Devil's Advocates: Midterm
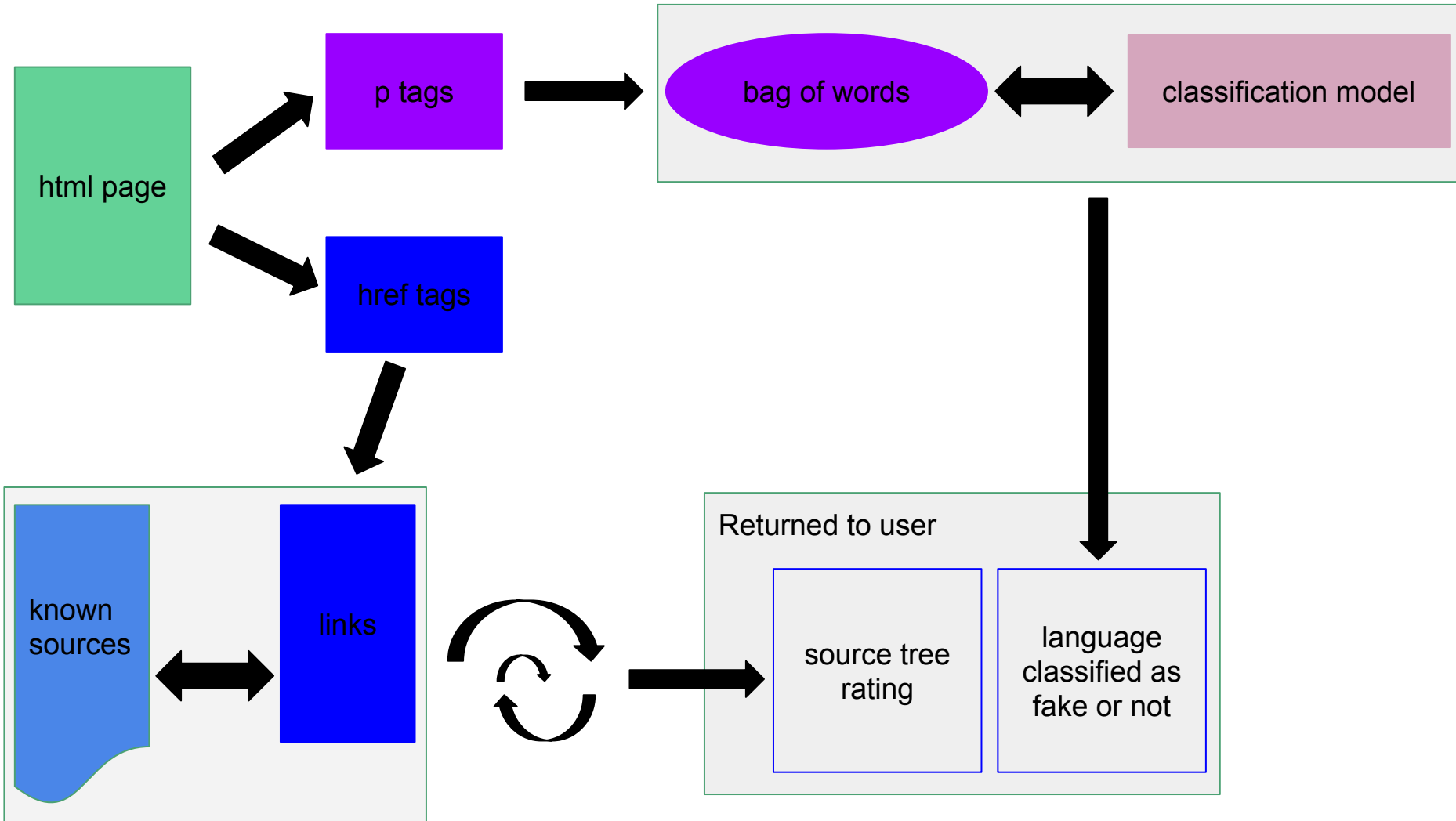
James Brothers, Julia Evans, Joseph Turner

Road Map

- Where we are now
  - Demos

- What's next

- Plan of activities

```
html page  →  p tags  →  [ bag of words  ⟷  classification model ]

html page  →  href tags  →  [ known sources  ⟷  links ]  ↻  →  Returned to user [ source tree rating | language classified as fake or not ]
```

html page

p tags

bag of words

classification model

href tags

known sources

links

Returned to user

source tree rating

language classified as fake or not

Where we are now: Practicum 1

- Practicum 1
  - LinkTree

Where are we now: Link Analysis

- LinkTree => LinkAnalyzer

Where are we now: whoisit

- Gives information about who owns a website

- Websites from certain owners and certain countries are suspicious

- For example (from our data) websites from Macedonia or Scottsdale Arizona are often Fake or Imposter news sites

What's next

- Analysis
  - LIWC and bag of words method
- R-Studio

Analysis Approach

Classification dictionary

Article	"Phrase 1"	"Phrase 2"	"Phrase 3"	...	...	Biased	Fake
name	count/wc	count/wc	count/wc			0 | 1	0 | 1
.	.	.	.	.	.	.	.
.	.						
.		.					
.			.				
.				.			

Plan of Activities

- Still on track with original plan
  - With a few tweaks to what we're doing
  - No longer focused Facebook API
- 11/9 - 11/16: Practicum 2 (Julia)
  - Data Analysis and R-studio
- 11/20: Full functionality complete (EVERYONE)
  - Language analysis complete (Julia)
  - Source Tree (time permitting)
- 12/1
  - Final Presentation complete
- 12/5 - 12/7: Final Project Presentation
  - Refined functionality (James)
  - Refined user interface (Joseph)

Questions?