Softmax layer



DATASET

(VAN HEERDEN & BAS, 2021B; 2021C)

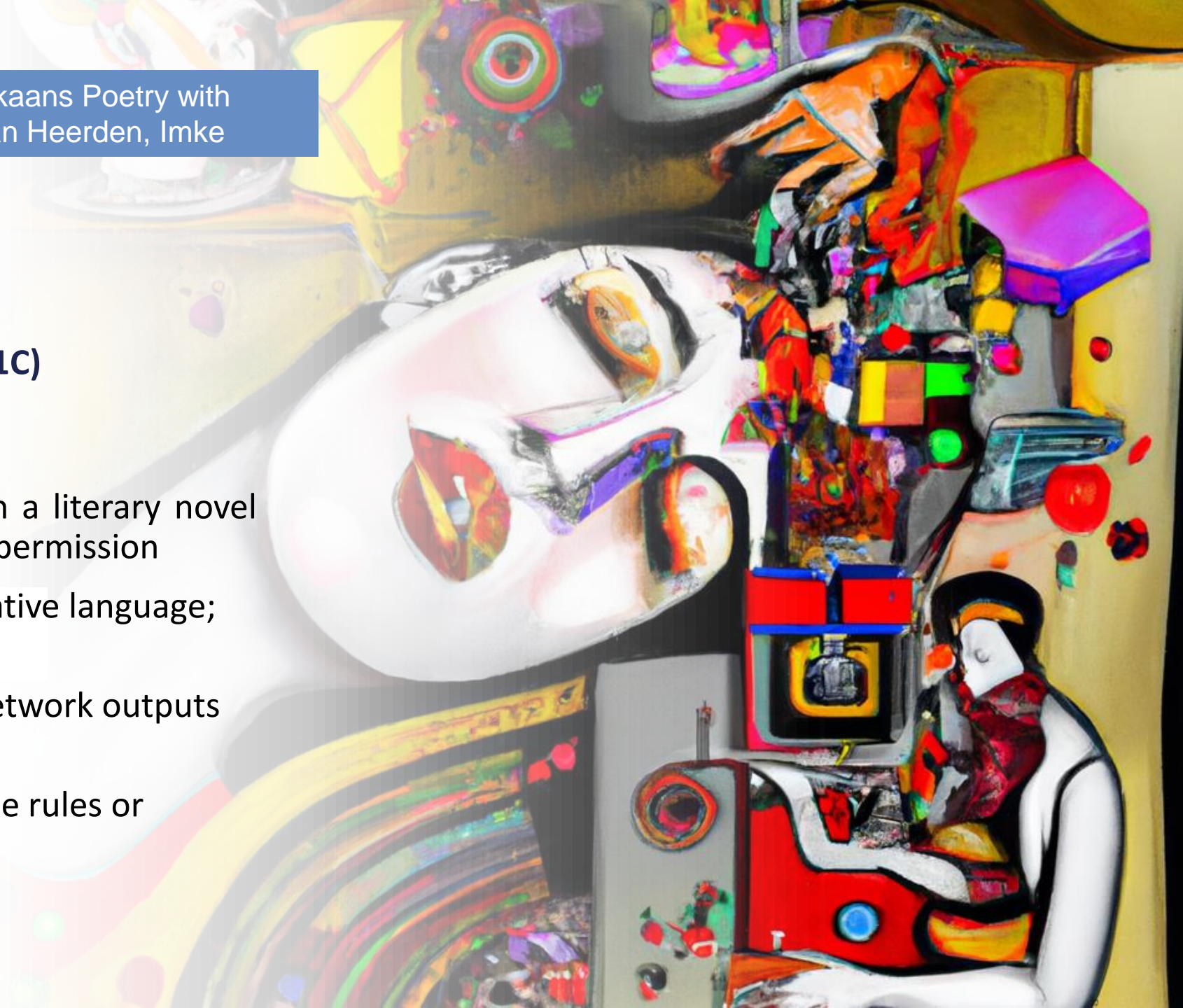
- Poetry dataset availability
- Alternative approach: Trained on a literary novel (Van Heerden, 2019) with author's permission
- Novel's writing style: rich in figurative language; near-poetic language use
- Similar to the trained data, the network outputs unique descriptive phrases
- No specific components to enforce rules or constraints on the model



DATASET

(VAN HEERDEN & BAS, 2021B; 2021C)

- Poetry dataset availability
- Alternative approach: Trained on a literary novel (Van Heerden, 2019) with author's permission
- Novel's writing style: rich in figurative language; near-poetic language use
- Similar to the trained data, the network outputs unique descriptive phrases
- No specific components to enforce rules or constraints on the model
- **Boutique Language Model**





PART 5

RESULTS

Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

RESULTS

Original (Afrikaans)

ons biblioteek by die werkwoord gekaap
die wêreld sê ek met boeke
wêreldletterkunde in armoede
saggies soos 'n spokerigheid
in die vlug van papier
sy vingers draai om haar gevoel
ek het 'n gloeiende noordgrens
woede is jou mond
brand my in die oggendlug
die wind stoppelbaard vorentoe
sy kyk verras op, sy oë verlate
verandering speel as foto's van die wind
die petrolbomme wat nie vertel nie
sukkel is hulle kuns
onbeskermd skittering in die woord
my rug se wit greep
ek is geld want niks kan bloei nie
aarselend weerskante van die staar
demokrasie was 'n daktuin
begin die sonsopkoms voor die dak van my gesig
jou uitgespoel is 'n onderstebo losgewoel
gesprekke vir die oomblik skoongeskraap bleek

Translation (English)

our library hijacked at the verb
the world I say with books
world literature in poverty
softly like a ghostliness
in the flight of paper
his fingers wrap around her feeling
I have a glowing northern border
anger is your mouth
burn me in the morning air
the wind stubbles forth
she looks up in surprise, his eyes deserted
change plays as photos of the wind
the petrol bombs that do not tell
struggling is their art
unprotected brightness in the word
my back's white grip
I am money because nothing can bleed
hesitant on either side of the stare
democracy was a roof garden
begins the sunrise before the roof of my face
your rinsed-out is an upside-down tossed-loose
conversations momentarily clean-scraped pale



(Van Heerden & Bas, 2021b)



