THE UNIVERSITY OF WESTERN AUSTRALIA

SEEK WISDOM

NovelPersepective

Identifying point of view characters

Lyndon White, Roberto Togneri, Wei Liu, Mohammed Bennamoun

What? Why? Why are you doing this to books?

Many novels, especially epic fantasy series, are written from the Point of View (POV) of many different characters.

They feature parallel sub-stories tracking the journey of each POV character.

Readers sometimes wish to read just one character's story; for example, on a second read through.

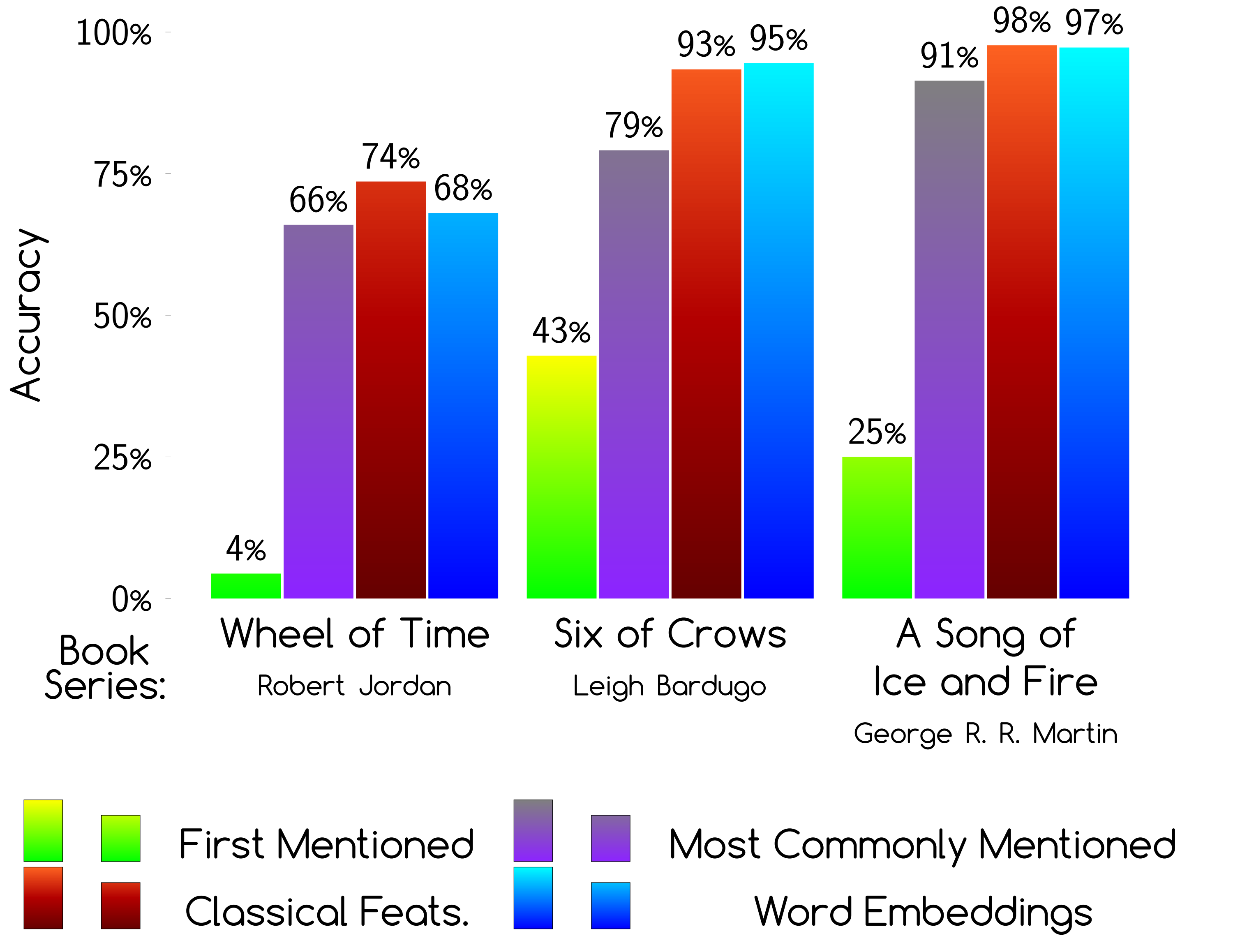
We have made a tool that allows the user to slice-up and restitch their ebooks around each POV character.

