

I am an **Interactive Machine Translation** researcher at Lilt with expertise in **Data Science**, **Machine Learning**, **Natural Language Processing**, and **Human-Computer Interaction**. I have a Ph.D in Computer Science from Stanford University, and have been awarded the NSF and NDSEG Fellowships, and a Stanford Human-Centered AI Grant. I have written software used by millions of users – I built HabitLab, a data-driven behavior change system with **12,000+ daily active users**; I also wrote and maintain UNetbootin, an open-source project with **40 million downloads**.

## Education

<b>Ph.D.</b> , Computer Science	Stanford University	GPA: 4.0/4.0	2019
<b>B.S.</b> and <b>M.Eng.</b> , Computer Science	Massachusetts Institute of Technology	GPA: 5.0/5.0	2013

## Work Experience

<b>Principal Research Scientist</b>	Lilt	San Francisco	Feb 2021 - present
<b>Senior Research Scientist</b>	Lilt	San Francisco	Aug 2019 - Feb 2021

I am the head of HCI research and manage a team of researchers at Lilt, an interactive machine translation startup. Improved interactive MT system speed by shifting computation client-side via TensorflowJS and heuristics. Built named entity transliteration system based on Transformer architecture using tensor2tensor and Tensorflow. Developed metrics and logging to determine how translators spend their time and predict translator performance. Ran A/B tests to evaluate website translation ROI, and developed system that recommends pages to translate.

<b>Research Intern</b>	Microsoft Research	Redmond	Summer 2015
------------------------	--------------------	---------	-------------

Designed and built an educational social feed app usable by pre-literate children. Published at CSCW 2017.

<b>Research Intern</b>	Microsoft Research	Beijing	Summer 2014
------------------------	--------------------	---------	-------------

Built QuizCram, a quiz-driven MOOC lecture viewer that improves learning outcomes. Presented at CHI 2015.

<b>Software Engineering Intern</b>	Google	Mountain View	Summer 2013
------------------------------------	--------	---------------	-------------

Developed a machine learning system for detecting taps on the phone bezel, for use in Android input methods.

<b>Software Engineering Intern</b>	Google	Mountain View	Summer 2012
------------------------------------	--------	---------------	-------------

Developed an NLP system to automatically generate glossaries from book text. Patent granted **US9483460B2**.

<b>Software Engineering Intern</b>	Google	Mountain View	Summer 2011
------------------------------------	--------	---------------	-------------

Developed a machine learning system to predict the quality of user reviews, now deployed on Google Play.

<b>Software Development Engineer Intern</b>	Microsoft	Redmond	Summer 2010
---	-----------	---------	-------------

Built the IntelliSense code completion system for a scientific computing language, and contributed to its compiler.

## Skills and Technologies

**Programming Languages:** Python, JavaScript, C, C++, Java, TypeScript, R, C#, Ruby, Scala, Haskell, Bash, SQL  
**Machine Learning + Deep Learning:** Tensorflow, PyTorch, TensorflowJS, Keras, scikit-learn, xgboost, MLFlow  
**Natural Language Processing + Machine Translation:** SpaCy, tensor2tensor, fairseq, HuggingFace, LASER, NLTK  
**Data Science + Visualization:** NumPy, SciPy, Pandas, Jupyter, RStudio, Plotly, D3.js, Superset, Hadoop, MapReduce  
**Web + Mobile Development:** HTML, CSS, React, AngularJS, NodeJS, Express, Flask, MySQL, Docker, Android  
**Languages:** Fluent: English, Chinese (Mandarin), Hungarian. Intermediate: Japanese, Vietnamese, Spanish.

## Publications in Journals and Conference Proceedings

**Geza Kovacs**, Zhengxuan Wu, Michael Bernstein. "Not Now, Ask Later: Users Weaken Their Behavior Change Regimen Over Time, But Expect To Re-Strengthen It Imminently." *ACM annual conference on Human Factors in Computing Systems (CHI) 2021*.

Samuel Läubli, Patrick Simianer, Joern Wuebker, **Geza Kovacs**, Rico Sennrich, Spence Green. "The Impact of Text Presentation on Translator Performance." *Target: International Journal of Translation Studies, 2021 (to appear)*.

