Let's Talk Open-Source -- An Analysis of Conference Talks and Community Dynamics

Kimberly Truong
Oregon State University, USA
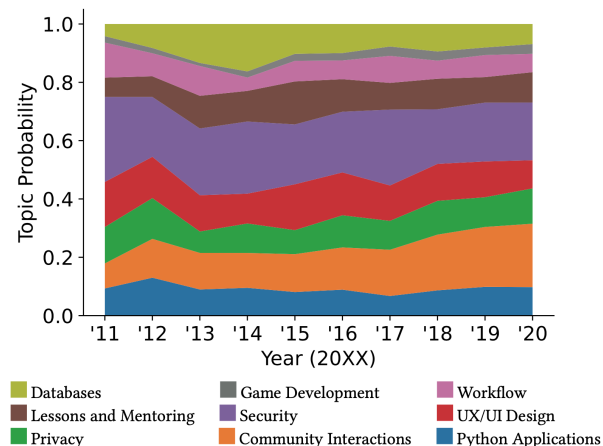
Oregon State University

Motivations

o There is a gap between open-source researchers and practitioners.
o Community interactions in a project are an indication of the project's success. So what do practitioners choose to share?
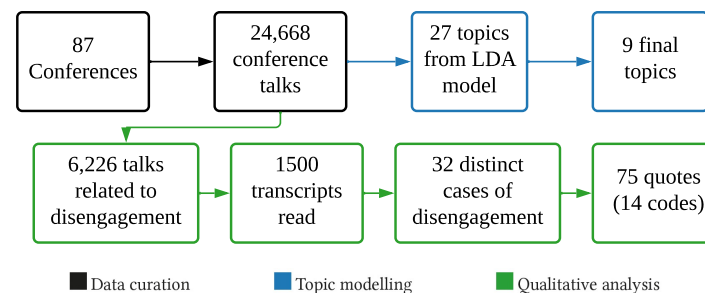o How can we further sustain and support open-source software?

Source of Grey literature

o Large and diverse untapped source (24,669 talks) publicly available on YouTube
   o Disciplines – different tools, specialties, fields
   o Size – 50 to 4,000 attendees
   o Experiences/roles – contributor, core maintainer, user, CEO
   o Regions – North America, Europe, Australia, parts of Africa
   o Times – 10 years (2010 - 2021)
o Provides first-hand accounts from practitioners

Distribution of nine most common topics in open-source software from 2011-2020 from LDA model.

Methodology

■ Data curation   ■ Topic modelling   ■ Qualitative analysis

Curating a Dataset

1. Establish a conference list through Google search
   o Get metadata from top 30 search results
   o Add conferences that meet our constraints:
      o Two documented editions
      o 50 attendees or 10+ speakers/talks
2. Extract video data.
3. Establish a sample by filtering with keywords:
   o 'leave', 'abandon', 'pause', 'hostility', etc.
4. Topic model over all transcripts.

Analyzing a Disengagement Sample

1. Thematically code relevant transcripts with with two people.
2. Record reasons cited for disengagement, GitHub activity, and recommended interventions

Cultural (32%)	External (18%)	Volunteering (50%)

Contributions

Dataset of 24,669 Open-Source Software Talks

o Name of video
o Publication date
o Playlist (conference edition)
o Description
o Transcript
o YouTube URL

o Conference metadata
o Scripts to extract YouTube video data

Reasons and Interventions for Disengagement

o Promote inclusive communities

Not Enough Time (Internal)   No Longer Enjoyable

Who?

Github:

Project: Django

Why?

Anna cited three reasons for leaving their open source communities at DjangoCon US in 2018. These reasons arose because they had to change their priorities as new events unfolded in their life. The first issue was the lack of a work life balance as newer, more important responsibilities arose.

"People's living situations change or your priorities change and so people may say after a while I can't do it anymore for whatever reason." (15:44)

"Tech kind of takes over our whole life. I'm definitely guilty of that so I'm kind of like actively saying no." (30:00)

This led to a lack of time internally due to their shift in priorities. Although they still appreciated the Django community, Anna found contributing to open source to take away from more significant priorities, like a new job and changing living situtations.

"I'm actually moving cities so I also kind of want to prioritize my personal life a little more I've served the community a lot and now it's kind of like time to look after myself a little more I'm out I also took on this

Screenshot from Disengagement Diaries (https://disengagement-diaries.github.io) showing a contributors reasons for disengagement.