The challenging part is that most books do not label the sections with the name of the POV character, rather the reader works it out.

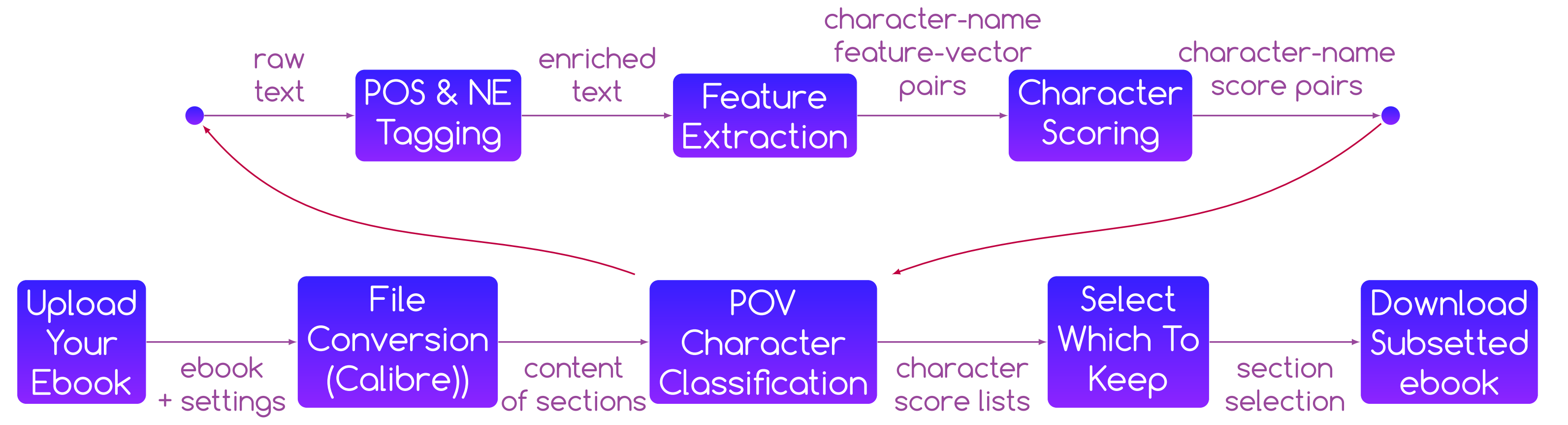
The source code is publicly available

MIT Licensed
https://github.com/oxinabox/NovelPerspective
Built on CherryPy, NLTK, Scikit-Learn, EbookLib and Calibre

Results on popular series of novels



The process for subsetting ebooks by POV



Baseline methods for determining POV

First Mentioned Named Entity

Features: first occurrence of named entity token in the section.
Scoring: earliest mentioned scores highest, $S_i = 2^{-rank(f_i)}$
Result: terrible. Other named entities often occur before POV.

Most Commonly Mentioned Named Entity

Features: count of occurrences of named entity token in section.
Scoring: most mentioned scores highest, $S_i = \frac{f_i}{\sum_{\forall j} f_j}$
Result: generally solid, but fooled by descriptions focusing on others.

