The Trevor Project: Building AI to Support LGBTQ Youth

Presenters: Taki Hasegawa (she/they)
Taki Hasegawa
(she/they)

- Senior Machine Learning Engineer
- Role: Align with stakeholders on AI/ML ideas, research and prototype, operationalize products
- Previously: Machine Learning Engineer @ Mercedes-Benz Research
- Why Trevor: Important mission, empathetic culture
- What makes you smile: snowboarding, Simon (my cat), playing video games!
<Content warning>
This presentation mentions suicide and mental health topics.
Introduction to The Trevor Project
Mission, Impact and Technology
The Trevor Project is the world’s largest suicide prevention and crisis intervention organization for LGBTQ young people.

Our mission is to end LGBTQ youth suicide — and it is urgent.
Our research shows that there are an estimated **1.8 million LGBTQ youth** who seriously consider suicide each year.
Crisis services
Direct suicide prevention and crisis intervention services to support LGBTQ youth 24/7 via phone, text, and chat

Peer support
The world’s largest safe space social networking community for LGBTQ youth

Research
Evaluations and external research that support The Trevor Project in significantly improving its services while maintaining preeminence in scientific inquiry

Education and public awareness
Programs, trainings, and content promoting awareness around issues and policies relevant to LGBTQ youth and the adults who support them

Advocacy
Advocacy at the federal, state, and local levels to fight for policies and laws that protect LGBTQ youth
Building AI at Trevor: A Narrative
The Crisis Contact Simulator and Beyond
Problem Space: **Training**

Volunteers go through a 40-hour virtual training over 10 weeks that covers counseling skills, LGBTQ identity, the structure of a conversation on our crisis services.
The Crisis Contact Simulator (CCS) is an AI chatbot designed to act like an LGBTQ youth in crisis.
CCS allows aspiring counselors to experience realistic practice conversations called role-plays.

Each trainee completes at least 4 role-plays, which were previously all human-led.

The Trevor Project used to train roughly 30 new counselors each quarter.

For human-led roleplays, trainees must schedule a time with a coordinator.

2 role-plays are now conducted with CCS.

With the help of CCS and other key initiatives, our team now trains over 200 counselors each month.

With CCS, trainees can complete the role-play on their own time.
The Evolution of CCS
From Prototype to Product
Chatbots are largely evaluated on their **sensibility** and **specificity**

**Sensibility**: Does the chatbot’s response make sense in context?  
**Specificity**: Does the chatbot avoid vague answers?

**Example**: How are you doing today?

**Sensible and Specific**

Human: How are you doing today?  
Chatbot: I’m doing well! I had a nice weekend.

**Not Sensible, but Specific**

Human: How are you doing today?  
Chatbot: My name is Wendy

**Sensible, but not Specific**

Human: How are you doing today?  
Chatbot: I don’t know...
Counselor: What’s going on?

Youth: I think I’m very gay

Youth: I think and like like like like like like and
Youth: thank you so much your friend meh used to have actually, he even picked me up when i was that way and came out to my parents and put me in a trash chanel

Youth: bag for the night and began, shit, dating him at the library cause everyone saw him as that guy and he had ties with two of the people i

Youth: d like that

Youth: im literally so f****** dumb

Youth: and he'll send me something with the logo of art bullshit memes in it and it literally reminds me of him

Counselor: Thank you for reaching out to [PERSON_NAME] for support and for sharing all this with me. I'm so glad you did reach out today. I'm here to listen and talk through anything you have going on. What name can I call you while we chat?

Youth: jake is fine

Counselor: Nice to meet you, [PERSON_NAME]. How old are you?
Counselor: Hi welcome to TrevorChat. My name is Alex. What’s going on?

Youth: idk I guess i’ve just been feeling depressed lately. I’m feeling really down and anxious about thing

Counselor: Thank you for sharing that with me. It sounds like you are having a difficult time. Before we start, what name can I call you today?

Youth: riley

Counselor: Thanks Riley. Nice to meet you. How old are you?

Youth: 16

Counselor: Thank you for letting me know. What city and state are you chatting from?

