InSite and Short Text

Reaching Into an Interview and Speeding Input

Kim Patch

Boston, MA, USA, kim@scriven.com

Introduction

I'll demonstrate two related tools for journalists – the InSite open source publishing system for interactive transcripts [1] and Short Text, a systematic shorthand for paper and keyboard.

What if, from the text of your story, you could point not just to the beginning of video, but to any particular sentence in that video? InSite makes it easy for journalists to link to source material at the sentence level so readers can click to see a quote in the full context of the source interview. Enabling this doesn't require manually adding timecodes. The system works for videos, podcasts and also text.

InSite Interactive Transcripts

I worked with Duke University professor and former Washington Post managing editor Phil Bennett to enable sentence-level linking for the Rutherfurd Living History project [2] at Duke.

The InSite software that we built to do this is an open source template for WordPress. The system makes it easy for content providers to attach supporting content – text, documents, audio, video, maps and links – to any particular sentence. It also contains some sophisticated search functions, including cross-collection searches, and a hit map that shows up on the video seek bar so you can see the pattern of hits within an interview.

We also helped PBS Frontline implement a version of the server software as part of their transparency project [3]. We've enabled interactive transcripts for six documentaries: Putin's Revenge, The Facebook Dilemma, Trump's Showdown, Supreme Revenge, Zero Tolerance and America's Great Division. So now anyone can point to any sentence in any of those 218 interactive source interviews.

Here are links to the Duke oral history collection and one of the Frontline interactive transcript projects:

- Duke University's Rutherfurd Living History collection

Here's an example of a direct link from that collection:
> I think that one of the big issues we're going to have to talk about in our country, that isn't being talked about enough, is the fact that we're incarcerating larger percentages of our people. And that is eating up a lot of our economic strength.
> http://livinghistory.sanford.duke.edu/interviews/cory-booker/#3610

- The Frontline Interviews: America's Great Divide

Here's an example of a direct link from that collection:
> The economic collapse of 2008 and the economy that Barack Obama inherited, I think, is the defining event of this generation, even more so than 9/11, and it profoundly reshaped American politics.
> https://www.pbs.org/wgbh/frontline/interview/steve-schmidt/#558

- There are other examples here:
patchontech.com/interactive-transcripts

1.1 Sentence-Level Linking

A key InSite innovation is sentence-level linking. It's possible to copy a URL in a YouTube video that includes a timestamp to a particular moment in the video [4]. But it's tedious to collect a list of direct URLs that

    a.   Start at the beginnings of sentences.
    b.   Include labels or excerpts from the transcript so it's clear what the links go to.

The InSite system allows researchers, journalists and other viewers to use a well-formatted transcript to rapidly collect linked excerpts. The annotation features allow content creators to efficiently place any type of media alongside any sentence in an interview, including explanatory text. The advanced search features make it easy to find and compare quotes within and among interviews.

The transcript and annotation workflows don't require content creators to manually input timestamps. This makes it possible to enable large amounts of content. [5] [6]

We want to encourage others to make video, audio and text accessible at the sentence level. We want to increase the available ecosystem of source interviews that people can reach into and make part of their stories. We also want to encourage programmers to contribute to the open-source project.

1.2 Workflow

I'm working on improving the tools we use to help content creators along the whole workflow, including recording, organizing, sharing locally, and posting. One thing I pay a lot of attention to is how the tools fit together into an efficient system.

The part of the workflow that needs to be improved the most is transcription. Automatic transcription is useful, but falls short when you want a near-perfect transcript including punctuation. Today's best speech engines still make hundreds of mistakes per hour of audio, even given high-quality recordings.

Short Text

I'm also working on Short Text, a shorthand that sidesteps the type-ahead error issues that frustrate people with typing on phones. If you learn the system you can use it on both paper and keyboard. On the keyboard the shortcuts expand into text. It speeds all kinds of writing including transcription.

REFERENCES

[1] Patch, K. (2018). The InSite System: Enabling Transparency with Searchable, Shareable Interactive Transcripts, presented at I Annotate, San Francisco, 2018.

[2] Rutherfurd Living History, Duke University, livinghistory.sanford.duke.edu.

[3] PBS FRONTLINE Transparency Project, http://www.pbs.org/wgbh/frontline/about-frontlines-transparency-project.

[4] How to Link to a Certain Time in a YouTube Video, wikiHow, wikihow.com/Link-to-a-Certain-Time-in-a-YouTube-Video.

[5] Our Research, Rutherfurd Living History, Duke University, livinghistory.sanford.duke.edu/interactives/our-research.

[6] Interactive Transcripts, patchontech.com/interactive-transcripts.