Challenges in Recruiting Andean Indigenous Participants for Software Engineering Research

Mary Sánchez-Gordón
Department of Computer Sciences
Østfold University College
Halden Norway
mary.sanchez-gordon@hiof.no

Ricardo Colomo-Palacios
Department of Computer Sciences
Østfold University College
Halden Norway
ricardo.colomo-palacios@hiof.no

ABSTRACT

This paper reports on the recruiting challenges we faced in our study on the factors influencing software engineering career choice of Andean indigenous. It also emphasizes that Indigenous people are hard-to-reach, hidden, and vulnerable samples.

CCS CONCEPTS

• The Computing Profession → Occupations

KEYWORDS

Recruiting, Ethnicity, Software Engineering, Indigenous people

ACM Reference format:

Mary Sánchez-Gordón and Ricardo Colomo-Palacios. 2022. Challenges in Recruiting Andean Indigenous Participants for Software Engineering Research. In *Proceedings of 1st International Workshop on Recruiting Participants for Empirical Software Engineering (RoPES'22). Figshare, 3 pages.* https://doi.org/10.6084/m9.figshare.19379876

1 Introduction and Context

Diversity studies have gained growing relevance in software engineering (SE) [6]. It is widely discussed that there are some underrepresented profiles in Computer Science and SE [1,6]. In particular, ethnicity has been the least studied diversity aspect in SE that characterizes underrepresented minorities such as indigenous people. The estimated number of indigenous people ranges between 370 and 500 million worldwide [9]. Based on the latest censuses available in Latin America (LATAM) [3], in 2010, there were about 42 million indigenous people in the region — representing nearly 8% of the total population. Although it is difficult to estimate increases in the indigenous population across the region, the only country that reported a decrease in its indigenous population in the past decade is Bolivia [3]. It is also reported an indigenous people's overall lower attainment of secondary and tertiary education, regardless of their location. Despite that fact, few Andean indigenous persevere and graduate from a university or college and enter the workforce [7]. The identification of factors that influence the career choices of ethnic minorities requires qualitative approaches that could provide a rich, nuanced, and actionable view of this area [4]. In our previous exploratory study [7], we conducted ten interviews with seven Andean indigenous men and three women from Ecuador (Quechuas) and Bolivia (Aymaras) to gain insights into their SE career choices. They were very driven and showed great interest in this study. Each interview was voice recorded, transcribed, and then, analyzed using coding techniques proposed by Grounded Theory. The findings provide a good overview of this topic by reveling seven factors: social support, exposure to digital technology, autonomy of use, purpose of use, digital skill, identity, and work ethic. However, these factors may not be representative of the larger indigenous population of IT/SE students and IT/SE professionals and, consequently, not generalizable to such a population or other contexts.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
RoPES'22, May, 2022, Pittsburg, PA, USA
© 2022 Copyright held by the owner/author(s).
https://doi.org/10.6084/m9.figshare.19379876

2 Recruiting Challenges

The recruiting challenges identified in this section deal with Andean indigenous because our study involved this population [7], but challenges may apply to ethnic minorities in SE.

Identifying participants. The most relevant aspect of the recruiting challenge is the identification of participants' ethnicity. The conditions of indigeneity vary over time and are, in some cases, context and country-specific. We used the most common method: participants' self-identification of ethnicity via self-reports. However, some people with indigenous heritage may not self-identify because of prejudices and stereotypes held by others against indigenous people [3]. Ethnicity may be defined at a broad level (e.g., Andean Indigenous) or a more specific level (e.g., Quechuas, Aymaras). However, indigenous communities, including Andean ones, differ in their customs, language, and types of family structure. As a result, differences in some variables between specific groups may be greater than differences between Andean indigenous and Whites. Researchers must also decide how to classify people of mixed racial or ethnic backgrounds. In LATAM, "mestizo" is a well-recognized category, so we used it. Another way to identify potential participants' ethnicity is the surname identification method, i.e., surname-based community sampling methodology. However, once individuals have identified as belonging to a particular

ethnic group, researchers must validate that they identify with such an ethnic group and share a common understanding of it.

