My research aims to design **equitable natural language processing (NLP) systems that are adaptable to individuals and groups**. Despite the ubiquity and popularity of large language models (LLMs), numerous studies have shown that their benefits are not equal across everyone. This is partly because today's LLMs are primarily optimized to generate the most average responses for the most average person based on training data that is insufficient to learn nuanced, inclusive responses. For example, when responding to questions like 'What precautions should I take when visiting Saudi Arabia?', language models should consider the questioner's cultural background, age, gender, sexuality, and personal preferences, all factors that influence the kinds of situations they may encounter.

I address these challenges by **centering the sociocultural context in modern NLP** and developing computational approaches that **identify fairness issues** that arise from failures to incorporate social factors. Specifically, I address the following key areas at the intersection of AI ehhics, NLP, and computational social science (CSS):

- **Measuring social biases in sociotechnical system (AI Ethics)**. How do we identify and robustly measure social biases in language technologies, at *all* stages of development? [1, 2, 3]
- **Incorporating social contexts into NLP models (NLP)**. How do we develop socially aware NLP models that incorporate semantic and pragmatic sociocultural knowledge? [4, 5, 6, 7, 8, 9]
- **Explaining social phenomena with NLP (CSS)**. How can we use socioculturally aware NLP models to better understand people and cultures through the lens of language? [10, 11, 12]

In addressing these questions, I develop general and theoretically grounded computational methods, model architectures, analysis algorithms, and datasets. I enhance the **performance and equity of NLP systems by directly incorporating social context into their learning procedures**, solving challenges that cannot be resolved solely by scaling up models indefinitely. Using advanced machine learning (ML) and NLP techniques, my work spans beyond an English-centric view of NLP to **a variety of languages with cross-lingual methods** and multilingual models, encompassing structured sources like news articles and Wikipedia to the unstructured landscape of social media. **Bridging social science with NLP systems**, my research not only develops state-of-the-art technologies to equitably serve diverse users but also sheds new light on our understanding of people and cultures.

Robustly Measuring Social Biases in Language Technologies

It is well established that NLP models learn and amplify social biases [13, 14]. While considerable work addresses social biases in language technologies, it generally focuses on a limited set of biases (e.g., gender or racial bias) [15] within limited scenarios [16] in a single language (primarily, English) [17]. This makes findings less generalizable and less robust [18]. I aim to develop novel computational methods to identify and measure more diverse social biases comprehensively.

Controlled Multilingual Affect Analyses for Measuring Social Biases As a step towards more robust algorithms to analyze social biases, I introduced a series of methods that quantify various social biases, including those towards LGBT, non-binary, and intersectional identity groups in Wikipedia biographies across multiple languages [2, 3]. My collaborators and I proposed **biography matching algorithms grounded in causal inference methods** to control for confounding factors [3] and a **model for multilingual affective analysis** [2] that leverages crosslingual contextual sentence embeddings to measure implied affect towards a person. We built the **first multilingual dataset for the contextual affective analysis task** to train the model. One of our analyses revealed that Russian articles tend to use verbs with more negative connotations when describing LGBT people than English or Spanish articles, confirming different perceptions across cultures.

The Wikimedia Foundation, the primary stakeholder with power to use our research to mitigate social biases in people's narratives, recognized our work with the **Wikimedia Foundation Research Award of the Year** and is implementing our methods.[1] Our work has also received interest from the **journalism community**, leading to a collaboration with *The Washington Post* to examine anti-Black discrimination on Chinese social media [12].

Tracking the Trails of Social Biases Leading to Unfair NLP Models Building upon our investigation into biases present in Wikipedia, a critical extension of this work is to understand if and how the biases get amplified throughout the development lifecycle of LLMs [18]. Recently, we studied how political

[1] https://phabricator.wikimedia.org/T290447

biases in training data propagates to LLMs and eventually leads to unfair NLP models [1]. We devised a framework, grounded in established *political compass tests* from political science to **measure political leanings of LLMs**. Using a stance detector, we assessed models' positions on various topics. We then created controlled training datasets to demonstrate that political biases in training data propagate to LLMs during pretraining. Subsequently, we finetuned the models with varied political beliefs on downstream task datasets to reveal that **LLMs' distinct political leanings affect the types of content they classify** as hate speech and misinformation. For example, left-leaning models showed increased sensitivity to hate speech against minority groups, such as Women and LGBTQ+ individuals, while right-leaning models exhibited greater sensitivity to hate speech targeting white Christian men. Our work won the **Best Paper Award at ACL 2023**, gaining extensive media attention, including coverage and interviews by *The Washington Post* [19] and the *MIT Tech Review* [20].