Machine learning methods for determining POV

Classical Features + Logistic Regression

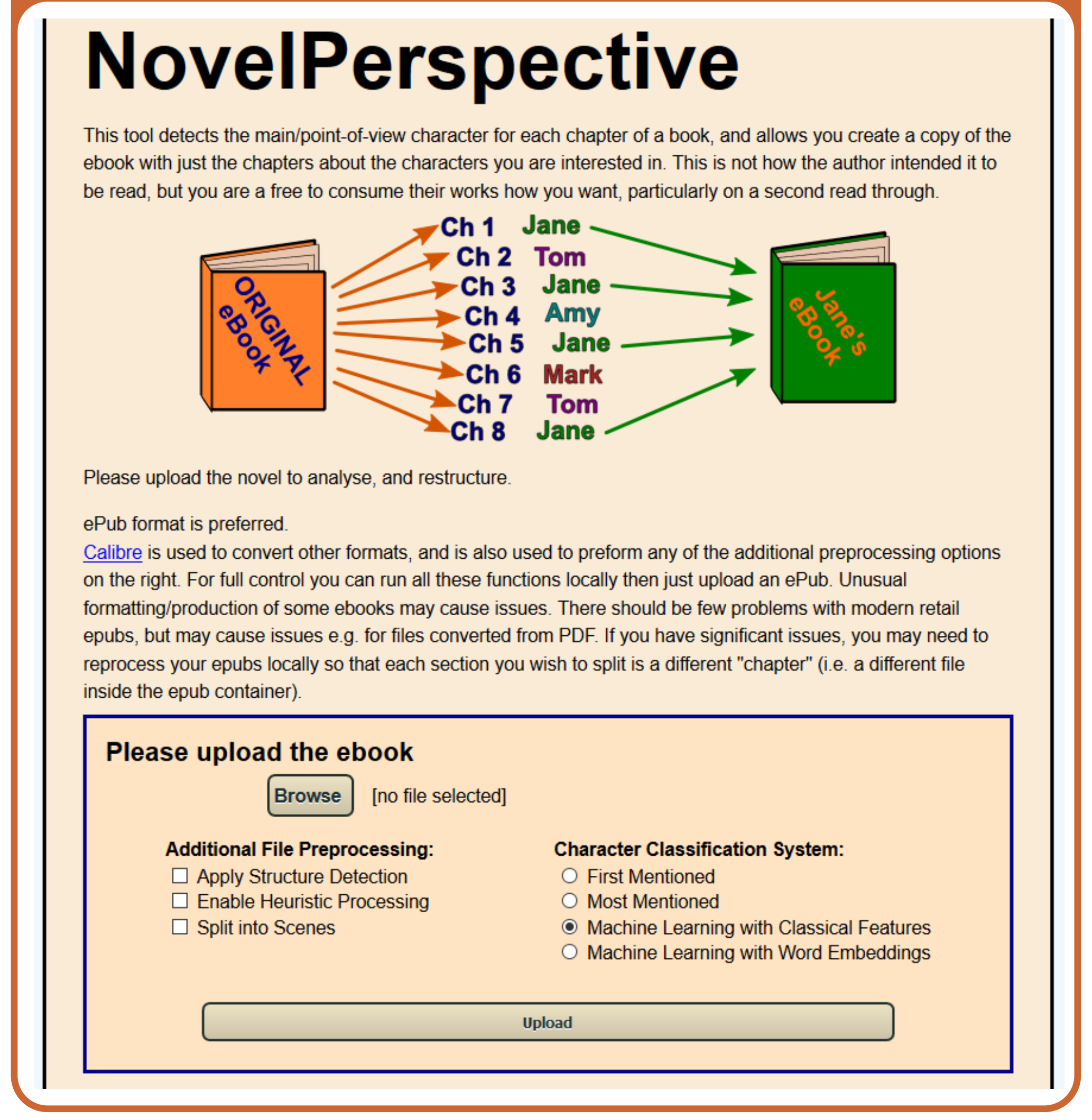
Features: position, and occurrence frequency, plus parts of speech co-occurring frequency. Total 200 dims.
Scoring: use logistic regression model on if POV or not, $S_i = \frac{P(f_i)}{\sum_{\forall j} P(f_j)}$
Result: generally great. Main characters occur near verbs and grammar. This gives an edge over frequency information alone.

Word Embeddings + RBF-SVM

Features: concatenation of FastText word embeddings for the adjacent words, averaged over all occurrences. Total 600 dims.
Scoring: use RBF-SVM model on if POV or not, $S_i = \frac{P(f_i)}{\sum_{\forall j} P(f_j)}$
Result: generally great. However due to high dimensionality, this method needs a lot of training data from other labelled books.

What does it look like?

Upload your ebook



NovelPerspective

This tool detects the main/point-of-view character for each chapter of a book, and allows you create a copy of the ebook with just the chapters about the characters you are interested in. This is not how the author intended it to be read, but you are a free to consume their works how you want, particularly on a second read through.

