Individual Reflective Report

Mindan Chen

Innovation Project Course-Self Report (2026)

Advisors:

Professor Marianne Graves Petersen

Associate Professor Majken Kirkegaard Rasmussen

PostDoc Mille Skovhus Lunding

PhD Student Maja Dybboe

Academic Year: June 2026

1. Introduction

MoveML is a team project. My individual work encompassed three aspects: high-fidelity interface design for students and teachers, analysis of evaluation results, and subsequent improvements to the physical wristband, ultimately forming the conceptual product I call "Animal Buddies." This report focuses on illustrating the thought process behind this work rather than listing my achievements.

MoveML belongs to a highly competitive field that includes numerous tools designed to guide children in learning machine learning by building and testing their own models, such as network classifiers like Teachable Machine (Carney et al., 2020) and the micro: bit entity model behind CreateAI for micro: bit (micro: bit Foundation, 2022). It also aligns with the core concept of "learning" in the AI4K12 framework (Touretzky et al., 2019). Within this field, my approach was more specific. In the context of machine learning education, interface design is not merely about making the prototype look more polished, but about translating the technological system into a tool that children and teachers can understand and implement. The system includes features such as device connectivity, gesture recording, shared dataset construction, model training, and predictive feedback, all of which are clear from an engineering perspective. However, this is not the case for children in a classroom. Most of my design decisions stemmed from bridging this gap, and the following sections will elaborate on what I consider to be the most important aspects.

2. Interface as Translation

My first design judgment was that the student and teacher sides couldn't be two visual versions of the same interface. Their users, responsibilities, and information needs are different. Students need to answer one question at any given moment: "What should I do now?" while teachers need to answer another: "What's happening in class? When should I move the activity forward?" Treating them as the same screen with different styles wouldn't satisfy either.

After discussions with my teammates, this distinction shifted from visual composition to information architecture. On the student side, I designed an action-centric flow: first, guidance and avatar selection to establish identity in shared sessions; then, a data collection screen built around gesture cards to record status and simple progress indicators. The goal was to concretize the abstract concept of training data, allowing children to see their actions generating samples for the model, not just "what happened."

The teacher's side is a coordination interface. It needs to display connected students, collected samples, training progress, and prediction results, because the dashboard is more than just a management tool. It's a projection object shared by the whole class, and

MoveML's learning value depends on how students see how their collective data shapes the model. Supporting this visibility is one of the requirements of AI literacy for learning tools, because children need a clear understanding of what the system is doing, not just the ability to operate it (Long & Magerko, 2020). The principle behind this predates the history of AI itself: the system should maintain visibility into its current state so that users can build an accurate model (Norman, 2013). For MoveML, this means the dashboard must expose the model's state, rather than hiding it behind a running or completed binary.

The takeaway I've learned is that high-fidelity work is often misunderstood as simply making things look prettier. Here, it's quite the opposite. The visualization layer exists to make the interaction logic clear and understandable, and the more difficult questions are: what should be displayed, when it should be displayed, and how it should support the learning process.

3. Designing for Visible Failure

Given that the dashboard is a shared projection object, we made a more specific decision. If the class model produces an incorrect prediction, this failure should not be simply dismissed as a technical error. It should be something that teachers and students can view and discuss together. This aligns with research showing that children understand machine learning more easily when its inner workings are revealed rather than hidden as a black box (Hitron et al., 2019). It also resonates with the viewpoint in learning science that a clearly visible and explainable failure can be more productive than an unhindered success (Kapur, 2008). My dashboard design ensures that model status and predictions are always visible and easily accessible, as suggested by research on children using machine learning interfaces (Dwivedi et al., 2021).

The evaluation results supported this approach in an unexpected way. At one point, the model began to favor the gesture with the most samples, and students noticed that this gesture was trained far more often than others, and were able to explain why the model consistently predicted this category. They no longer simply viewed model predictions as right or wrong; they connected the model's behavior to class-generated data, closely aligning with the key data-driven connections that data literacy research suggests children should develop (Bilstrup et al., 2022a).

After discussions with my teammates, we decided to visualize this shared structure on a dashboard, making the data imbalance clearly visible. Tseng et al. (2023) argue that individual modeling makes it difficult for learners to access data diversity and class imbalance because an individual's dataset rarely reflects these characteristics, while collaborative model building makes these dataset design practices readily apparent. Models trained by the whole class enabled children to experience and understand data imbalance. Some children explained their failures through differences, such as noticing a classmate was left-handed or performing gestures "their own way." These comments are my strongest evidence that this embodied interaction does much more than just make activities fun; embodied cognition takes on some of the conceptual load, just as collaboration, creativity, and embodied cognition play a role in the AI learning experience (Long et al., 2021). It helps children understand that data is shaped by human behavior.

4. Decorative vs. Meaningful Interactions

Not all decisions stand up to scrutiny, and avatar selection is a prime example. Students spent considerable time choosing avatars, indicating that personalized experiences were appealing to them. The problem is that the avatars played almost no role in subsequent learning. They attracted attention but contributed nothing to the core activities, creating a mismatch between where children invest their energy and where learning actually occurs.

This is a known risk, not an isolated case. Interesting but irrelevant content can distract students from key concepts, leading to decreased comprehension—a phenomenon known as the "enticing detail effect" (Harp & Mayer, 1998). Avatar selection taught me a principle I will always uphold: attention is not equivalent to contribution. Personalization should complement primary interactions, not serve as decorative embellishments; research on AI literacy for children views the relationship between children and systems as part of the core interaction, not as superficial decoration (Druga et al., 2022). Virtual avatars could function by linking a child's identity to samples of their contributions and how those samples influence shared models throughout the data collection and dashboard process. However, in terms of its design, it is merely a decorative entry point, a fact confirmed by the evaluation results.

5. From Findings to Proposal: Animal Buddies

Following the evaluation, I independently developed a redesign of the wristband, which I considered an evidence-based proposal rather than a final product. The original wristband exposed the micro: bit on a regular watchband. It served as a test device, but didn't feel like a finished product to the children. The evaluation results showed that students valued personalization, and some even commented on the wristband's appearance. During Dokk1 activities, I also noticed that children tended to integrate the worn device with interactive experiences on the screen—a so-called "physical-digital" interaction. This indicated that the wearable device was not simply a container for the micro: bit, but rather part of the children's experience. The concept of "Animal Companions" was designed based on these findings.

Instead of using a traditional watchband, I redesigned the wearable device as a series of animal characters, each retaining the visibility of the micro: bit as the core of the interaction while giving the device a child-friendly image. Viewing wearable devices as items children want to keep and personalize, rather than simply electronic devices, has precedents in the field of educational wearable devices, such as the LilyPad Arduino (Buechley et al., 2008). Viewing the physical device as a physical interface for learning machine learning aligns with the logic of machine learning machines (Kaspersen et al., 2021). The redesign connects the physical object with the concept of a virtual avatar within the interface, allowing personalization throughout the entire experience, from wearing the device to contributing data to a shared model.

This redesign also addressed a student's suggestion that the wristband should respond more strongly to incorrect predictions. While I didn't follow the student's suggestion exactly, it does reflect their expectation for tangible feedback. Future versions could utilize the micro: bit's light, sound, or display to indicate recording, prediction, success, or error, transforming the wristband from a passive sensor into part of a feedback loop.

This concept also has some practical limitations, so it's more than just a shell. The micro: bit's screen must remain visible and the buttons easy to operate; the battery module cannot make the wearable too heavy or unstable on a child's wrist; and the case and strap must fit wrists of various sizes. Overcoming these limitations constituted a large part of the design work.

Animal Buddies was not evaluated, so I am presenting it as a proposal to expand on the evaluation results to achieve tighter integration between wearable devices, virtual avatar systems, and classroom interfaces.

6. Learning Reflection

This project changed my understanding of my role far more than it improved any particular skill, and the lessons I learned mainly focused on its limitations.

The first lesson was that I initially thought my main contribution was making the interface clear and coherent; however, I later realized that the quality of the interface depended on my understanding of the underlying system. I had to deeply understand the training data, model behavior, and why predictions failed in order to take these factors into account in the design, because without this knowledge, the interface might be visually appealing but conceptually flawed. The "visible failure" decision-making mentioned in Part 3 only became possible after I understood the model bias.

The second lesson was the need to fully consider the user's perspective before making decisions. The UEQ evaluation did not consider whether users could understand the meaning of the options and make the corresponding choices. UEQ scores were encouraging, but some students struggled with the questionnaire format, so I treated this data as supplementary information rather than a measure of learning outcomes. My most trusted evidence was observing children explaining the model's behavior in their own words. Adhering to this is crucial because it's easy to interpret evaluation results as proof of the design's effectiveness. The evaluation results were actually more concrete: fourteen children were able to begin reasoning about the training data within three short lessons. The same caution applied to the "Animal Companion" project, which I viewed as a proposal based on our research findings rather than claiming it was a final solution.

The third lesson, most relevant to my own work, is a form of self-criticism. In Part Two, I argued that the interface should maintain the visibility of the system's state. I applied this principle to the model but not to the data recording process. During the evaluation, some students couldn't determine when to make a gesture or when the system was collecting data; one student suggested adding a countdown sound effect. This was a principle I set for myself but hadn't fully implemented: I made the model's state clear and understandable, but the communication of the recorded state was insufficient, leading children to sometimes operate in a system they couldn't understand their current pattern. Time and pattern visibility are interaction issues, not visual ones, and solving these problems should have been my responsibility.

A deeper issue lies in how the interface handles collaboration. The core idea of MoveML is that the classroom is a learning unit, but the student interface I designed was essentially a parallel repetition of a single-user process: each child recorded their own sample on their own screen. Collaboration existed in the shared projection and the facilitator's questioning, not in the interactive structure I constructed. The evaluation results revealed this. The temporary constraint of two students taking turns using a single computer due to a shortage of computers led to more negotiation and role division than my interface anticipated, even though quieter students often secretly observed. The first lesson was also more competitive than others, entirely due to the facilitator's approach; this competitiveness lessened once the approach changed. The sense of collaboration relied heavily on the framework, a scaffolded teaching method (Vygotsky, 1978), rather than the interaction itself, partly because my interface design failed to reflect this. