

## **Data Protection Impact Statement**

### **Need:**

SPROK's data collection has specific goals divided among 6 separate parts. The first is to observe and interview teachers in immigrant preparatory classrooms to understand how they interact with language learning materials. The second is to understand the differences between native and non-native language tracts via teacher interviews. The third is to understand issues in teaching subjects to non-native-speaking students. The fourth is to know how students with language issues end up in Tutkintokoulutukseen valmentava koulutus (TUVA) courses via teacher interviews. The fifth, and most relevant, is to collect data regarding using prototypes developed for SPROK. The last is interviewing administrators to gain a general overview of the issues in 1-4.

Referring to the attached research proposal, it can be noted that data collection will take the form of audio recordings for points parts 1, 2, 3, 4, and 6. However, we must note that goal one only involves recording the teacher interview, not the classroom observations. No children will be recorded using any visual or audio device during the data collection.

### **Part 1, Observations: RISK ASSESSMENT: LOW**

#### **Process:**

- i. Consent forms will be taken before classroom observations, which will be taken with hand-written notes by two separate observers using the same protocol. Parents will be informed beforehand.
- ii. Documents will use anonymized numbers in their identification, which are connected to consent forms.
- iii. One observer will keep the notes, and one will keep the forms.
- iv. No names or identifying features of anyone present during the observations will be recorded.
- v. The observers will carry these documents at all times until they reach their offices at Åbo Akademi University (ÅAU) in Vasa. They will be stored in a locked cabinet when not in direct use.
- vi. These notes and observations will be transcribed, stored on ÅAU computers, and transferred to local ÅAU servers. Said servers are only accessible via ÅAU

credentials.

- vii. These notes will be entered into Nvivo for processing on ÅAU laptops.
- viii. This data will only be shared with those involved in processing it or those involved in the writing process.

**Scope:**

This data focuses on what materials teachers use and how they interact with them. These observations will take place throughout their lessons for one day. Ideally, there will be 8 – 10 observations of both Swedish and Finnish language classrooms. Geographically, we wish the collection to be as broad as possible, including smaller schools in less urbanized areas if applicable, and covering most of the major cities throughout Finland. Notes may indicate the classroom’s age range to compare usage variations between age groups, and variances between instructions in Swedish and Finnish will be noted, but this is not our focus. The data will be kept for no more than two years. It will be deleted after being used to create an article.

**Processing:**

The processing of the data will be entirely on ÅAU computers. The Nvivo files will be held on no more than 2 computers. When shared, these files will be printed or transferred via secure means on the ÅAU server. The individuals we will observe are not currently known to us personally. Great care will occur in the process to ensure that they are not identifiable in any way. However, we may have interviewed some of them before in tangential research. Participants will have control in these situations, and while children will be present, they will not be the focus. Data processing will comply with the ethical guidelines of ÅAU and within GDPR.

**Purpose:**

The processing aims to provide insight into how to better design technology for utilization in Finnish language classrooms, utilizing modern methodologies. While targeted explicitly at language classrooms, the resulting observational protocol and process could be adapted to develop digital or traditional materials in other subjects that may align more directly with modern teaching methodologies. Additionally, some of the findings would be directly applicable to the development of other language materials.

**Part 5, Log Data Collection:**

**RISK ASSESSMENT: LOW**

**Process:**

1. Consent forms will be taken from parents, and only students willing to participate will be

chosen for the data collection.

