

CONCRETE PROJECT PLAN

EXISTING LOCAL VOICES NETWORK PLATFORM

MIT CCC and Cortico have been designing, building, testing, and deploying the [Local Voices Network \(LVN\) platform](#) for four years (see the [LVN Overview](#) and [Freeport](#) and [Boston](#) Case Study docs uploaded and at <https://www.ccc.mit.edu/future-fund/> for background and impact).

LVN's key elements include:

1. **Recorded small-group conversations about participants' lived experiences** – which tap into ancient, time-proven social dynamics that are largely suppressed today, and which people find meaningful and enjoyable as they realize there are few opportunities for such conversations elsewhere in their lives.
2. **Cross-pollination by playing conversation highlights from one conversation to another** – which can happen in specific conversations, or by hearing amplified highlights, and which empirical evidence has shown increases mutual respect across divides.
3. **Cross-conversation pattern analysis to identify, understand, and surface connections across people's experiences** – which we believe is a significant improvement over traditional public opinion research.
4. **Scalable transparency that makes it easy for the public to see how community input is being interpreted and used in public processes** – which we believe will improve public perception of legitimacy of the decision- and policy-making processes.
5. **Design for decentralized access and visibility control by those organizing the conversations** – which encourages community agency.

Critical in terms creating a development feedback cycle, we have several LVN deployments currently in flight, including:

- With support from their District Superintendent and Principal, High School students in Freeport, Maine facilitated recorded conversations on the experiences, thoughts, and feelings of 50+ peers. Student leaders owned every step of this process from project planning to question design and they used LVN to identify over 220 conversation highlights to share directly with administration, teachers, and fellow students to hear. Shared content will be used by leadership to inform strategic planning efforts and make hiring decisions.
- [RealTalk@MIT](#) is a new community-based civic space for deep and substantive conversations among faculty, students, and staff to ensure that the Institute's ongoing values initiative accurately represents the thoughts, concerns, and aspirations of the entire MIT community. Through participation in small, virtual, facilitated dialogues that will be recorded, community members will have the opportunity to share their diverse lived experiences at MIT, hear the stories of others, and reflect on the connections, tensions, and commonalities that emerge. A team of MIT students, staff, and faculty will identify themes and patterns that emerge from conversation recordings and share findings back to the entire MIT community.

- The City of Cambridge recently collected input on the selection of a new City Manager from over 120 Cambridge residents. Everyone from small business owners, to faith community leaders, to racial justice organizers has come together to weigh in on the hiring process. Vice Mayor Alanna Mallon has been actively making highlights on the LVN platform, sharing some [powerful examples](#) of experiences and feedback shared by Cambridge residents. Input has been directly incorporated into the job description itself.
- Child Poverty Action Lab is collaborating with Cortico and MIT CCC to launch an effort to collect, make sense of, and share conversations with over 1,000 underheard Dallas voices. The project will inform a roadmap, resource allocation, and community action to significantly reduce child poverty in the upcoming years.

ADDRESSING THE TWO BOTTLENECKS IN THE SYSTEM

Through our experience in deploying and iterating LVN over these four years, we have encountered two primary bottlenecks that impede mainstream adoption and define the work we're doing.

First, in order to hold an LVN conversation, our current model requires partner organizations (e.g., community-based organizations or municipalities) to gather small groups (in person, and since the pandemic, over Zoom), provide a trained facilitator, and upload audio recordings of the conversation onto our platform. This approach is limited by the capacity and expertise of partner organizations. In addition, by solely relying on organizations the system does not leverage the network power of self-organized communities of trust.

Second, in order to make use of LVN conversation recordings – either to find audio highlights that can be shared with other groups for cross-divide listening, or to analyze patterns across conversations as a way to synthesize insights and guide public decision making – our current model requires a high level of manual effort to sift through audio recordings and transcripts supported by only basic keyword search and keyword-based topic classification tools.

PROPOSED SOLUTIONS

These bottlenecks motivate two work streams:

Stream 1: Design of a mobile social network app optimized for teens

(For more background on the Blink app, see the [Overview](#) and [Update](#) docs uploaded and at <https://www.ccc.mit.edu/future-fund/>.)

Just as text messaging or Facebook groups are formed bottom up, we envision a new participant pathway to LVN that is user-driven and does not require top-down organizational support. We have begun the design of a mobile-first social network app – code named “Blink” – based on the core design concepts that have emerged in LVN (small-group facilitated dialogue, speech recording and sharing under user consent, human + machine analysis of patterns across conversations for communication across divides). We anticipate Blink will be as easy to download, join, and gather conversations as any mainstream social network app. But in contrast to current mainstream social media platforms that promote viral content and audience growth, Blink inherits its design philosophy from LVN, with a focus on small-scale (2-6 person) live, unrehearsed conversations to foster

nuanced sharing of personal stories and perspectives in a safe, trusted space. Blink will incorporate state-of-the-art speech analytics to streamline selection and sharing of conversation highlights within the app. We are taking great care to design user-controlled mechanisms for lifting excerpts from conversations that can be shared more widely, or combined for sense-making across conversations.