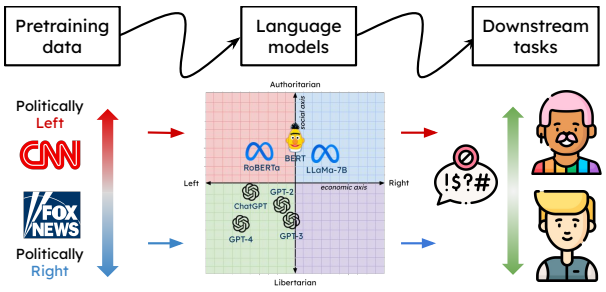

Figure 1: We measure political biases of LLMs and systematically examine how they lead to unfair NLP models for downstream tasks.

Future Work My research identifies open challenges in human-centered NLP, particularly in robustly measuring **cultural biases in NLP systems**. I am excited to investigate how well LLMs adapt to diverse cultures and how to quantify their understanding of each culture to enhance model fairness. For example, I am currently working on building a challenging benchmark dataset to test LLMs' reasoning ability for both explicit cultural knowledge and implicit cultural norms. Additionally, I aspire to continue developing methodologies for multilingual controlled analysis, contributing to uncovering and addressing cultural biases in platforms like Wikipedia, thereby promoting equitable representations across languages.

Incorporating Social Contexts to Improve NLP Models

While identifying social bias in NLP systems is essential to make them more equitable, another key challenge lies in integrating a nuanced understanding of social contexts into technology development. To build socially aware and adaptable NLP systems, models must effectively represent and account for social contexts [14]. However, as much social information is not explicitly stated in language, social context is absent in most training datasets and current models [21]. My work addresses this challenge by (1) building **datasets that encode social contexts** along with the texts, including information about writers and readers and the social settings in which the texts are contextualized, and (2) developing **new ML models that integrate** such information.

Community Context for Norm Violation Detection Today's automated tools for moderating online communities (e.g., hate speech detectors) do not take social context into account [22]. I hypothesized that incorporating explicit knowledge about a community and its rules is crucial for detecting violations of community norms more accurately. To validate this hypothesis, we collected NormVio [4], a dataset that contains 52K comments from Reddit, their communities (i.e., subreddits), their respective community rules, prior conversation information, and labels indicating whether they violated any community rules and were moderated by human moderators. We then introduced BERT-based **context-sensitive norm violation classifiers** which are capable of incorporating community context and rules as additional input; unlike existing hate speech classifiers that rely solely on text input, they consider community-specific information. Our best model outperformed context-insensitive baselines in detecting norm violations by nearly 50% in terms of F1, and our models can pinpoint specific violated rules in a community. Context-sensitive classifiers thus provide a key **practical assistive technology**, helping human moderators identify inappropriate content for their specific communities and communicate their rationale to users. This lessens the burden of managing the overwhelming influx of content. This work **led to a startup's interest** in developing a similar model for other platforms that suffer from intractable amounts of toxic comments, e.g., real-time chat platforms like Twitch, resulting in a collaboration paper accepted at EMNLP 2023 [5].

Generative Zero-Shot Classifier with Text Labels for Personalization Since a limited number of datasets offer social context, zero-shot classifiers that can account for social context without training data can be

especially valuable. We introduced a *generative classifier* for zero-shot classification [9] that enables the simple personalization and adaptation of models by incorporating social context through a text label (Figure 2). For example, a comment "go get it girl" might be empowering when addressed to a woman but sarcasm when addressed to a man. Our model calculates the probability of generating the comment, given the contextual text label, such as "The comment written by a woman empowers the addressed woman", to determine whether the comment is empowering. Our model, evaluated across 18 different tasks including hate speech and empowerment prediction [6], shows a better classification performance than strong in-context learning baselines. For this line of work – making LLMs more socioculturally aware – I received CMU's competitive **K&L Gates Presidential Fellowship for Ethics and Computational Technologies**.