Please upload the novel to analyse, and restructure.

ePub format is preferred.
Calibre is used to convert other formats, and is also used to preform any of the additional preprocessing options on the right. For full control you can run all these functions locally then just upload an ePub. Unusual formatting/production of some ebooks may cause issues. There should be few problems with modern retail epubs, but may cause issues e.g. for files converted from PDF. If you have significant issues, you may need to reprocess your epubs locally so that each section you wish to split is a different "chapter" (i.e. a different file inside the epub container).

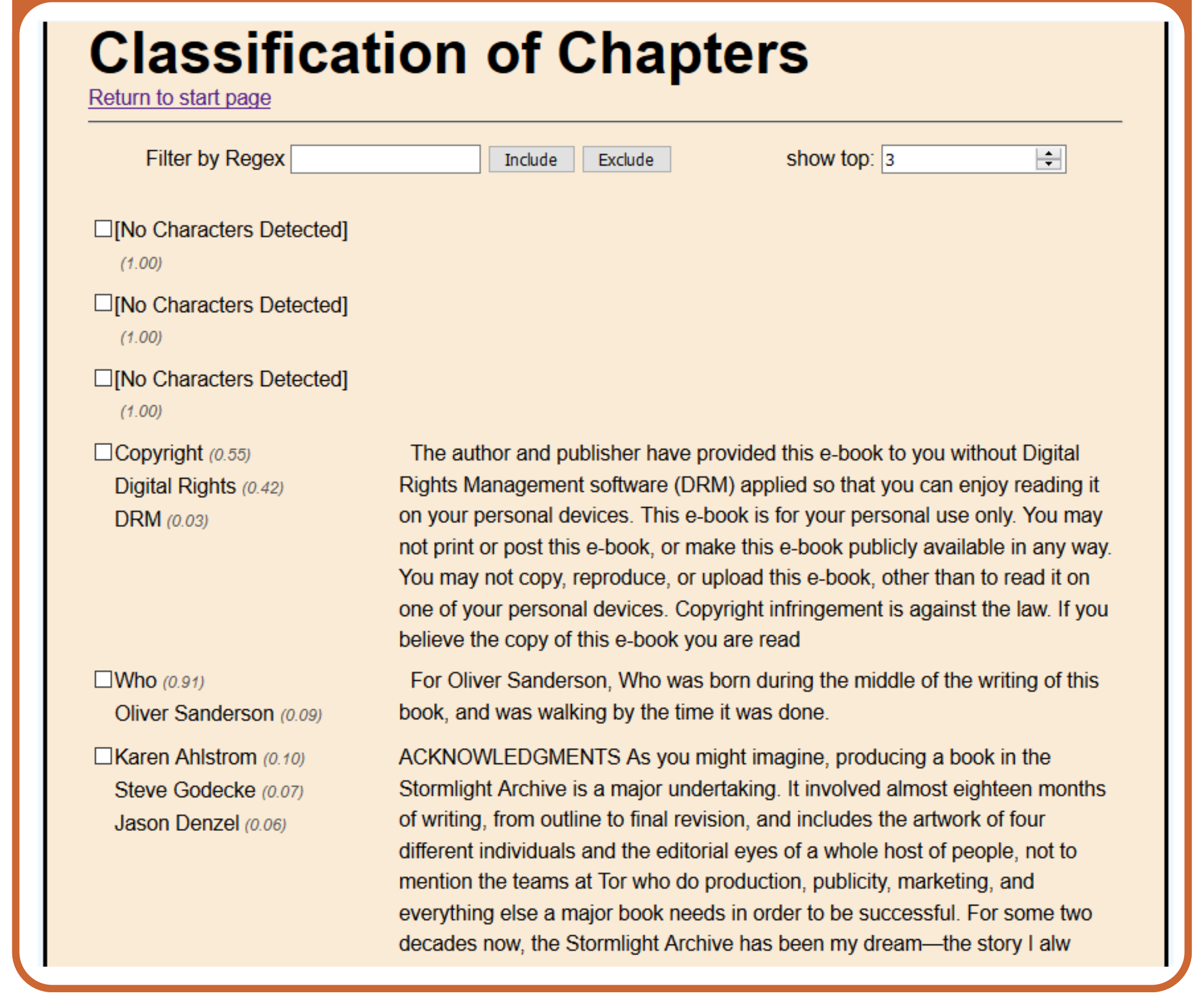
Please upload the ebook

[no file selected]