PART 6

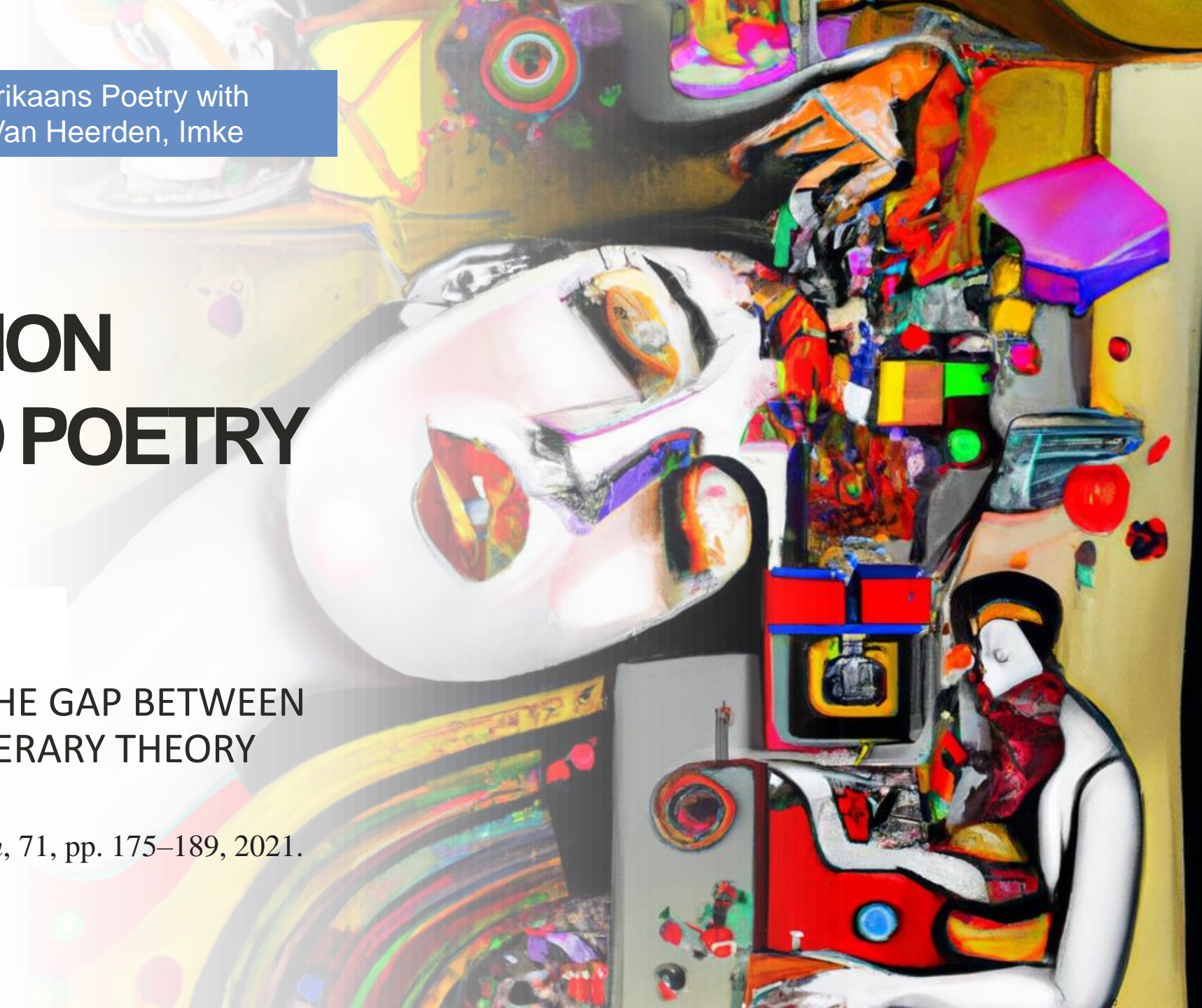
TOWARDS EVALUATION

Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

ON THE EVALUATION OF GENERATED POETRY

AI AS AUTHOR – BRIDGING THE GAP BETWEEN
MACHINE LEARNING AND LITERARY THEORY

Journal of Artificial Intelligence Research, 71, pp. 175–189, 2021.



PROBLEM OF EVALUATION

- **Evaluation:** Required in NLG; important stage in the development process.
- **Aim:** Achieving human likeness.
- Evaluation is conducted using predetermined criteria, e.g. *fluency*, *coherence* and *poeticness*; Likert scale.
- We need to rethink categories of human evaluation in creative text generation.
- Real need for expert evaluators and literary perspectives.

EXAMPLE: RADFORD ET AL. (2019A; 2019B)

- Intend to produce clear, informative text?
Unsuitable topic
- Intend to read figuratively? Intriguing phrase
- Reconceptualising failure in creative language models: What might be considered a weakness in standard text generation might be considered a strength in a creative text generation
- “The kinds of insight [literary texts] provide are qualitatively different from those of pragmatic texts” (Gross, 1997)

fires happening under water
– GPT-2

EXAMPLE: HOPKINS AND KIELA (2017)

- Generated high-quality rhythmic verse
- (Human) Evaluation Results: Generated poems are more humanlike than human-written poems

(a)

The crow crooked on more beautiful and free,
He journeyed off into the quarter sea.
his radiant ribs girdled empty and very -
least beautiful as dignified to see.

Example of automatically generated poetry with
rhythm regularisation and iambic pentameter
(Hopkins and Kiela, 2017)

EXAMPLE: HOPKINS AND KIELA (2017)

- Generated high-quality rhythmic verse
- (Human) Evaluation Results: Generated poems are more humanlike than human-written poems
- Unexplored: Half misjudged the writing of Dickinson, Dryden, Tennyson and Shakespeare as AI-generated
 - Lowest human likeness score: Shakespeare’s “A Fairy Song”
 - Dickinson’s “I’m Nobody” misjudged as well
- Does this reveal an assumption that:
 - Clarity = Human?
 - Complexity = AI?

“A Fairy Song” William Shakespeare

Over hill, over dale,
Thorough bush, thorough brier,
Over park, over pale,
Thorough flood, thorough fire!
I do wander everywhere,
Swifter than the moon’s sphere;
And I serve the Fairy Queen,
To dew her orbs upon the green;
The cowslips tall her pensioners be;
In their gold coats spots you see;
Those be rubies, fairy favours;
In those freckles live their savours;
I must go seek some dewdrops here,
And hang a pearl in every cowslip’s ear.

ON POETRY

- Does this reveal a misunderstanding of the nature, purpose and diversity of poetry?
- Typically dense and polysemous form of literature
- May employ ambiguous and abstract language
- May offer interpretive difficulties (Fabb, 2010):
“Some poetry, including traditional poetry, is for social or aesthetic reasons intended to be difficult”
(Fabb, 2015)

Poetic texts have been described as *ambiguous, confusing, elusive, inaccurate, incorrect, peculiar, unreliable, unclear and uncertain* (Bennett & Royle, 2016)

ON AMBIGUITY

- Literary scholars frequently pay attention to ambiguity in texts (Bennett & Royle, 2016)
- “Any verbal nuance, however slight, which gives room for alternative reactions to the same piece of language” (Empson, 2004)
- “Among the very roots of poetry” (Empson, 2004)

A poem may say one thing
but mean another
(Riffaterre, 1978)

ORDINARY VS LITERARY LANGUAGE

- Ordinary and literary/poetic language may differ (Leung & Durant, 2018; Fabb, 2010) and may have different aims
- Literary language:
 - “Novel uses of language” (Gross, 1997)
 - “Deviations from or distortions of ordinary language” (Bennett & Royle, 2016)
 - “Organised violence committed on ordinary speech” (Jakobson, 1923)
 - “Makes poetry poetry and not a weather report” (Rivkin & Ryan, 2017)

Defamiliarisation
(Shklovsky, 1917): Artistic
technique of presenting to
readers familiar things in an
unfamiliar way – inspiring new
insights and emotions

BUT IT DOESN'T RHYME...

- Poets use various tools and techniques:
 - Asyntactic Structure
 - Anastrophe
 - Anadiplosis
 - Ambiguity
- Consider, e.g., typographic experimentation as well as the absence of rhyme, meter or punctuation

l(a

le

af

fa

ll

s)

one

l

iness

—e.e. cummings

IN A STATION OF THE METRO

The apparition of these faces in the crowd :
Petals on a wet, black bough .

Ezra Pound

Typographic experimentation (Cummings, 1991; Pound, 1913)

RULE-BREAKING

- Creativity as the departure from norms (Gruber, 1988)
- Strict adherence to rules is not necessarily an indication of human likeness
- If poetry is thought to bend/break the rules of ordinary speech, what appears to be “rule-breaking” in AI-generated poetry is not necessarily a failure

Creative endeavours
must deal not with *the
predictable and repeatable* –
the stuff of normal science –
but with the *unique and
unrepeatable*
(Gruber, 1988)

PROBLEMATISING EVALUATION

- We believe difficulty, peculiarity, complexity and ambiguity are not failures
- Prevalent human evaluation criteria: Appropriate for standard text generation
- However, the prioritisation of these evaluation categories in poetry generation might be counterproductive if it loses sight of (or seeks to eliminate) essential poetic qualities

Prevalent human evaluation criteria in NLG (Van der Lee et al., 2021)

Criterion	Total	Criterion	Total
Fluency	40 (27%)	Readability	9 (6%)
Overall quality	29 (20%)	Appropriateness	7 (5%)
Informativeness	15 (10%)	Meaning preservation	6 (4%)
Relevance	15 (10%)	Clarity	5 (3%)
Grammaticality	14 (10%)	Non-redundancy	4 (3%)
Naturalness	12 (8%)	Sentiment	4 (3%)
Coherence	10 (7%)	Consistency	4 (3%)
Accuracy	10 (7%)	Answerability	4 (3%)
Correctness	9 (6%)	Other criteria	124 (48%)*