**Geza Kovacs**, Drew Mylander Gregory, Zilin Ma, Zhengxuan Wu, Golrokh Emami, Jacob Ray, Michael Bernstein. "Conservation of Procrastination: Do Productivity Interventions Save Time Or Just Redistribute It?" *ACM annual conference on Human Factors in Computing Systems (CHI) 2019*.

**Geza Kovacs**, Zhengxuan Wu, Michael Bernstein. "Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition." *ACM annual conference on Computer-Supported Cooperative Work and Social Computing (CSCW) 2018*.

Rajan Vaish, Neil Gaikwad, **Geza Kovacs**, Andreas Veit, Ranjay Krishna, Imanol Arrieta Ibarra, Camelia Simoiu, Michael Wilber, Serge Belongie, Sharad Goel, James Davis, Michael Bernstein. "Crowd Research: Open and Scalable University Laboratories." *ACM Symposium on User Interface Software and Technology (UIST) 2017*.

Kiley Sobel, **Geza Kovacs**, Galen McQuillen, Andrew Cross, Nirupama Chandrasekaran, Nathalie Riche, Ed Cutrell, Meredith Morris. "EduFeed: A Social Feed to Engage Preliterate Children in Educational Activities." *ACM annual conference on Computer-Supported Cooperative Work and Social Computing (CSCW) 2017*.

**Geza Kovacs**. "Effects of In-Video Quizzes on MOOC Lecture Viewing." *ACM annual conference on Learning at Scale (L@S) 2016*.

**Geza Kovacs** and Robert C. Miller. "Smart Subtitles for Vocabulary Learning." *ACM annual conference on Human Factors in Computing Systems (CHI) 2014*.

## Publications in Adjunct Conference Proceedings

Stanford Crowd Research, **Geza Kovacs**, Rajan Vaish, Michael Bernstein. "Daemo: A Self-Governed Crowdsourcing Marketplace". *ACM Symposium on User Interface Software and Technology (UIST) 2015, Poster*.

**Geza Kovacs**. "FeedLearn: Using Facebook Feeds for Microlearning." *ACM annual conference on Human Factors in Computing Systems (CHI) 2015, Extended Abstracts*.

**Geza Kovacs**. "QuizCram: A Question-Driven Video Studying Interface." *ACM annual conference on Human Factors in Computing Systems (CHI) 2015, Extended Abstracts*.

Joseph Jay Williams, **Geza Kovacs**, Caren Walker, Samuel G Maldonado, Tania Lombrozo. "Learning Online via Prompts to Explain." *ACM annual conference on Human Factors in Computing Systems (CHI) 2014, Extended Abstracts*.

**Geza Kovacs** and Robert C. Miller. "Foreign Manga Reader: Learn Grammar and Pronunciation while Reading Comics." *ACM Symposium on User Interface Software and Technology (UIST) 2013, Demo*.

**Geza Kovacs**. "Smart Subtitles for Language Learning." *ACM annual conference on Human Factors in Computing Systems (CHI) 2013, Extended Abstracts*.

**Geza Kovacs**. "ScreenMatch: providing context to software translators by displaying screenshots." *ACM annual conference on Human Factors in Computing Systems (CHI) 2012, Extended Abstracts*.

## Patents

Tania Bedrax-Weiss, **Geza Kovacs**, Ulas Kirazci. "Automated formation of specialized dictionaries." [US9483460B2](#). Filed 10/2013, Published 11/2016, Expires 01/2034.

Meredith Morris, Nathalie Henry Riche, Edward B. Cutrell, Andrew C. Cross, Natasa Milic, Nirupama Chandrasekaran, Galen McQuillen, Kiley Sobel, **Geza Kovacs**. "Presenting educational activities via an extended social media feed." [US20180068578A1](#). Filed 09/2016, Published 03/2018.

## Open Source Projects

**UNetbootin** (LiveUSB Creator) <https://en.wikipedia.org/wiki/UNetbootin>  
**40 million downloads**. UNetbootin creates bootable USB flash drives for various (50+) Linux distributions.

**Wubi** (Ubuntu Installer for Windows) [https://en.wikipedia.org/wiki/Wubi\\_\(software\)](https://en.wikipedia.org/wiki/Wubi_(software))  
**Now part of Ubuntu**. Built the first versions of Wubi, which allows Ubuntu to be installed from Windows.