Small sample size. Collecting data from a large enough sample of ethnic minorities has long posed a challenge, partly because of the small overall population size. In our study, Andean indigenous people make up less than 8% of the population in Ecuador. University students provide recent experience on the topic under study but only around 3% of Ecuadorian university students self-reported as indigenous in 2016. So, we considered practitioners as they can provide genuine and valuable experience, even if they can only recall part of their experiences.

Although it was an overwhelming challenge, we contacted four well-known Ecuadorian universities to get the contact information of the Indigenous students under a data protection agreement that preserves their privacy and security. It resulted in lengthy bureaucratic delays, and only two of these universities provided the requested information from 153 students who self-identified as indigenous during the enrollment process in careers related to SE. We also contacted the student communities at the other two universities, but just the indigenous students and graduates community was interested in collaborating with us and provided a list of six members. Despite the scarce number of indigenous people in those four universities, we planned to recruit at least ten participants to generate adequate data and provide comprehensive information.

The small sample size often leads to researchers combining the data from some of the ethnic groups with some common origin (e.g., combining Quechuas and Aymaras into one group) or across communities in the case of Quechuas (e.g., Tomabela, Otavalo, Puruhá, Kayambi into one group). Considering the case of Ecuadorian indigenous people, we realized that locating an adequate sample size of indigenous participants is very difficult, if not impractical. Indigenous people also vary in their degree of acculturation and exposure to community or Western cultures, whether they live in rural or urban areas, ethnic identification, experience with racism, and so forth. Thus, when researchers aggregate individuals into an ethnic group classification, must face the decision of ignoring the causes of variability. This discussion aims to promote a more refined treatment of ethnicity while not underestimating the diversity within the dominant groups, such as Mestizos.

Recruiting participants. When recruiting ethnic minority participants, researchers must consider possible ethnic and cultural differences in their willingness to engage in research studies. Are ethnic minorities, like indigenous people, less likely to give consent and contribute to a research endeavor? For some ethnic groups, cultural values and oppressive systems may influence their participation or response patterns in research. Among the indigenous populations of LATAM, there is a long history of oppression, poverty, and struggle for survival. The potential vulnerability of the participants raised serious concerns about conducting sensitive research. We engaged in careful, critical reflection on factors like mistrust of researchers, time and scheduling constraints, economic barriers related to time off work. Moreover, although many are proud of their indigenous identity, others prefer to hide or deny it wherever possible. Thus, our study design included active recruitment using direct communication, e.g. in person, by phone, or by mail, and avoiding passive recruitment such as advertising, mailing lists, forum invitation messages, or crowdsourcing tools. To build rapport and trust, we ensured participants' confidentiality and data anonymization. An indigenous software professional also collaborated with us to ensure appropriate language use and a culturally sensitive perspective. Moreover, there was a designated recruiter and one professor from each university endorsed our study to encourage participation, but we did not provide incentives (e.g., cash, prizes) to do so. We sent 159 customized email invitations, with reminders sent to non-respondents on two occasions, one week apart. We expected a low participation rate since previous research [8] reported that this rate is often less than 10% in voluntary surveys with practitioners. Although the 2% response rate discouraged us, we recruited two participants. As we also had a few contact mobile phone numbers (12), we noticed that all used WhatsApp accounts. It allowed us to send WhatsApp messages as invitations, with two reminders sent to non-respondents. Moreover, WhatsApp allowed us to call the non-respondents, and it was the most effective communication channel since the response rate was 84%. Although no one strategy is effective for all contexts, such a rate suggests that participant stimulation is important to boost response rates after the invitation was sent out. We reached out to ten individuals by voice call to provide them with information about the study, privacy rights, rights to withdraw their consent, and inquire about potential interests. Although every reasonable effort was taken to protect participants' information, the informed consent also included potential risks related to technology. After taking voluntary consent, scheduling interviews was difficult because some individuals were in rural areas and had limited internet connection. Four individuals told us that they do not self-identify as indigenous, and confirmed that the university record was wrong, and the other two did not agree to participate. In the end, we recruited six individuals, two from university records and four from the community of students.