RESULTS

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

Original (Afrikaans)

ons biblioteek by die werkwoord gekaap
die wêreld sê ek met boeke
wêreldletterkunde in armoede
saggies soos 'n spokerigheid
in die vlug van papier
sy vingers draai om haar gevoel
ek het 'n gloeiende noordgrens
woede is jou mond
brand my in die oggendlug
die wind stoppelbaard vorentoe
sy kyk verras op, sy oë verlate
verandering speel as foto's van die wind
die petrolbomme wat nie vertel nie
sukkel is hulle kuns
onbeskermd skittering in die woord
my rug se wit greep
ek is geld want niks kan bloei nie
aarselend weerskante van die staar
demokrasie was 'n daktuin
begin die sonsopkoms voor die dak van my gesig
jou uitgespoel is 'n onderstebo losgewoel
gesprekke vir die oomblik skoongeskraap bleek

Translation (English)

our library hijacked at the verb
the world I say with books
world literature in poverty
softly like a ghostliness
in the flight of paper
his fingers wrap around her feeling
I have a glowing northern border
anger is your mouth
burn me in the morning air
the wind stubbles forth
she looks up in surprise, his eyes deserted
change plays as photos of the wind
the petrol bombs that do not tell
struggling is their art
unprotected brightness in the word
my back's white grip
I am money because nothing can bleed
hesitant on either side of the stare
democracy was a roof garden
begins the sunrise before the roof of my face
your rinsed-out is an upside-down tossed-loose
conversations momentarily clean-scraped pale

FLUENCY



INFORMATIVENESS

COHERENCE

CLARITY

Some prevalent categories of evaluation (Van der Lee et al., 2021)

RESULTS

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

Original (Afrikaans)	Translation (English)
ons biblioteek by die werkwoord gekaap die wêreld sê ek met boeke wêreldletterkunde in armoede saggies soos 'n spokerigheid in die vlug van papier sy vingers draai om haar gevoel ek het 'n gloeiende noordgrens woede is jou mond brand my in die oggendlug die wind stoppelbaard vorentoe sy kyk verras op, sy oë verlate verandering speel as foto's van die wind die petrolbomme wat nie vertel nie sukkel is hulle kuns onbeskermd skittering in die woord my rug se wit greep	our library hijacked at the verb the world I say with books world literature in poverty softly like a ghostliness in the flight of paper his fingers wrap around her feeling I have a glowing northern border anger is your mouth burn me in the morning air the wind stubbles forth she looks up in surprise, his eyes deserted change plays as photos of the wind the petrol bombs that do not tell struggling is their art unprotected brightness in the word my back's white grip
ek is geld want niks kan bloei nie aarselend weerskante van die staar demokrasie was 'n daktuin begin die sonsopkoms voor die dak van my gesig jou uitgespoel is 'n onderstebo losgewoel gesprekke vir die oomblik skoongeskraap bleek	I am money because nothing can bleed hesitant on either side of the stare democracy was a roof garden begins the sunrise before the roof of my face your rinsed-out is an upside-down tossed-loose conversations momentarily clean-scraped pale

FLUENCY



INFORMATIVENESS



COHERENCE



CLARITY



Some prevalent categories of evaluation (Van der Lee et al., 2021)

RESULTS

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

Original (Afrikaans)

ons biblioteek by die werkwoord gekaap
die wêreld sê ek met boeke
wêreldletterkunde in armoede
saggies soos 'n spokerigheid
in die vlug van papier
sy vingers draai om haar gevoel
ek het 'n gloeiende noordgrens
woede is jou mond
brand my in die oggendlug
die wind stoppelbaard vorentoe
sy kyk verras op, sy oë verlate
verandering speel as foto's van die wind
die petrolbomme wat nie vertel nie
sukkel is hulle kuns
onbeskermd skittering in die woord
my rug se wit greep
ek is geld want niks kan bloei nie
aarselend weerskante van die staar
demokrasie was 'n daktuin
begin die sonsopkoms voor die dak van my gesig
jou uitgespoel is 'n onderstebo losgewoel
gesprekke vir die oomblik skoongeskraap bleek

Translation (English)

our library hijacked at the verb
the world I say with books
world literature in poverty
softly like a ghostliness
in the flight of paper
his fingers wrap around her feeling
I have a glowing northern border
anger is your mouth
burn me in the morning air
the wind stubbles forth
she looks up in surprise, his eyes deserted
change plays as photos of the wind
the petrol bombs that do not tell
struggling is their art
unprotected brightness in the word
my back's white grip
I am money because nothing can bleed
hesitant on either side of the stare
democracy was a roof garden
begins the sunrise before the roof of my face
your rinsed-out is an upside-down tossed-loose
conversations momentarily clean-scraped pale

FLUENCY

INFORMATIVENESS

COHERENCE

CLARITY

Some prevalent categories of evaluation (Van der Lee et al., 2021)

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

1. Die lewe is 'n draaiende wiel wat almal draai en val. (Life is a spinning wheel that everyone turns and falls.)
2. Die liefde is 'n warm deken wat ons in die koue nagte omhul. (Love is a warm blanket that envelops us on cold nights.)
3. Die tyd is 'n onverbiddelike vyand wat altyd aan ons knaag. (Time is an unforgiving enemy that always gnaws at us.)
4. Die vriendskap is 'n lig wat ons pad in die donker verlig. (Friendship is a light that illuminates our path in the dark.)
5. Die toekoms is 'n onbekende land wat ons met hoop en vrees betree. (The future is an unknown country that we enter with hope and fear.)
6. Die geheue is 'n skatkis vol ou herinneringe wat ons soms laat glimlag en soms laat huil. (Memory is a treasure trove of old memories that sometimes make us smile and sometimes make us cry.)
7. Die woede is 'n brandende vuur wat alles in sy pad vernietig. (Anger is a burning fire that destroys everything in its path.)
8. Die musiek is 'n stroom van emosies wat ons hart laat sing. (Music is a stream of emotions that makes our heart sing.)
9. Die gedagtes is 'n wilde perd wat ons soms laat galoppeer en soms moet tem. (Thoughts are a wild horse that sometimes makes us gallop and sometimes we have to tame.)
10. Die vryheid is 'n voël wat vlerke het om te vlieg en die blou lug te verken. (Freedom is a bird with wings to fly and explore the blue sky.)

FLUENCY



INFORMATIVENESS



COHERENCE



CLARITY



Metaphors generated by ChatGPT. Prompt provided: *Create a list of ten unrelated original metaphors in Afrikaans.*

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke



Here are five unrelated original phrases in Afrikaans with figurative meaning and imagery:

1. 'n Rots in die branding wees (To be a rock in the surf) - To be a reliable and steadfast support in times of difficulty or uncertainty.
2. Die pad vorentoe is vol struikelblokke (The road ahead is full of stumbling blocks) - To describe a situation that is challenging and difficult, with many obstacles and setbacks.
3. Sy het 'n hart van goud (She has a heart of gold) - To describe someone who is kind, generous, and compassionate.
4. Die wêreld is 'n oester en ek sal dit kraak (The world is an oyster and I will crack it) - To express a determination to succeed and make the most of opportunities.
5. Hy loop op eiers (He walks on eggs) - To describe someone who is cautious and careful in their actions and words, as if walking on fragile eggs.

Text generated by ChatGPT. Prompt provided: *Create a list of five unrelated original phrases with figurative meaning and imagery in Afrikaans.*

FLUENCY



INFORMATIVENESS



COHERENCE



CLARITY



Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke



Hier is 'n innoverende poëtiese frase met figuurlike betekenis in Afrikaans:

Die maan se glimlag is 'n kitsvlam van hoop wat die donker verjaag.

Translation: The moon's smile is a quick flame of hope that dispels the darkness.

