One day, it is said, the Tortoises held a council how they might hunt Ostriches, and they said, "Let us, on both sides, stand in rows near each other, and let one go to hunt the Ostriches, so that they must flee along through the midst of us..."

A wolf resolved to disguise himself in order that he might prey on a flock of sheep without fear of being detected. So he clothed himself in a sheepskin and slipped among the sheep when they were out at pasture...

West African vs. Western European
Using Natural Language Processing Techniques to Identify West African Features in Folk Narratives

1st thesis presentation
By Gossa Lô

Introduction

➢ ICT4D
➢ Voice-based tools
➢ Communicative differences
➢ 2CoolMonkeys internship

Source: https://2coolmonkeys.nl/
Storytelling in West Africa

➢ Communicative interaction
➢ Storytellers (Griots)
➢ Cohesion, shared identity
➢ Oral → written

Source: http://bcrcmontreal.com/
Fables

- Fictional story
- Anthropomorphized
- Moral message
- Global similarities
- Main characters - setting

Source: http://toubibadakar.mondoblog.org/
Natural Language Processing (NLP)

➢ Subfield of Artificial Intelligence
  ○ Speech recognition
  ○ Machine translation
  ○ Natural Language Understanding/Generation

➢ Machine/Deep learning

Source: https://vaughanrockets.typepad.com/
Research questions

How can Natural Language Processing Techniques be used to analyze and generate West African folktales?

➢ Can we build an AI story generator that generates stories with West African features?
➢ Can we build a classifier that distinguishes between West African and Western European stories?
➢ How can we incorporate these techniques in ICT4D tools?

Source: https://mcfarlinumc.org/
Methodology
Bag of Words classification

Input sentence: the dog is on the table

Vectorized sentence:

<table>
<thead>
<tr>
<th>are</th>
<th>cat</th>
<th>dog</th>
<th>is</th>
<th>now</th>
<th>on</th>
<th>table</th>
<th>the</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Vocabulary

Accuracy classifier: 83%

Source: https://medium.com/
Recurrent Neural Network (RNN)

Source: http://karpathy.github.io/
Text generation

- Character vs. word level
- More data = better
- Unicorn coherence

**Input seed:** hiding all the time and found the leopard lying dead with okun standing over it then in very strong language effiong began to upbraid his friend and asked him why he had killed his old friend the leopard that nothing would satisfy him but that he should report the whole matter...

**Generated continuation:** ...to the king who had been a good life but angry where a poor fellow would be able to catch the cunning man with all the ministers and punish him that he could not move because she was very hungry he could not be allowed to answer his people to remain in the dry profession and instructed his way back to reveal the year

Source: https://carrollwoodpta.org
Result analysis: human evaluation

➢ Rank texts according to coherence.
  ○ Original texts ranked best

➢ Classify texts
  ○ 7/10 correct, 3/10 unclear
  ○ Puma vs. wolf

\[ t=0.5 \]
"A very long time ago, before the drum was close to the water and he started to climb to the spring the tortoise had deceived his teacher to the boy. The Iyawo never had enough to stop the pot of stew, so that he could not get it on his head."

\[ t=1.0 \]
"There had been a long walk ahead of the big river without seeding his great joy, sats. His mother had died it was time for he had made slaves. "My brother, may I can beat me you will have to take this shady it."
T-SNE visualization RNN

Accuracy classifier: 88%
Future work

➢ Improve classifier accuracy
➢ Improve coherence generated texts
➢ Visualize differences
➢ Field trip Ghana
  ○ Interview storytelling experts
➢ Second human evaluation

Source: http://ramenrainbow.blogspot.com
Questions?

And they lived happily ever after - as in, like 'forever'!

Source: https://madoutherere.wordpress.com/