## ECS 193AB Winter/Spring2017

## Realtime Podcast RSS Feed Parser for ShortOrange.com

**Background/Introduction/Business Value:**

ShortOrange is a cross platform podcast player that allows people to listen, manage, and organize their favorite podcasts. One of the hardest challenges for our site is to keep all 10-20k podcasts feeds up to date in realtime so we can send push notification and send updates to peoples devices. We currently have around 5k to 10k thousand users so your code will go into production and be used by real users if you succeed. If you enjoy scalability, cross server communication, and performance challenges you will love this project. Also, a great talking point for job interviews.

**Description/Design Issues/Project scope:**

Create a rss job processing queue that can scale across multiple machines.

**Requirements:**

1) Rss Parser: Podcasts are sometimes in different formats, using different xml tags for the same data. We need to support the majority of podcasts supported by iTunes. Example feed: http://feeds.serialpodcast.org/serialpodcast?format=xml (http://feeds.serialpodcast.org/serialpodcast?format=xml)

2) Scalable / Performant ( +- 15-60 mins delay ): People don't want to be notified a new podcast is available 5 hours after it is released. Server process should be able to run on multiple machines.

3) Reliable: Support a smart retry policy and remove unhealthy podcasts from the queue.

**Deliverables:**

Working multi-server scalable rss podcast parser that can replace our current version.

**Contact:**

Josh Ketellapper, jjketellapper@gmail.com (mailto:jjketellapper@gmail.com) – UC Davis Alumni, currently working at Facebook, can meet as much as needed but hopefully once a week.