Text generated by ChatGPT. Prompt provided:

*Write an **incredibly innovative** poetic phrase with figurative meaning in Afrikaans.*

FLUENCY



INFORMATIVENESS



COHERENCE



CLARITY



RESULTS

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

Original (Afrikaans)

ons biblioteek by die werkwoord gekaap
die wêreld sê ek met boeke
wêreldletterkunde in armoede
saggies soos 'n spokerigheid
in die vlug van papier
sy vingers draai om haar gevoel
ek het 'n gloeiende noordgrens
woede is jou mond
brand my in die oggendlug
die wind stoppelbaard vorentoe
sy kyk verras op, sy oë verlate
verandering speel as foto's van die wind
die petrolbomme wat nie vertel nie
sukkel is hulle kuns
onbeskermd skittering in die woord
my rug se wit greep
ek is geld want niks kan bloei nie
aarselend weerskante van die staar
demokrasie was 'n daktuin
begin die sonsopkoms voor die dak van my gesig
jou uitgespoel is 'n onderstebo losgewoel
gesprekke vir die oomblik skoongeskraap bleek

Translation (English)

our library hijacked at the verb
the world I say with books
world literature in poverty
softly like a ghostliness
in the flight of paper
his fingers wrap around her feeling
I have a glowing northern border
anger is your mouth
burn me in the morning air
the wind stubbles forth
she looks up in surprise, his eyes deserted
change plays as photos of the wind
the petrol bombs that do not tell
struggling is their art
unprotected brightness in the word
my back's white grip
I am money because nothing can bleed
hesitant on either side of the stare
democracy was a roof garden
begins the sunrise before the roof of my face
your rinsed-out is an upside-down tossed-loose
conversations momentarily clean-scraped pale

ORIGINALITY



AMBIGUITY



PECULIARITY



DEFAMILIARIZATION



Alternative categories for creative text evaluation (Van Heerden & Bas, 2021a)

RESULTS

Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

Original (Afrikaans)	Translation (English)
ons biblioteek by die werkwoord gekaap die wêreld sê ek met boeke wêreldletterkunde in armoede saggies soos 'n spokerigheid in die vlug van papier sy vingers draai om haar gevoel ek het 'n gloeiende noordgrens woede is jou mond brand my in die oggendlug die wind stoppelbaard vorentoe sy kyk verras op, sy oë verlate verandering speel as foto's van die wind die petrolbomme wat nie vertel nie sukkel is hulle kuns onbeskermd skittering in die woord my rug se wit greep ek is geld want niks kan bloei nie aarselend weerskante van die staar demokrasie was 'n daktuin	our library hijacked at the verb the world I say with books world literature in poverty softly like a ghostliness in the flight of paper his fingers wrap around her feeling I have a glowing northern border anger is your mouth burn me in the morning air the wind stubbles forth she looks up in surprise, his eyes deserted change plays as photos of the wind the petrol bombs that do not tell struggling is their art unprotected brightness in the word my back's white grip I am money because nothing can bleed hesitant on either side of the stare democracy was a roof garden
begin die sonsopkoms voor die dak van my gesig jou uitgespoel is 'n onderstebo losgewoel gesprekke vir die oomblik skoongeskraap bleek	begins the sunrise before the roof of my face your rinsed-out is an upside-down tossed-loose conversations momentarily clean-scraped pale

ORIGINALITY



AMBIGUITY



PECULIARITY



DEFAMILIARIZATION



Alternative categories for creative text evaluation (Van Heerden & Bas, 2021a)

ARGUMENT

- *Literariness* is understudied in relation to text generation
- Literary-theoretical perspectives are beneficial
- Collaboration:
 - Rethinking purpose and priorities of poetry generation
 - Reconceptualising failure in creative language models
 - Development and explanation of suitable evaluation criteria
 - (Truly) expert evaluation

Originality

- There is no operational scoring system that prioritises originality in generated text (Klebanov and Madnani, 2020).
- Traits and measurement criteria of originality have yet to be determined in computational linguistics (Klebanov & Madnani, 2020).
- Some have considered this unachievable outside this context (Gruber, 1988).
- Attempts at categorisation may fail to do justice to the uniqueness and power of poetry and, therefore, require great dexterity (Gross, 1997).

(Van Heerden & Bas, 2021a)




Novel or Nonsensical?

- **Results:** Unique generated phrases containing metaphor and personification.
- **Metaphorical language** is open to interpretation.
- Invites the reader to participate in the process of meaning-making (White, 1996).
- Metaphor creates new, often surprising, likeness between concepts (Indurkha, 1992; Veale et al., 2016).
- Contributes to the impression of intent and sentience.

(Van Heerden & Bas, 2021b)

Original (Afrikaans)	Translation (English)
ons biblioteek by die werkwoord gekaap die wêreld sê ek met boeke wêreldletterkunde in armoede saggies soos 'n spokerigheid in die vlug van papier sy vingers draai om haar gevoel ek het 'n gloeiende noordgrens woede is jou mond brand my in die oggendlug die wind stoppelbaard vorentoe sy kyk verras op, sy oë verlate verandering speel as foto's van die wind die petrolbomme wat nie vertel nie sukkel is hulle kuns onbeskermde skittering in die woord my rug se wit greep ek is geld want niks kan bloei nie aarselend weerskante van die staar demokrasie was 'n daktuin begin die sonsopkoms voor die dak van my gesig jou uitgespoel is 'n onderstebo losgewoel gesprekke vir die oomblik skoongeskraap bleek	our library hijacked at the verb the world I say with books world literature in poverty softly like a ghostliness in the flight of paper his fingers wrap around her feeling I have a glowing northern border anger is your mouth burn me in the morning air the wind stubbles forth she looks up in surprise, his eyes deserted change plays as photos of the wind the petrol bombs that do not tell struggling is their art unprotected brightness in the word my back's white grip I am money because nothing can bleed hesitant on either side of the stare democracy was a roof garden begins the sunrise before the roof of my face your rinsed-out is an upside-down tossed-loose conversations momentarily clean-scraped pale

The background of the slide is a complex, abstract painting. It features a central, somewhat realistic face with large, expressive eyes. Surrounding this face are numerous colorful, geometric and organic shapes, including circles, squares, and irregular forms in shades of red, yellow, blue, green, and purple. The overall style is reminiscent of mid-20th-century abstract art, possibly influenced by artists like Piet Mondrian or Wassily Kandinsky. The colors are bright and saturated, creating a sense of dynamic energy and visual complexity.

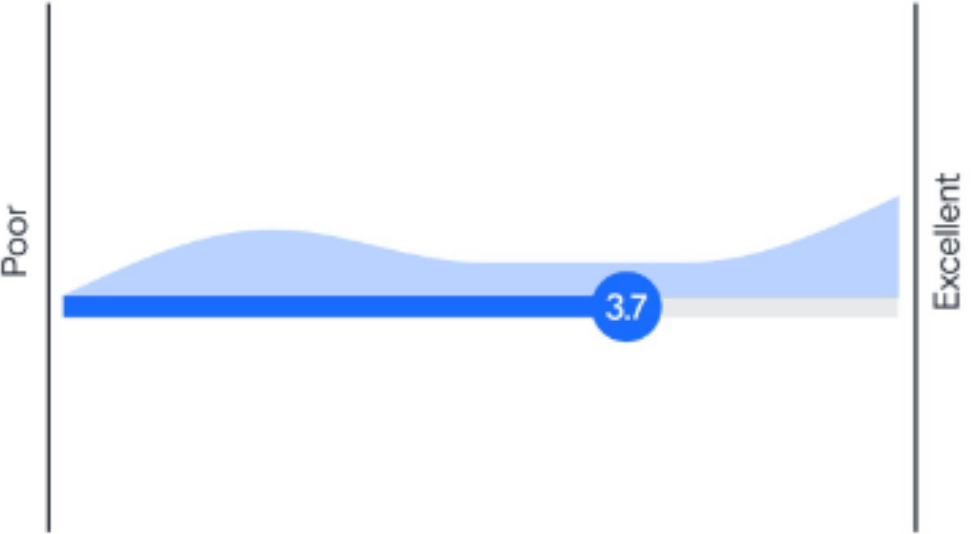
We now have machines that can mindlessly generate words, but we haven't learned how to stop imagining a mind behind them.
– Emily Bender (qtd. in Tiku, 2022)

Although such models are sometimes impressive – generating poetry or correctly answering trivia questions – they have no sense of the meaning of language, which causes them to also create gibberish.
– Elizabeth Gibney (2022)