Additional File Preprocessing:
☐ Apply Structure Detection
☐ Enable Heuristic Processing
☐ Split into Scenes

Character Classification System:
○ First Mentioned
○ Most Mentioned
● Machine Learning with Classical Features
○ Machine Learning with Word Embeddings

Select which sections to keep



Classification of Chapters

Return to start page

Filter by Regex show top: 3

☐[No Characters Detected] (1.00)

☐[No Characters Detected] (1.00)

☐[No Characters Detected] (1.00)

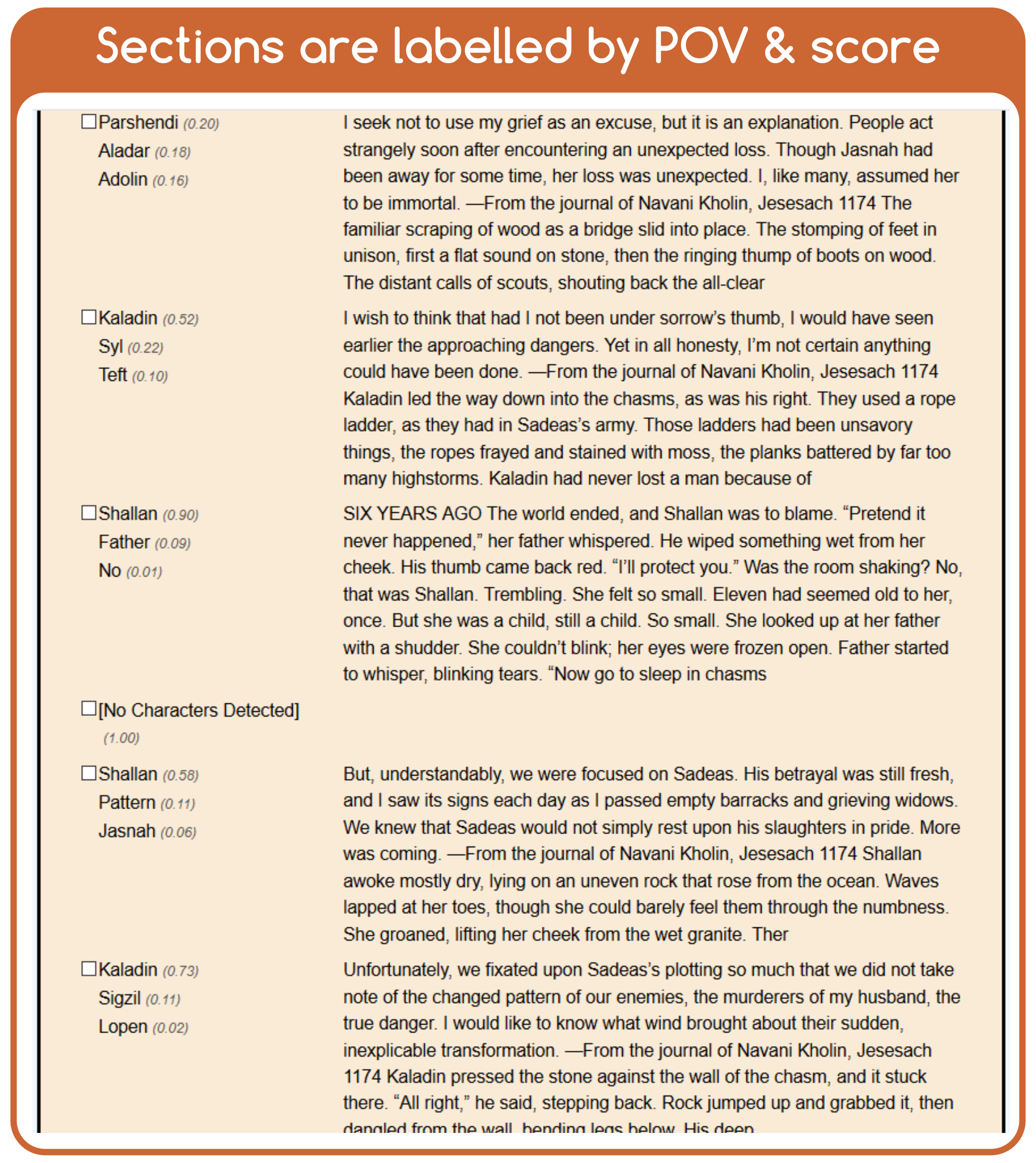
☐Copyright (0.55)
Digital Rights (0.42)
DRM (0.03)
The author and publisher have provided this e-book to you without Digital Rights Management software (DRM) applied so that you can enjoy reading it on your personal devices. This e-book is for your personal use only. You may not print or post this e-book, or make this e-book publicly available in any way. You may not copy, reproduce, or upload this e-book, other than to read it on one of your personal devices. Copyright infringement is against the law. If you believe the copy of this e-book you are read

