Creating Crowdsourced Research Talks at Scale

Rajan Vaish*
Stanford University
rvaish@cs.stanford.edu

Shirish Goyal
Stanford University
shirish.goyal@stanford.edu

Amin Saberi
Stanford University
saberi@stanford.edu

Sharad Goel
Stanford University
scgoel@stanford.edu

ABSTRACT

There has been a marked shift towards learning and consuming information through video. Most academic research, however, is still distributed only in text form, as researchers often have limited time, resources, and incentives to create video versions of their work. To address this gap, we propose, deploy, and evaluate a scalable, end-to-end system for crowdsourcing the creation of short, 5-minute research videos based on academic papers. Doing so requires solving complex coordination and collaborative video production problems. To assist coordination, we designed a structured workflow that enables efficient delegation of tasks, while also motivating the crowd through a collaborative learning environment. To facilitate video production, we developed an online tool with which groups can make micro-audio recordings that are automatically stitched together to create a complete talk. We tested this approach with a group of volunteers recruited from 52 countries through an open call. This distributed crowd produced over 100 video talks in 12 languages based on papers from top-tier computer science conferences. The produced talks consistently received high ratings from a diverse group of non-experts and experts, including the authors of the original papers. These results indicate that our crowdsourcing approach is a promising method for producing high-quality research talks at scale, increasing the distribution and accessibility of scientific knowledge.

1 INTRODUCTION

There is growing demand for learning and consuming scientific information through video [38]. This demand has in part been met by MOOCs [20], which typically focus on in-depth presentations of established areas, and by efforts such as "Two Minute Papers" and "Papers We Love", which distill scientific ideas for viewers with limited technical expertise. But the vast majority of contemporary research is still available only in the form of text, as traditional academic papers, in part because individual researchers often have limited time, resources, and incentives to produce video-based summaries of their work. This gap prompts a challenge: distilling the content of academic papers into short presentations suitable for students and researchers, and doing so at scale.

Here we introduce and evaluate a system for creating an open, multilingual repository of 5-minute lightening talks developed collaboratively by volunteers worldwide. These talks are catered to technically knowledgeable viewers, who after watching the video summary might read the original papers or attend a longer conference presentation. The initial videos are produced by distributed teams of individuals working in close collaboration; the videos can subsequently be edited and improved by any interested participant. Our project increases the accessibility of scientific knowledge by converting English-language research papers into video-based talks produced in multiple languages—and all without involving the authors of the paper or other domain experts. In the process of creating this content, volunteer contributors learn collaboratively, furthering educational opportunities and incentivizing participation.

Crowdsourcing such an open-ended expert task poses two key challenges. First it is not immediately clear how to collaboratively produce editable videos. Second, one must facilitate extended and complex coordination between large, distributed groups of individuals of varying expertise. To address the fi rst task, we standardize each talk to consist of slides, a written script, and voice-overs; we then programmatically stitch these components together to produce a complete video presentation. To streamline this process, we created an online tool that lets people collaborate and seamlessly record audio on a slide-by-slide basis. Our modular approach supports efficient editing and reduces retake time, both during and after the initial videos are created. We address the second challenge by designing a structured scaffolding process to coordinate volunteers [28, 41]. Specifically, we divide the talk creation process into three discrete phases spanning a period of 21 days (three weeks): (1) on-boarding the crowd and forming teams; (2) generating a slide deck that includes both the talk slides and a slide-by-slide script of the talk; and (3) converting the script to slide-by-slide audio recordings, and reviewing the complete video presentation.

To test this system, we issued an open call for participation, attracting 840 people from 52 countries. This crowd of volunteers created 107 lightening talks in 12 languages based on 40 recent papers from top-tier computer science conferences and scientific journals. These talks were entirely created by the crowd, from designing the structure to producing the content. To evaluate the talks,