- a. These students must be within the age ranges of 6-12.
  - b. They must not have learned the language being taught within the prototype.
  - c. These forms will contain only the name and location of the signee. They will be pseudonymized for privacy and enable connection to the data logs.
  - d. The pseudonym will enable the researchers to identify the user's age and mother tongue from which data will be collected.
  - e. These will be kept in a locked cabinet in a locked office at ÅAU.
2. Two researchers will be present with the students as they follow the testing protocol. This protocol will be developed with Experience Lab at ÅAU, Vasa. It may be adapted to suit different locations.
  3. The data collected will be gameplay data. Specifically, the times and accuracy of actions within the game. The processing of this data will look at performance and learning.
    - a. The devices used may be laptops, tablets, and mobile phones.
  4. Additionally, eye-tracking or the monitoring of heart rates may occur. This enables the researchers to measure interest, motivation, and enjoyment.
  5. Surveys will be used afterward to gauge the students' levels of reported interest and verify the technical findings. Additionally, these will be used to make alterations to the design.
  6. This data will be collected only using only ÅAU devices. At no point will it be installed on private or school devices.
  7. No personal data will be collected at any time during the testing process. No recordings of the students will be taken, which may identify their voice or likeness.
  8. The data will be processed using ÅAU computers. Besides the pseudonym and the above-stated data logs, no data will be entered into Firebase or ÅAU computers.
  9. The processing of this data will be quantitative, and subsequent analyses will utilize SPSS or similar software, which may include Mplus or Python.
  10. The data will be kept until articles and/or alterations to the prototypes for which it is needed are made.

**Scope:**

This data will be used to identify needed alterations to the prototypes and test their overall effectiveness. At a minimum, an experiment will be run on each individual prototype. The exact numbers will vary depending on the reported effect sizes (if any are shown) and the availability of students. The sampling will be of a wide variety and will try to focus on one language if possible. This data will be kept for no more than three years or until the associated articles and/or alterations to the prototypes are made.

**Processing:**

The processing of the data will be entirely on ÅAU computers. The data will be stored in Firebase under the pseudonyms which will be alphanumeric designations. Files will be held on no more than 2 computers. The data collection will involve no more than two researchers, and the subjects will not be known to the researchers. Great care will occur in the process to ensure that they are not identifiable in any way. Data processing, storage, and collection will comply with the ethical guidelines of ÅAU and within GDPR.

**Purpose:**

The overall purpose of the data collection is to enable processing to optimize features for utilization in digital language learning materials. These recommendations will be written into articles that will be relevant to the creation of materials in language learning as well as other academic subjects.

**Part 1, 2, 3, 4, & 6, Interviews:**

**RISK ASSESSMENT: LOW**

**Process:**

1. Consent forms will be taken before all interviews. All are semi-structured and will follow specific protocols, varying between the different parts of the research proposal.
  - a. Participants will be informed that they will be recorded, transcribed, and translated using Microsoft 365 and Teams software or Zoom if necessary.
  - b. Participants will be instructed not to mention place names, any personal identifying information, or any information that may reveal the location of their schools.
  - c. Only after this will recordings take place.
    - i. A digital voice recorder will be used for in-person interviews. These will be transferred immediately to ÅAU computers

- ii. Teams will be used to record and transcribe those interviews conducted online.
2. All recordings, transcriptions, and translations will be reviewed to ensure accuracy and anonymity before storage and processing.
3. Transcribed and translated materials will be entered into Nvivo.
4. Two researchers will analyze the data as described in the research plan.
5. This data will only be shared with those involved in processing it or those involved in the writing process.
6. It will be deleted after its intended usage and kept up to no more than 3 years.

**Scope:**

This data focuses on various information concerning language education in Finland. For further clarification on the specific foci, consult the Research Plan. These interviews will be done at the teachers' convenience. These interviews span a large geographical distance and are intended to span the education from lower primary to upper secondary. Care will be taken to remove and avoid any information that may reveal the identity of those interviewed or the location at which they work. This data will be kept for no more than three years or until it is used to create articles based on the information. The overall intention is to support the creation of materials that can be used as a first step in language education. However, this data also hopes to reveal the nuanced issues in language education in Finland and provide recommendations for administrators and policymakers on structures more suitable for current affairs.

**Processing:**

The processing of the data will be entirely on ÅAU computers. The Nvivo files will be held on no more than 2 computers. When shared, these files will be printed or transferred via secure means on the ÅAU server. The individuals we will observe are not currently known to us personally. Great care will occur in the process to ensure that they are not identifiable in any way. However, we may have interviewed some of them before in tangential research. Participants will have control in these situations, and while children will be present, they will not be the focus. Data processing will comply with the ethical guidelines of ÅAU and within GDPR.

**Purpose:**

The process will look at issues within each level of education. A final paper will combine these to examine the overall process. For details on each, consult the Research Plan. Overall, the examination will focus on how the issues with the levels combine to create the current problems related to language-level attainment in Finland.