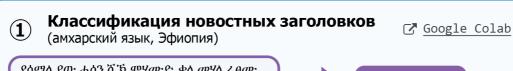
Кирилл Карпенко

## Проекты

GitHub: mssrchapelier



Чтобы ссылки стали активными при просмотре этого файла на GitHub, пожалуйста, загрузите файл.

የሶማሊያው ሐሰን ሼኽ ምሃሙድ ቃለ መሃላ ፈፀሙ world news (Somalia's Hassan Sheikh Mohamud swears in)

ኢትዮጵያ በአፍሪካ ዋንጫ የጣጣርያ ጨዋታ ግብፅን 2 - 0 አሸነፈች

(Ethiopia defeats Egypt 2-0 in Africa Cup of Nations qualifiers)

sports

Python NumPy, pandas

Технологии: SciPy

matplotlib, seaborn, statannotations

## Внутритекстовые ссылки: 2 конвертация формата

了 GitHub

... крупнейшие неразмеченные включают корпус Walta Information Center - WIC (описан в [4])... и корпус An Crúbadán [30] в 17 миллионов слов...

 $\begin{tabular}{ll} \blacksquare & \begin{tabular}{ll} \blacksquare & \begin{tabular}{ll} A & \bed$ 2007, pages 104-110, 2007.

lacktriangledown ... крупнейшие неразмеченные включают корпус Walta Information Center - WIC (описан в [Argaw and Asker 2007])... и корпус An Crúbadán [Scannell 2007] в 17 миллионов

🖹 Argaw and Asker 2007 — Atelach Alemu Argaw and Lars Asker. An Amharic stemmer... In Proceedings ... - ACL 2007, pages 104-110, 2007.

Технологии: Python, regex

## Морфологический парсинг **(3**)

☑ GitHub

(язык тигре, Эфиопия/Эритрея)

ለመጽአው

la - maS> - aw DEF - mS>:PRF - 3.M.PL

እትሕጹይ

>t - HSuy CAUS - (HSy).PTCP.PASS.M.SG

Технологии: Java, regex

## **(4**) Затекстовые ссылки: сегментация

☑ GitHub

🖹 Аркадьев П. М. О некоторых особенностях склонения в адыгских языках // Плунгян В. А. (отв. ред.). Язык. Константы. Переменные: Памяти Александра Евгеньевича Кибрика. СПб.: Алетейя, 2014. С. 552—563.

Кибрика",

"pagination": "C. 552-563", roity": "СПб.", "year": "2014", "collection-editors": "Плунгян В. А. (отв. ред.)",
"collection-title": "Язык. Константы. Переменные: Памяти Александра Евгеньевича "publishers": "Алетейя", "article-title": "О некоторых особенностях склонения в адыгских языках", "authors": "Аркадьев П. М."

Технологии: Java, regex