EVALUATION

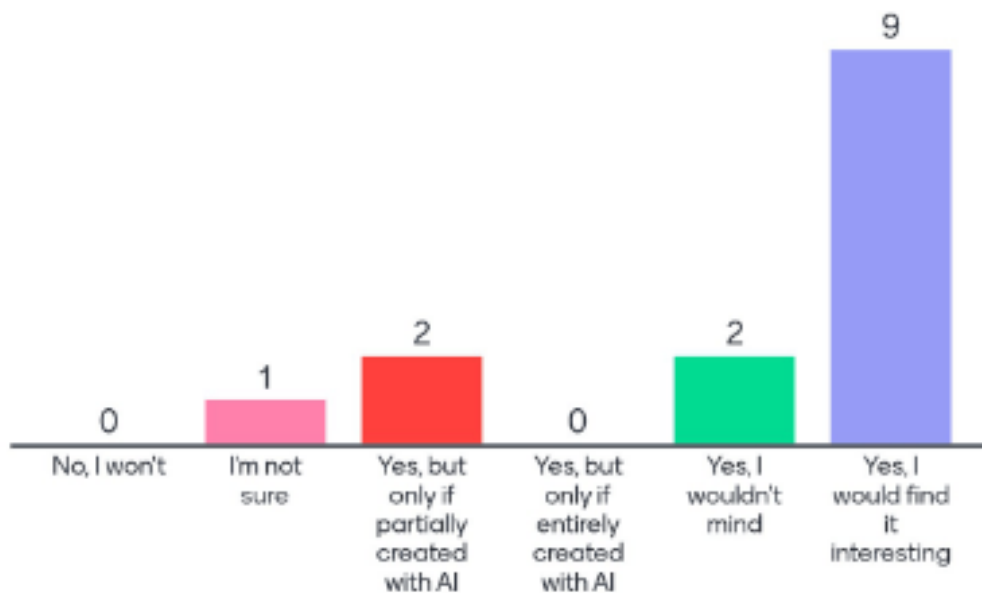


How would you rate the quality of AfriKI's generated phrases?



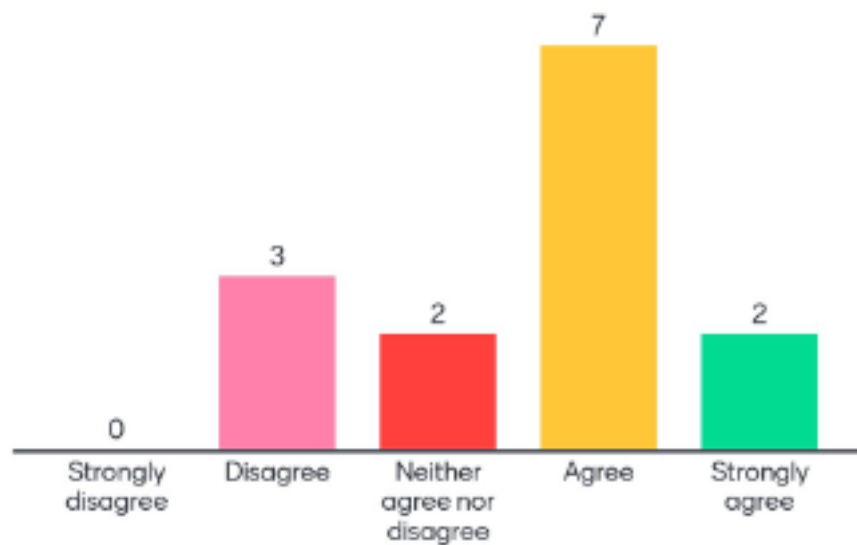
EVALUATION

Would you read a literary text created with AI?



EVALUATION

Do you think AI-assisted writing will become the norm?





PART 7

MACHINE-IN-THE-LOOP

AFRIKI

CO-CREATIVE PROCESS (VAN HEERDEN & BAS, 2021C)

- **Machine-in-the-loop** frameworks promote human creativity through computational assistance, as opposed to human-in-the-loop pipelines, which aim to strengthen machine learning models (Clark et al., 2018).
- **Human-centred approach**
- Creativity support tools “make more people more creative more often” (Schneiderman, 2002)
- AI is used as an **instrument** to augment creativity
- **Co-creativity** = human + machine
- Embracing, instead of eradicating, the **strangeness** of generated text

One criticism of artificial intelligence programs that claim to be creative is exactly that a human plays a role at some point, which reduces the autonomy of the machine. From the HCI perspective [...] these “failed” AI creativity programs are examples of successful human–computer interactions to facilitate creativity. – Lubart (2005)

AFRIKI

CO-CREATIVE PROCESS (VAN HEERDEN & BAS, 2021C)

■ Computational:

- Network generates thousands of individual lines (instead of stanzas of verse)
- Results are highly original (distinct from dataset; no repetition of word order)

■ Human:

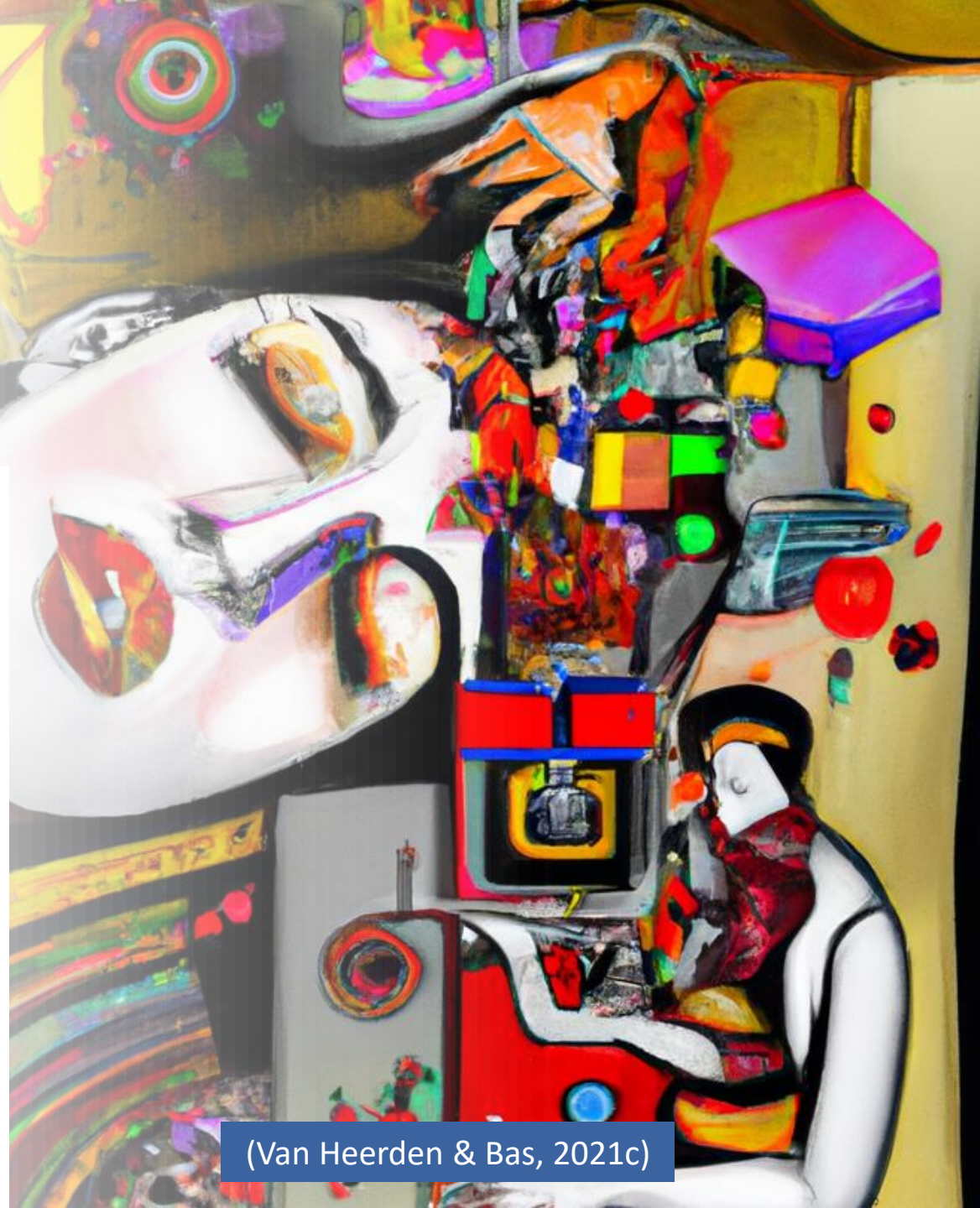
- Poet vertically arranges phrases into poems
- Selection and arrangement
- No modification of the word order
- Method highlights the model's contribution
- However, the human's role should not be underestimated