**HabitLab** (In-the-wild Behavior Change Research Platform) <https://habitlab.stanford.edu>  
**12,000+ daily active users**. I built HabitLab over my Ph.D, and it is still used for research at Stanford Medical School.

## Invited Keynote Talks

**Geza Kovacs.** "Predictive Translation Memory in the Wild: A Study of Interactive Machine Translation Use on Lilt." *Association for Machine Translation in the Americas (AMTA) Workshop on the Impact of Machine Translation 2020.*

## Select Awards and Honors

Stanford Human-Centered AI Grant (for my research project HabitLab)	2018
National Defense Science and Engineering Graduate Fellowship	2013
National Science Foundation Graduate Research Fellowship	2013
1 <sup>st</sup> place, ACM UIST (User Interface Software and Technology) Student Innovation Contest	2012
1 <sup>st</sup> place, ACM CHI (Human Factors in Computing Systems) Student Research Competition	2012
Phi Beta Kappa (top 10% of students at MIT), Tau Beta Pi (top 12.5% of Engineering students at MIT)	2012

## Academic Conference Reviewing and Committees

Organizing Committee, WMT 2022 Shared Task on Word-Level Auto-Completion	2021
Program Committee, EACL 2021 Bridging HCI and NLP Workshop	2021
Reviewer, ACM Conference on Human Factors in Computing Systems (CHI)	2015, 2018, 2019, 2021, 2022
Reviewer, ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)	2021
Reviewer, ACM Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)	2019
Reviewer, ACM Symposium on User Interface Software and Technology (UIST)	2017, 2018

## Researchers Managed

Sai Gouravajhala, Senior Research Scientist at Lilt.	August 2020 – present
Hannah Yan, Senior Data Scientist at Lilt.	September 2020 – present
Jordan Huffaker, Research Intern at Lilt. Now a Ph.D student at University of Michigan.	Summer 2021
Jessy Lin, Research Engineer at Lilt. Now a Ph.D student at UC Berkeley.	August 2019 – August 2020
Ming-Chang Chiu, Data Science Intern at Lilt. Now a Ph.D student at USC.	Summer 2020
Zhengxuan Wu, MS student researcher at Stanford.	Fall 2018 – Spring 2019
Zilin Ma, undergraduate researcher at Stanford. Now a Ph.D student at Harvard.	Summer 2018

## Teaching Experience

<b>Understanding Users</b> (CS 377U) – Teaching Assistant, at Stanford	Spring 2019
<b>Human Computer Interaction Research</b> (CS 376) – Teaching Assistant, at Stanford	Fall 2018
<b>Natural Language Processing</b> (6.863) – Teaching Assistant, at MIT	Fall 2012

## Select Press Coverage

### HabitLab

<b>WIRED</b> - The HabitLab Browser Extension Curbs Your Time Wasted on the Web.	January 2019
<b>Lifehacker</b> - Prevent Procrastination With This Chrome Extension.	February 2019
<b>The New York Times</b> - Finding It Hard to Focus? Maybe It's Not Your Fault.	August 2018
<b>Lifehacker</b> - Be More Mindful of the Time You Waste Online With HabitLab.	August 2018
<b>Entrepreneur</b> - Use These Strategies to Maximize Productivity Without Inventing an Extra Weekday.	May 2018

### Crowd Research / Daemo

<b>Stanford University News</b> - A Stanford-led platform for crowdsourced research	October 2017
<b>WIRED</b> - Amazon's Turker Crowd Has Had Enough.	August 2017

### UNetbootin

<b>Forbes</b> - How To Try Linux Without Making Any Changes To Your PC.	September 2018
<b>PCWorld</b> - Create a Bootable Linux Flash Drive in Three Easy Steps.	February 2012
<b>Lifehacker</b> - The Complete Guide to Saving Your Windows System with a Thumb Drive.	March 2010
<b>Network World</b> - Installing Ubuntu on an old netbook with hair tearing and profanity.	August 2014

### Wubi

<b>Ars Technica</b> - Wubi arrives: a look at Ubuntu 8.04 alpha 5.	February 2008
<b>Lifehacker</b> - Install Ubuntu on a Windows Netbook, No Partitioning Needed.	May 2008