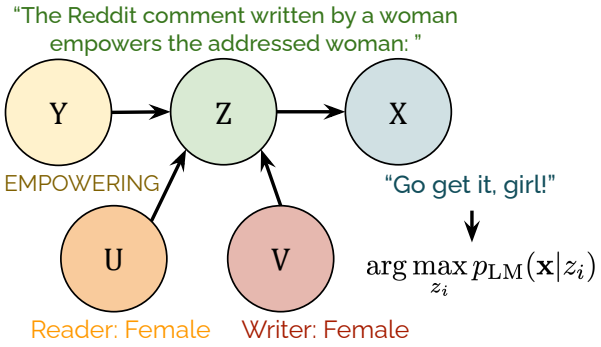

Figure 2: Our generative framework measures the likelihood of LLMs generating input text $\mathbf{x}$, conditioned on natural language descriptions of labels z to incorporate social contexts.

Future Work. My work has advanced the development of socioculturally aware NLP models to make them more effective and equitable, but much work remains to be done. In addition to classification, I am particularly interested in designing LLMs that can *generate* responses that are more appropriate and useful by considering users' sociocultural backgrounds. I am currently working on contextualizing reinforcement learning from human feedback (RLHF), known to be a critical step in aligning models with human preferences, with social contexts. Specifically, we constructed a set of paired Reddit comments with human preference labels indicated by users' upvotes across diverse subreddits. We are testing whether contextualizing models during RLHF helps the model generate responses that are better suited to each subreddit.

Leveraging NLP Models to Explain Social Phenomena

People communicate through language, and social scientists analyze it to understand society. Despite remarkable advances in NLP, not all systems have been adopted by social scientists because many language technologies are not developed with real-world applications in mind [23]; as such, they might not perform adequately in new target domains [24] and might overlook requirements essential for social scientists, such as interpretability [25]. My research aims to **identify shortcomings** in NLP models for CSS [11], propose **NLP solutions** to bridge the gap [10], and demonstrate how **state-of-the-art NLP** methods can uncover new societal insights [10, 11].

Analyzing the Driving Forces of Activism using Robust Emotion Classifiers In my 2022 **PNAS** paper [10], we showed how to analyze the relationship between social movements and emotions expressed on social media with domain adaptation. Specifically, we compared various unsupervised and few-shot training methods for domain adaptation to provide **guidelines for social scientists** who want to use NLP models on their own target domains and tasks. With our domain-adapted emotion classifiers, we collected and analyzed 34M tweets posted during the 2020 Black Lives Matter (BLM) protests in collaboration with a researcher from the Data for Black Lives organization. We found that expressions of positivity, like hope and camaraderie, are more significantly linked to the movement than negative ones, like anger (see Figure 3), thereby **countering a common stereotype of "angry Black people."**

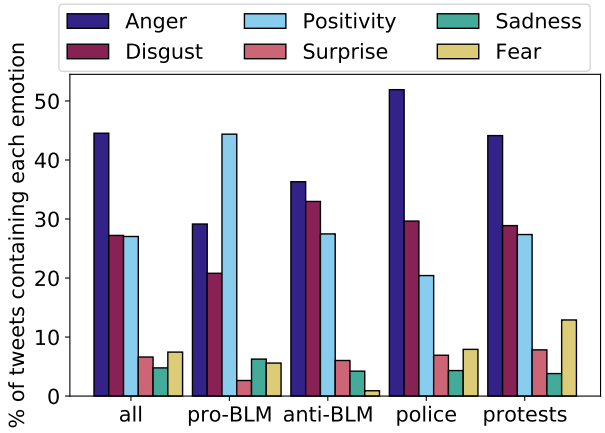

Figure 3: Emotion distribution of tweets by hashtags reveals that positivity is the main emotion pro-BLM tweets exhibit in contrast to other hashtags.