Original (Afrikaans)	Translation (English)
<i>Die konstabel se skiereiland</i>	<i>The constable's peninsula</i>
Afrika drink onheil in die water. Die landskap kantel sy rug in sigbewaking en vlam. Ons oopgesnyde sake brandtrappe vir die ander state. Hierdie grond word intimidasie.	Africa drinks disaster in the water. The landscape tilts its back in surveillance and flame. Our cut-open affairs fire escapes for other states. This soil becomes intimidation.
<i>Gedigte, daar by die brul van 'n brander</i>	<i>Poetry, there near the roar of a wave</i>
Hier is die oë katvoet vir die speelrotse onder uitdrukkings die golwe van gister wat getol en woens en water saam met die son skuim in hul woorde die ingedagte see lig die geure en praat 'n asemhaal	Here the eyes are cautious of the sea rocks under expressions the waves of yesterday whirled and wild and water that froth with the sun in their words the introspective sea lifts the scents and utters a breath

AFRIKI

EXAMPLE POEMS

Original (Afrikaans)	Translation (English)
<p><i>Die konstabel se skiereiland</i></p> <p>Afrika drink onheil in die water. Die landskap kantel sy rug in sigbewaking en vlam. Ons oopgesnyde sake brandtrappe vir die ander state. Hierdie grond word intimidasie.</p>	<p><i>The constable's peninsula</i></p> <p>Africa drinks disaster in the water. The landscape tilts its back in surveillance and flame. Our cut-open affairs fire escapes for other states. This soil becomes intimidation.</p>
<p><i>Gedigte, daar by die brul van 'n brander</i></p> <p>Hier is die oë katvoet vir die spoelrotse onder uitdrukkings die golwe van gister wat getol en woes en water saam met die son skuim in hul woorde</p> <p>die ingedagte see lig die geure en praat 'n asemhaal</p>	<p><i>Poetry, there near the roar of a wave</i></p> <p>Here the eyes are cautious of the sea rocks under expressions the waves of yesterday whirled and wild and water that froth with the sun in their words</p> <p>the introspective sea lifts the scents and utters a breath</p>



(Van Heerden & Bas, 2021c)

AFRIKI

EXAMPLE POEMS

Kaapstad

Vandag is ons nie net die stad nie
maar
die vertaler van die son

Vanaand se gordyne
glinster by skuifvensters
in die stadsliggies

Die uur van die winde
sorg dat dit rondom klink
Sy wil die glasvensters deurkosyn
eens iets te beskerm

Tafelberg
maak 'n vraag waarbinne ons
'n duisend name
genoem word

Cape Town

Today we are not just the city
but
the translator of the sun

Tonight's curtains
glitter at sliding windows
in the city lights

The hour of the winds
takes care it sounds around
She wants to doorframe the glass windows
to protect something

Table Mountain
creates a question in which we
are given
a thousand names

Sonlig son die promenade
-
Sunlight suns the promenade

AFRIKI

EXAMPLE POEMS

Kaapstad

Vandag is ons nie net die stad nie
maar
die vertaler van die son

Vanaand se gordyne
glinster by skuifvensters
in die stadsliggies

Die uur van die winde
sorg dat dit rondom klink

Sy wil die glasvensters deurkosyn
eens iets te beskerm

Tafelberg
maak 'n vraag waarbinne ons
'n duisend name
genoem word

Cape Town

Today we are not just the city
but
the translator of the sun

Tonight's curtains
glitter at sliding windows
in the city lights

The hour of the winds
takes care it sounds around

She wants to doorframe the glass windows
to protect something

Table Mountain
creates a question in which we
are given
a thousand names

Sonlig son die promenade
-
Sunlight suns the promenade

AFRIKI

SUMMARY

- **Dataset:** Etienne van Heerden's novel *Die biblioteek aan die einde van die wêreld* (2019). Approximately 210K words in length.
- **Language model:** AfriKI – *Afrikaanse Kunsmatige Intelligensie* (Afrikaans Artificial Intelligence). LSTM architecture.
- **Results:** New phrases containing compelling figures of speech such as metaphor, simile and personification.
- **Approach:** Co-creative, i.e. human + machine. Human poet arranges the generated phrases to create poetry.



tinyurl.com/AfrikaansNLG
(Van Heerden and Bas, 2021c)



tinyurl.com/figlanggen
(Van Heerden and Bas, 2021b)



PART 8

CREATING A POETRY COLLECTION

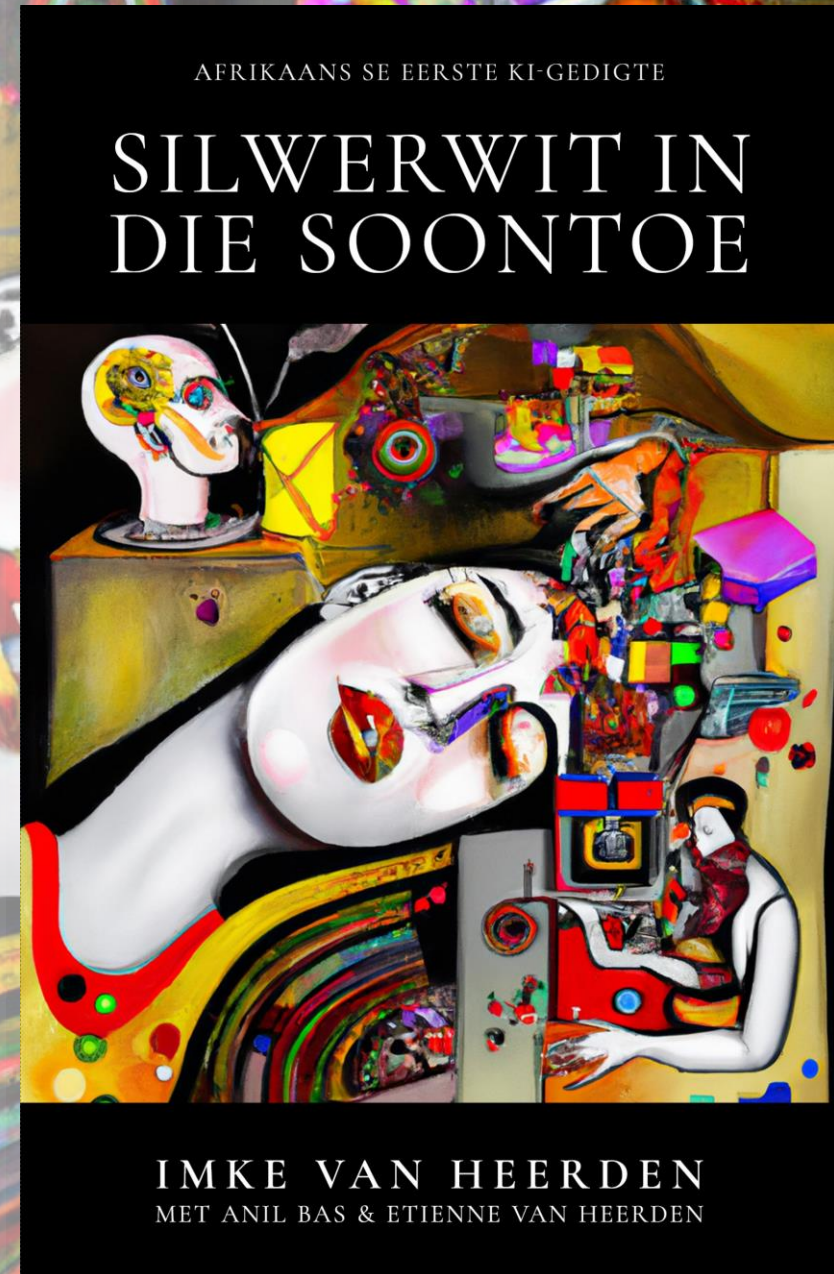
Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

ABOUT THE BOOK:

Silwerwit in die soontoe (transl. *Silverwhite in the Distance*) is a groundbreaking attempt to explore the creative potentialities of language with the help of artificial intelligence (AI) – a first in Afrikaans, as well as in South Africa. The poems in this volume were composed co-creatively – a human in collaboration with a machine. The poet, Imke van Heerden, delicately interweaves phrases of AI-generated text to create verse, as part of an experiment that examines the following timely question: How might this technology augment and challenge the art of poetry? In poems on AI, Africa, Cape Town and the Karoo, waves glisten, air burns and a machine dances on the outskirts of language.