Given that our convenience sample was small, we also applied referral-chain/snowball sampling. However, only one participant was able to identify an additional contact. We also tried to contact some indigenous organizations in Ecuador, but only one ex-coordinator of the Indigenous Organizations of the Amazon Basin told us that she had heard about another individual who studied at another university. In turn, this last participant had heard about one Aymara professional who accepted to participate and identified an additional Aymara individual who also wanted to collaborate. Despite the inherent challenges, we kept recruiting participants until we reached a sample of ten participants —four students and six practitioners.

There are significant barriers to minority participation in SE studies that are not addressed in available guidelines [2,5]. We hope that sharing our recruitment journey lends visibility to the original inhabitants of the land and encourages more research on how to enhance diversity and inclusion in the software industry.

REFERENCES

[1] Khaled Albusays, Pernille Bjorn, Laura Dabbish, Denae Ford, Emerson Murphy-Hill, Alexander Serebrenik, and Margaret-Anne Storey. 2021. The Diversity Crisis in Software Development. *IEEE Software* 38, 2 (March 2021), 19–25. DOI:https://doi.org/10.1109/MS.2020.3045817

[2] Sebastian Baltes and Paul Ralph. 2021. *Sampling in Software Engineering Research: A Critical Review and Guidelines*. Retrieved from http://arxiv.org/abs/2002.07764

[3] Germán Freire, Steven Daniel Schwartz Orellana, Melissa Zumaeta Aurazo, Damasceno Costa Costa, Jonna Maria Lundvall, Martha Celmira Viveros Mendoza, Leonardo Ramiro Lucchetti, Laura Liliana Moreno Herrera, and Liliana Do Couto Sousa. 2015. *Indigenous Latin America in the twenty-first century : the first decade*. The World Bank.

[4] K. D. Joshi, L. Kvasny, P. Unnikrishnan, and E. Trauth. 2016. How Do Black Men Succeed in IT Careers? The Effects of Capital. In *49th Hawaii International Conference on System Sciences*, 4729–4738. DOI:https://doi.org/10.1109/HICSS.2016.586

[5] Austen Rainer and Claes Wohlin. 2021. Recruiting credible participants for field studies in software engineering research. *arXiv:2112.14186 [cs]* (December 2021). Retrieved March 16, 2022 from http://arxiv.org/abs/2112.14186

[6] Gema. Rodríguez-Pérez, Reze Nadri, and Meiyappan Nagappan. 2021. Perceived diversity in software engineering: a systematic literature review. *Empir Software Eng* 26, 5 (July 2021), 102. DOI:https://doi.org/10.1007/s10664-021-09992-2

[7] Mary Sánchez-Gordón and Ricardo Colomo-Palacios. 2020. Factors influencing Software Engineering Career Choice of Andean Indigenous. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Companion Proceedings* (ICSE '20), ACM, New York, NY, USA, 264–265. DOI:https://doi.org/10.1145/3377812.3390899

[8] Marco Torchiano, Daniel Mendez Fernandez, Guilherme Horta Travassos, and Rafael Maiani de Mello. 2017. Lessons Learnt in Conducting Survey Research. In *2017 IEEE/ACM 5th International Workshop on Conducting Empirical Studies in Industry (CESI)*, IEEE, Buenos Aires, Argentina, 33–39. DOI:https://doi.org/10.1109/CESI.2017.5

[9] UNESCO. 2017. *Indigenous peoples*. Retrieved from https://en.unesco.org/indigenous-peoples