Challenges of NLP models in Detecting Information Manipulation In another work [11], we analyzed

challenges and opportunities of NLP models used to detect information manipulation by examining Russian media during the 2022 Ukraine-Russian war. We **collected a dataset** of 10M+ Russian social media posts by state-affiliated and independent media outlets along with public reactions to them (e.g., likes and comments), and we applied state-of-the-art NLP models, such as topic modeling and media framing classifiers. While we identified numerous opportunities for NLP research to make positive contributions in combating real-world information manipulation campaigns (e.g., **uncovering agenda setting and framing strategies** of stat-affiliated media), we also recognized challenges in developing more deployable technology in practice. This project generated a high demand for our dataset (**20+ requests**), highlighting its value in advancing research on real-world misinformation and information campaigns. One data request resulted in a collaboration and **DoD funding** ($160K) on using NLP systems to identify evolving narratives in information operations.

Future Work One of the important factors that hinders the wide adoption of NLP methods is their interface [26]; for example, social scientists need to figure out how to finetune and run the models. However, with the introduction of generative LLMs such as ChatGPT, NLP systems now possess an exciting potential to offer powerful reasoning ability in convenient, easily learnable ways [27]. I plan to continue investigating and demonstrating various applications of LLMs to address research questions in diverse social science fields, including *political science*, *linguistics*, *communication*, and *cultural studies*. As one example, collaborators and I are investigating the utility of LLMs in identifying social norms that govern language in online communities by computationally modeling community's recognition signals (e.g., the number of upvotes on Reddit).

Future Directions

Aligned with my long-term goal of building socioculturally aware NLP systems to make them equitable and accessible, I plan to expand my research in various **interdisciplinary directions**, including the following.

AI Safety and Public Policy My research has shown how social biases in NLP models can disproportionately impact users [2, 9]. My primary goal is to ensure AI model safety and prevent harm to users. I aim to develop rigorous evaluation methods and benchmarks that can get widely adopted, both in the research community and industry. The keys to broad adoption are a 1) comprehensive benchmark framework applicable to various domains and problems, and 2) an active evaluation framework that stays effective by updating methods and data to adapt to model and language changes. Finally, to ensure that AI safety research translates into real-world impact, I will collaborate with researchers and practitioners in **public policy** to explore effective implementations of measures developed within the NLP community to guide one of the most powerful yet opaque technologies ever created.

Socially and Culturally Aware Multilingual Models Numerous studies have highlighted performance discrepancies and social biases exhibited by LLMs across languages [28, 29]. I intend to develop computational approaches to mitigate these performance gaps within multilingual models. One promising avenue of exploration involves implementing a *teacher-student model framework* [30] *between resource-rich and low-resource languages*, with the help of translation models. However, one foreseeable challenge in pursuing this direction is in determining what teachers can teach to students [31, 32]. For example, the answer to a question like "What is 1+1?" may be universally transferable to all languages, whereas answers to questions like "How much should I tip at a restaurant?" can vary significantly based on language and culture. To address this, I will leverage my computational expertise to collaborate with experts in **linguistics, social psychology, and anthropology** in developing more equitable and socially aware multilingual models.

Personalization and User Privacy To deploy models that incorporate users' sociocultural context in the real-world, an understanding of users' privacy requirements is indispensable [33]. Similar to personalized ads, to make this technology inviting, users should be able to control what information models can access and know why models make certain decisions [34]. In future work, I will focus on ways to build models that are **controllable** and **interpretable** by users, making them socially aware without being intrusive. Furthermore, I plan to collaborate with **Human Computer Interaction** experts to investigate how much personalization is appropriate for users and how it should be implemented in applications.

I am excited to take further steps to build equitable, inclusive, ethical, and trustworthy NLP systems as a faculty!

References

[1] Shangbin Feng, **Chan Young Park**, Yuhan Liu, and Yulia Tsvetkov. From pretraining data to language models to downstream tasks: Tracking the trails of political biases leading to unfair NLP models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11737–11762, Toronto, Canada, July 2023. Association for Computational Linguistics. URL https://aclanthology.org/2023.acl-long.656.