Youth: asheville, NC
Key Learnings from CCS
What makes projects like CCS successful?
Learning #1: Let the AI/ML Project Lifecycle guide but not define

Ideally....
Let the AI/ML Project Lifecycle guide but not define Realistically…

Step 1 Acquire data over time
Step 2 Create initial prototype
Step 3 Define project objectives
Step 4 Iterate and improve prototype
Step 5 Interpret and communicate results
Step 6 Iterate further while building software
Step 7 Align on go-ahead for model
Step 8 Launch product

Lather, Rinse, and Repeat
What can AI/ML do?

- Automate or streamline human processes that are labor-intensive
- Discover trends and patterns
- Extract meaning from language

What benefits does AI/ML have?

- Facilitates the scaling / expansion of operations
- Simplifies complex operational workflows
- Standalone processes provide flexible time schedules
- Automatic operations have quick response times

Learning #2: Involve cross-functional discussions early
However, thinking more concretely about AI/ML can get complex...

Concrete tasks that AI/ML can do with text data

- Named Entity Recognition (NER)
- Dependency Parsing
- Classification + Regression
- Semantic Embedding
- Linguistic Diversity
- Chatbot
- Topic Clustering
- Part-of-speech tagging

...and more!
In addition, AI/ML projects require a broad range of expertise.

Thus, early discussions amongst cross-functional team members are necessary.
Learning #3: Monitor your product, obtain feedback, and *act accordingly*

**Logging**: Current CCS logging system records inputs, outputs, and metadata for CCS

**User Feedback**: Trainees complete a post-chat survey after every CCS role-play

**User Research**: UX Research interprets survey results and conducts interviews

**Improvement Scoping**: Product analyzes information and determines path forward
Example: An early version of CCS occasionally responded with repeated phrases.

Logging: Logging system recorded cases in which repetitions were likely to occur

User Feedback: Post-chat survey identified trainees that were impacted

User Research: UX Research further assessed the impact

Improvement Scoping: Product worked with AI/ML experts to identify a fix and determined go/no-go
“We believe in technology enabling our work, but it does not replace our work. That person-to-person connection for people in crisis is critical. It’s the core of what we do. The way that we are using technology is to help facilitate that.”
<table>
<thead>
<tr>
<th>Area</th>
<th>Considerations</th>
<th>Avoid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Datasets &amp; Training</td>
<td>✓ Consider the size and source of your data</td>
<td>✓ Using data that is not understood or analyzed</td>
</tr>
<tr>
<td></td>
<td>✓ Ensure privacy through removing Personally Identifiable Information (PII)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓ Identify what biases, skews or mistakes are built-in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓ Determine how the data can be improved</td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>✓ Define intersectional groups that could be at a disadvantage</td>
<td>✓ Evaluating based on a single dimension, if possible</td>
</tr>
<tr>
<td></td>
<td>✓ Define intersectional groups that may have an advantage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓ Examine your success criteria/metrics among these groups</td>
<td></td>
</tr>
<tr>
<td>Feedback &amp; Collaboration</td>
<td>✓ Bring a diverse group of people together to develop and review model</td>
<td>✓ Developing or making decisions without domain experts</td>
</tr>
<tr>
<td></td>
<td>✓ Implement a feedback loop</td>
<td></td>
</tr>
<tr>
<td>Evolution</td>
<td>✓ Adjust as you learn more about the people you serve, technology, and other approaches</td>
<td>✓ Remaining static in the approach</td>
</tr>
</tbody>
</table>
What’s next?
The future of AI @ The Trevor Project
Strategic Vision

+ Significantly grow the number of crisis contacts served
+ Build on The Trevor Project’s history of quality and inclusivity to be the most effective and welcoming support for every LGBTQ young person
+ Lay the foundation for continued growth to support LGBTQ youth across the U.S. and abroad
Feel free to connect!

Taki Hasegawa  
(she/they)  
- Senior Machine Learning Engineer  
- Email: taki.hasegawa@thetrevorproject.org  
- LinkedIn: https://www.linkedin.com/in/taki-hasegawa
Feel free to connect!
Our trained counselors served more than 14,000 contacts via TrevorLifeline, TrevorChat, and TrevorText in 2020.