Based on testing and feedback from teens over the past year, we are now designing wireframes for a first version of the Blink app with an initial design scope that will permit Blink to replace Zoom as a front end for LVN. By doing so, we will address the first bottleneck described above. Blink will include features to support group formation, scheduling and holding live audio/video conversations, realtime highlighting of conversation experts, and user privacy and data sharing controls.

In late 2021, our staff engineering team prototyped a Blink cross-platform mobile app capable of recording conversations from multiple participants, recording real-time reactions, interactions, and transcribing the system for review. The application was constructed and deployed with a modern development stack (React Native, Twilio, and Amazon's speech-to-text system). In addition to the deep research talent MIT CCC brings to the table, our staff engineering team has the industry experience and development processes to move fast, like a startup. We have a fairly aggressive development timeline to build an MVP that will include core conversation features, a public feed designed to encourage cross discussion between groups, and the ability to support a few thousand users.

The Blink user's perspective will be qualitatively much different from that of an LVN participant. However, the fundamental structure of the data is such that we will be able to feed Blink conversations into the existing LVN technology stack, including its data pipelines, advanced speech processing, and natural language processing (NLP) systems. As such, we'll be able to leverage the years of LVN research and development on the backend while focusing on creating a delightful and empowering app experience for the user.

We are seeking support to accelerate engineering efforts to build the app, which will be fully compatible with the LVN platform, with a goal of piloting a first version of the Blink app with a network of high school students in late fall 2022 or spring 2023.

Stream 2: Harness large-scale language models for conversation analytics

A key feature of the LVN system is its ability to identify themes of lived experience across collections of conversations. To date, LVN thematic analysis has served as an input for leadership selection, as a basis for engaging candidates for public office in public dialogue, and as a basis for public policy recommendations. It also serves as a basis for individuals to understand their connections to others. To ease the burden of what is currently a highly manual process, we plan to develop advanced language modeling tools to augment human analysis.

Over the past year we have developed tools and processes to accelerate analysis of LVN conversation collections (see the [Analytic Sensemaking](https://www.ccc.mit.edu/future-fund/) overview uploaded and at <https://www.ccc.mit.edu/future-fund/>). The next challenge we face is that LVN conversations are replete with personal stories. Themes or patterns that connect stories often are not apparent in the surface form of words and phrases. Thus basic keyword search or keyword based topic models will fail to uncover connections that are apparent to human analysts.

Over the past few months we have run initial experiments in fine-tuning [GPT-3](#) – an AI system that can mimic human language – to automatically generate thematic analyses of LVN conversation highlights with very promising results. Encouraged by these early investigations, we propose a more extensive exploration in ways to harness modern large-scale language models in service of analyzing rich open-ended conversational data. When applied to large collections of lived experiences, we envision a new kind of public understanding that complements public opinion research with “public experience research” which we believe may ultimately play a more important role in our complex pluralistic democracy than polls and surveys by providing a systematic way for institutions to listen to people at scale.

Traditional sense-making tools include keyword-based search engines, which let the user find conversations that match their expressed interest; and topic models, which, in turn, yield a self-organizing map of the subjects being discussed. These, however, fall short of what most decisionmakers want when they approach a new conversational data set, which is a concise summary of why the conversations matter; how they add to what is already known; and how the discussions connect to outcomes in the world, such as public health policy. To address this:

1) We will first identify communities of interest from the conversational data collected by the LVN. Most LVN conversations are convened in a specific geographic location, giving us one readily available definition of community to use, but we are more interested in the deeper notions of community that transcend location. For example, we will model group demographics inferred from the way that speakers choose to identify themselves in these conversations.

2) We will [fine-tune large language models](#) corresponding to the words spoken within each of the communities of interest. We will first use the [OpenAI API to GPT-3](#), and then our own home-grown variants of it, to fine-tune a model for each of the communities’ associated transcribed speech. In particular, we will pair the self-identification speech for each participant with question and answer pairs extracted from the conversations in which the participant is the answerer. Using fine-tuning in this way we can thus push the model to answer questions in a manner consistent with the identities expressed in the conversations; we harness the immense world knowledge already encoded in the pre-trained base model of GPT-3, and overlay community-specific knowledge on top of it. Our home-grown variant of GPT-3 will be based on the smaller, open-source [GPT-J 6B](#) and trained on a curated subset of its Pile training data together with our own supplemental conversation data sets. This will give us greater control over the training data and allow us to investigate and remove sources of bias.

3) We will build a comprehensive tool for understanding the communities that includes:

- i) Semantic search and clustering: While we feel search is not *sufficient* in a conversation analysis tool, it is *necessary*, and our language models from (2) can help us do a better job than keyword search to identify key themes of interest. The analyst will then be able to enter a theme in natural language, even an abstract one such as “cases where someone expresses frustration over public health policy,” and get back an appropriate list of conversation snippets.
- ii) Natural language question answering: With the fine-tuned language models we can offer a natural language question-answering interface about the content itself. For example, users can ask “How does this community feel about police funding?” and get, in reply, a sensible summary of the conversations with respect to that topic.

iii) "Sounding Board": Even more usefully, the stakeholder can ask questions to be answered in the voice of the community members, or take simulated polls of them. For example, "Does this headline make you feel more optimistic about the city?" might be answered differently by different communities, depending on the headline. The tool will dispatch questions to each community in turn, and summarize the results and the differences between communities.