[2] **Chan Young Park***, Xinru Yan*, Anjalie Field*, and Yulia Tsvetkov. Multilingual contextual affective analysis of LGBT people portrayals in wikipedia. In *Proc. ICWSM'21*, 2021.

[3] Anjalie Field, **Chan Young Park**, Kevin Z. Lin, and Yulia Tsvetkov. Controlled analyses of social biases in Wikipedia bios. In *Proc. The ACM Web Conference '22*, 2022.

[4] **Chan Young Park**, Julia Mendelsohn, Karthik Radhakrishnan, Kinjal Jain, Tushar Kanakagiri, David Jurgens, and Yulia Tsvetkov. Detecting community norm violations in online conversations. In *Proc. EMNLP Findings'21*, 2021.

[5] Jihyung Moon, Dong-Ho Lee, Hyundong Cho, Woojeong Jin, **Chan Young Park**, Minwoo Kim, Jonathan May, Jay Pujara, and Sungjoon Park. Analyzing norm violations in live-stream chat. *arXiv preprint arXiv:2305.10731*, 2023.

[6] **Chan Young Park***, Lucille Njoo*, Octavia Stappart, Marvin Thielk, Yi Chu, and Yulia Tsvetkov. Talkup: A novel dataset paving the way for understanding empowering language. *arXiv preprint arXiv:2305.14326*, 2023.

[7] **Chan Young Park***, Jimin Sun*, Hwijeen Ahn*, Yulia Tsvetkov, and David R Mortensen. Ranking transfer languages with pragmatically-motivated features for multilingual sentiment analysis. In *Proc. EACL'21*, 2021.

[8] **Chan Young Park***, Hwijeen Ahn*, Jimin Sun*, and Jungyun Seo. NLPDove at SemEval-2020 task 12: Improving offensive language detection with cross-lingual transfer. In *Proc. SemEval-2020*, 2020.

[9] Sachin Kumar, **Chan Young Park**, and Yulia Tsvetkov. Gen-z: Generative zero-shot text classification with contextualized label descriptions. *arXiv preprint arXiv:2311.07115*, 2023.

[10] **Chan Young Park***, Anjalie Field*, Antonio Theophilo*, and Yulia Tsvetkov. An analysis of emotions and the prominence of positivity in #blacklivesmatter tweets. *National Academy of Sciences (PNAS)*, 2022.

[11] **Chan Young Park***, Julia Mendelsohn*, Anjalie Field*, and Yulia Tsvetkov. Challenges and opportunities in information manipulation detection: An examination of wartime Russian media. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 5209–5235, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. URL https://aclanthology.org/2022.findings-emnlp.382.

[12] Sarah Cahlan and Joyce Sohyun Lee. Video evidence of anti-Black discrimination in china over coronavirus fears. *The Washington Post*. URL https://www.washingtonpost.com/politics/2020/06/18/video-evidence-anti-black-discrimination-china-over-coronavirus-fears/.

[13] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. Language (technology) is power: A critical survey of "bias" in NLP. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5476, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.485. URL https://aclanthology.org/2020.acl-main.485.

[14] Dirk Hovy and Shrimai Prabhumoye. Five sources of bias in natural language processing. *Language and Linguistics Compass*, 15(8):e12432, 2021.

[15] Anjalie Field, Su Lin Blodgett, Zeerak Waseem, and Yulia Tsvetkov. A survey of race, racism, and anti-racism in NLP. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1905–1925, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.acl-long.149. URL https://aclanthology.org/2021.acl-long.149.

[16] Amandalynne Paullada, Inioluwa Deborah Raji, Emily M Bender, Emily Denton, and Alex Hanna. Data and its (dis) contents: A survey of dataset development and use in machine learning research. *Patterns*, 2 (11), 2021.

[17] Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. The state and fate of linguistic diversity and inclusion in the NLP world. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.560. URL https://aclanthology.org/2020.acl-main.560.