COVER IMAGE:

AI artwork generated with DALL·E 2 by OpenAI.

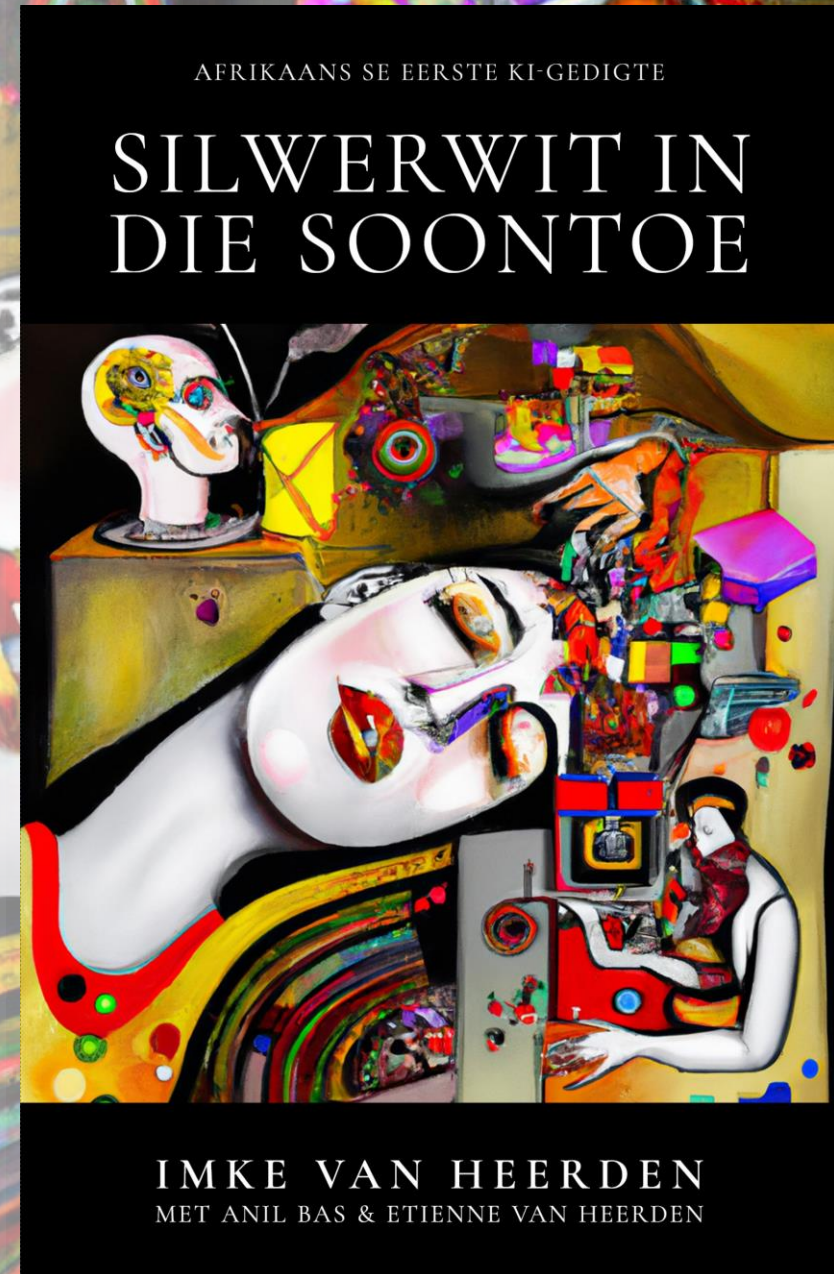


Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

CONCLUDING THOUGHTS

IDEAS

- ▀ Literary versus ordinary language
- ▀ Co-creative versus fully automatic text generation
- ▀ AfriKI: Instrument/device that “makes language strange”
- ▀ Enrichment and promotion of low-resource languages

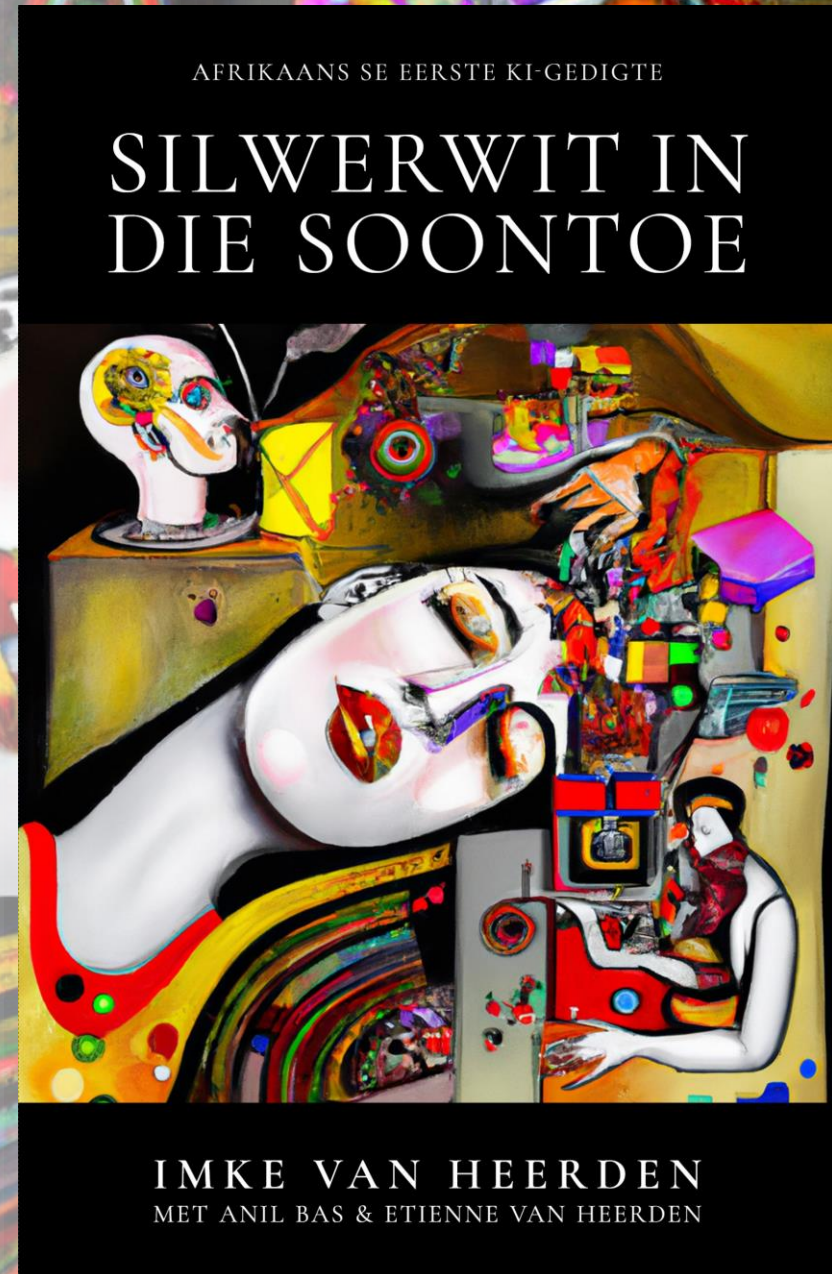


Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

CONCLUDING THOUGHTS

QUESTIONS (VAN HEERDEN & BAS, 2021A)

- (How) will AI impact our discipline?
- How might this technology encourage and challenge creative writing practices?
- Will AI writing always be read comparatively?
- Evaluation? New theoretical perspectives?
- Impact on disciplinary boundaries?
- Will it inspire reconceptualisations of *originality* and *creativity*? And *authorship*?
- **Literary value?**



Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

INTERDISCIPLINARITY & COLLABORATION



Making Strange: Co-Creating Afrikaans Poetry with
a Boutique Language Model – Van Heerden, Imke

ACKNOWLEDGEMENT

This work has been produced benefiting from the 2232 International Fellowship for Outstanding Researchers Program of TÜBİTAK (Project No: 118C285). However, the entire responsibility of the work belongs to the owner of the work. The financial support received from TÜBİTAK does not mean that the content of the presentation is approved in a scientific sense by TÜBİTAK.





Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

REFERENCES

- ACL. (2022). Language Diversity: From Low-Resource to Endangered Languages. *Association for Computational Linguistics*.
- Augustinus, L., et al. (2016). AfriBooms: An online treebank for Afrikaans. In *Proc. LREC*, pp. 677–682.
- Baldick, C. (1996). *The Concise Oxford Dictionary of Literary Terms*. Oxford University Press.
- Bennett, A., & Royle, N. (2016). *An Introduction to Literature, Criticism and Theory*. Routledge.
- Clark, E., et al. (2018). Creative Writing with a Machine in the Loop: Case Studies on Slogans and Stories. In *Proc. ACM IUI*, pp. 329–340.
- Cohen, T. (1978). Metaphor and the Cultivation of Intimacy. *Critical Inquiry*, 5(1), pp. 3–12.
- Chandler D., & Munday, R. (2011). *A Dictionary of Media and Communication*. Oxford University Press.
- Cummings, E. E. (1991). *Complete Poems, 1904–1962*. Liveright Publishing.
- Dirix, P., et al. (2017). Universal Dependencies for Afrikaans. In *Proc. NoDaLiDa*, pp. 38–47.
- Eiselen, R., & Puttkammer, M. (2014). Developing Text Resources for Ten South African Languages. In *Proc. LREC*, pp. 3698–3703.
- Empson, W. (2004). *Seven Types of Ambiguity*. Random House.
- Fabb, N. (2010). Is Literary Language a Development of Ordinary Language?. *Lingua*, 120(5), pp. 1219–1232.
- Fabb, N. (2015). *What is Poetry?: Language and Memory in the Poems of the World*. Cambridge University Press.
- Fainsilber, L. & Ortony, A. (1987). Metaphorical Uses of Language in the Expression of Emotions. *Metaphor and Symbol*, 2(4), pp. 239–250.
- Fussell, S. R., & Moss, M. M. (1998). Figurative Language in Emotional Communication. *Social and Cognitive Approaches to Interpersonal Communication*, pp. 113–141.
- Gibbs, R. W., Leggitt, J. S., & Turner, E. A. (2002). What's Special About Figurative Language in Emotional Communication. *The Verbal Communication of Emotions: Interdisciplinary Perspectives*, pp. 125–149.
- Grace, K., et al. (2018). When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, 62, pp. 729–754.
- Gross, S. (1997). Cognitive Readings; or, The Disappearance of Literature in the Mind. *Poetics Today*, 18(2), pp. 271–297.
- Gruber, H. E. (1988). The Evolving Systems Approach to Creative Work. *Creativity Research Journal*, pp. 27–51.
- Hall, S. B. (2018). AI will write a best-seller by 2049, experts predict. *World Economic Forum*.
- Hopkins, J., & Kiela, D. (2017). Automatically Generating Rhythmic Verse with Neural Networks. In *Proc. ACL*, pp. 168–178.
- Indurkha, B. (1992). *Metaphor and Cognition: An Interactionist Approach*. Kluwer Academic.
- Jakobson, R. (1923). *O cheshskom stikhe preimushchestvenno v sopostavlenii s russkim* [On Czech Verse, Primarily in Comparison with Russian]. Opoiaz – MLK.
- Joshi, P., et al. (2020). The State and Fate of Linguistic Diversity and Inclusion in the NLP World. In *Proc. ACL*, pp. 6282–6293.
- Klebanov, B., & Madnani, N. (2020). Automated Evaluation of Writing – 50 Years and Counting. In *Proc. ACL*, pp. 7796–7810.
- Kovecses, Z. (2010). *Metaphor: A Practical Introduction*. Oxford University Press.
- Leung, J. H., & Durant, A. (2018). *Meaning and Power in the Language of Law*. Cambridge University Press.
- Lubart, T. (2005). How Can Computers Be Partners in the Creative Process: Classification and Commentary on the Special Issue. *International Journal of Human-Computer Studies*, 63(4-5), pp. 365–369.
- Pound, E. (1913). In a station of the metro. *Poetry*, 2(1), 12.
- Radford, A., et al. (2019a). Language Models are Unsupervised Multitask Learners. OpenAI Blog.
- Radford, A., et al. (2019b). Better language models and their implications. OpenAI Blog.



Making Strange: Co-Creating Afrikaans Poetry with a Boutique Language Model – Van Heerden, Imke

REFERENCES

- Riffaterre, M. (1978). *Semiotics of Poetry*. Indiana University Press.
- Rivkin, J., & Ryan, M. (2017). *Literary Theory: An Anthology*. John Wiley & Sons.
- Roux, J. (2016). South African National Centre for Digital Language Resources. In *Proc. LREC*, pp. 2467–2470.
- Sanby, L., Todd, I., & Keet, M. C. (2016). Comparing the Template-Based Approach to GF: The Case of Afrikaans. In *Proc. WebNLG*, pp. 50–53.
- Shneiderman, B. (2002). Creativity support tools. *Communications of the ACM*, 45(10), pp. 116–120.
- Shklovsky, V. (1917). Art as technique. In *Literary Theory: An Anthology*, pp. 15–21. John Wiley & Sons.
- Steen, G. (1994). *Understanding Metaphor in Literature: An Empirical Approach*. Longman.
- Van der Lee, et al. (2021). Human Evaluation of Automatically Generated Text: Current Trends and Best Practice Guidelines. *Computer Speech & Language*, 67, 101151.
- Van Heerden, E. (2019). *Die Biblioteek aan die Einde van die Wêreld*. NB-Uitgewers.
- Van Heerden, I., & Bas, A. (2021a). AI as Author – Bridging the Gap Between Machine Learning and Literary Theory. *Journal of Artificial Intelligence Research*, 71, pp. 175–189.
- Van Heerden, I., & Bas, A. (2021b). Towards Figurative Language Generation in Afrikaans. In *Proc. NAACL Workshop on Research in Computational Typology and Multilingual NLP*, p. 10.
- Van Heerden, I., & Bas, A. (2021c). AfriKI: Machine-in-the-Loop Afrikaans Poetry Generation. In *Proc. EACL Workshop on Bridging HCI and NLP*, pp. 74–80.
- Van Heerden, I., Bas, A., & Van Heerden, E. (2023). *Silwerwit in die soontoë: Afrikaans se eerste KI-gedigte*. Naledi Publishers.
- Van Zaanen, M., & Van Huyssteen, G. (2003). Improving a Spelling Checker for Afrikaans. In *Proc. CLIN*, pp. 143–156.
- Veale, T., et al. (2016). Metaphor: A Computational Perspective. In *Synthesis Lectures on Human Language Technologies*. Morgan & Claypool Publishers.
- Weller, C., & Gould, S. (2017). Here's when robots will start beating humans at every task. *Business Insider* <www.businessinsider.com/when-robots-will-take-peoples-jobs-2017-6>.
- White, R. M. (1996). *The Structure of Metaphor: The Way the Language of Metaphor Works*. Wiley.
- Ziering, P., & Van der Plas, L. (2016). Towards Unsupervised and Language-Independent Compound Splitting Using Inflectional Morphological Transformations. In *Proc. NAACL*, pp. 644–653.
- ∇, et al. (2020a). Masakhane – Machine Translation for Africa. In *Proc. ICLR Workshop*.
- ∇, et al. (2020b). Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages. In *Proc. EMNLP*, pp. 2144–2160.
- ∇, et al. (2021). MasakhaNER: Named Entity Recognition for African Languages. In *EACL Workshop on AfricaNLP*.