☐Who (0.91)
Oliver Sanderson (0.09)
For Oliver Sanderson, Who was born during the middle of the writing of this book, and was walking by the time it was done.

☐Karen Ahlstrom (0.10)
Steve Godecke (0.07)
Jason Denzel (0.06)
ACKNOWLEDGMENTS As you might imagine, producing a book in the Stormlight Archive is a major undertaking. It involved almost eighteen months of writing, from outline to final revision, and includes the artwork of four different individuals and the editorial eyes of a whole host of people, not to mention the teams at Tor who do production, publicity, marketing, and everything else a major book needs in order to be successful. For some two decades now, the Stormlight Archive has been my dream—the story I alw

Demonstration on "*Words of Radiance*" by Brandon Sanderson.
Using the classical features method.

Sections are labelled by POV & score



☐Parshendi (0.20)
Aladar (0.18)
Adolin (0.16)
I seek not to use my grief as an excuse, but it is an explanation. People act strangely soon after encountering an unexpected loss. Though Jasnah had been away for some time, her loss was unexpected. I, like many, assumed her to be immortal. —From the journal of Navani Kholin, Jesesach 1174 The familiar scraping of wood as a bridge slid into place. The stomping of feet in unison, first a flat sound on stone, then the ringing thump of boots on wood. The distant calls of scouts, shouting back the all-clear

☐Kaladin (0.52)
Syl (0.22)
Teft (0.10)
I wish to think that had I not been under sorrow's thumb, I would have seen earlier the approaching dangers. Yet in all honesty, I'm not certain anything could have been done. —From the journal of Navani Kholin, Jesesach 1174 Kaladin led the way down into the chasms, as was his right. They used a rope ladder, as they had in Sadeas's army. Those ladders had been unsavory things, the ropes frayed and stained with moss, the planks battered by far too many highstorms. Kaladin had never lost a man because of

☐Shallan (0.90)
Father (0.09)
No (0.01)
SIX YEARS AGO The world ended, and Shallan was to blame. "Pretend it never happened," her father whispered. He wiped something wet from her cheek. His thumb came back red. "I'll protect you." Was the room shaking? No, that was Shallan. Trembling. She felt so small. Eleven had seemed old to her, once. But she was a child, still a child. So small. She looked up at her father with a shudder. She couldn't blink; her eyes were frozen open. Father started to whisper, blinking tears. "Now go to sleep in chasms

☐[No Characters Detected] (1.00)

☐Shallan (0.58)
Pattern (0.11)
Jasnah (0.06)
But, understandably, we were focused on Sadeas. His betrayal was still fresh, and I saw its signs each day as I passed empty barracks and grieving widows. We knew that Sadeas would not simply rest upon his slaughters in pride. More was coming. —From the journal of Navani Kholin, Jesesach 1174 Shallan awoke mostly dry, lying on an uneven rock that rose from the ocean. Waves lapped at her toes, though she could barely feel them through the numbness. She groaned, lifting her cheek from the wet granite. Ther

☐Kaladin (0.73)
Sigzil (0.11)
Lopen (0.02)
Unfortunately, we fixated upon Sadeas's plotting so much that we did not take note of the changed pattern of our enemies, the murderers of my husband, the true danger. I would like to know what wind brought about their sudden, inexplicable transformation. —From the journal of Navani Kholin, Jesesach 1174 Kaladin pressed the stone against the wall of the chasm, and it stuck there. "All right," he said, stepping back. Rock jumped up and grabbed it, then dangled from the wall, bending legs below. His deep

http://novelperspective.ucc.asn.au/