[18] Seraphina Goldfarb-Tarrant, Rebecca Marchant, Ricardo Muñoz Sánchez, Mugdha Pandya, and Adam Lopez. Intrinsic bias metrics do not correlate with application bias. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1926–1940, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.acl-long.150. URL https://aclanthology.org/2021.acl-long.150.

[19] Gerrit De Vynck. Chatgpt leans liberal, research shows. *The Washington Post*. URL https://www.washingtonpost.com/technology/2023/08/16/chatgpt-ai-political-bias-research/.

[20] Melissa Heikkilä. Ai language models are rife with different political biases. *MIT Technology Review*. URL https://www.technologyreview.com/2023/08/07/1077324/ai-language-models-are-rife-with-political-biases/?truid=&utm_source=the_algorithm&utm_medium=email&utm_campaign=the_algorithm.unpaid.engagement&utm_content=08-07-2023.

[21] Dirk Hovy and Diyi Yang. The importance of modeling social factors of language: Theory and practice. In Kristina Toutanova, Anna Rumshisky, Luke Zettlemoyer, Dilek Hakkani-Tur, Iz Beltagy, Steven Bethard, Ryan Cotterell, Tanmoy Chakraborty, and Yichao Zhou, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 588–602, Online, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.naacl-main.49. URL https://aclanthology.org/2021.naacl-main.49.

[22] Edoardo Mosca, Maximilian Wich, and Georg Groh. Understanding and interpreting the impact of user context in hate speech detection. In Lun-Wei Ku and Cheng-Te Li, editors, *Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media*, pages 91–102, Online, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.socialnlp-1.8. URL https://aclanthology.org/2021.socialnlp-1.8.

[23] Ana Macanovic. Text mining for social science–the state and the future of computational text analysis in sociology. *Social Science Research*, 108:102784, 2022.

[24] Lisa Kühnel and Juliane Fluck. We are not ready yet: limitations of state-of-the-art disease named entity recognizers. *Journal of Biomedical Semantics*, 13(1):26, 2022. doi: 10.1186/s13326-022-00280-6. URL https://doi.org/10.1186/s13326-022-00280-6.

[25] Tamás Rudas and Gábor Péli. *Pathways between social science and computational social science: theories, methods, and interpretations*. Springer, 2021.

[26] Simon Wibberley, David Weir, and Jeremy Reffin. Language technology for agile social media science. In Piroska Lendvai and Kalliopi Zervanou, editors, *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 36–42, Sofia, Bulgaria, August 2013. Association for Computational Linguistics. URL https://aclanthology.org/W13-2705.

[27] Caleb Ziems, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang. Can large language models transform computational social science? *Computational linguistics*, 49(4), December 2023.

[28] Jaeseong Lee, Dohyeon Lee, and Seung-won Hwang. Script, language, and labels: overcoming three discrepancies for low-resource language specialization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 13004–13013, 2023.

[29] Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. Unsupervised cross-lingual

representation learning at scale. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.747. URL https://aclanthology.org/2020.acl-main.747.

[30] Yun Chen, Yang Liu, Yong Cheng, and Victor O.K. Li. A teacher-student framework for zero-resource neural machine translation. In Regina Barzilay and Min-Yen Kan, editors, *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1925–1935, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1176. URL https://aclanthology.org/P17-1176.

[31] Mohammad Fazleh Elahi and Paola Monachesi. An examination of cross-cultural similarities and differences from social media data with respect to language use. In *LREC*, pages 4080–4086, 2012.

[32] Bill Yuchen Lin, Frank F. Xu, Kenny Zhu, and Seung-won Hwang. Mining cross-cultural differences and similarities in social media. In Iryna Gurevych and Yusuke Miyao, editors, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 709–719, Melbourne, Australia, July 2018. Association for Computational Linguistics. doi: 10.18653/v1/P18-1066. URL https://aclanthology.org/P18-1066.

[33] Eran Toch, Yang Wang, and Lorrie Faith Cranor. Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. *User Modeling and User-Adapted Interaction*, 22:203–220, 2012.

[34] S Shyam Sundar and Sampada S Marathe. Personalization versus customization: The importance of agency, privacy, and power usage. *Human communication research*, 36(3):298–322, 2010.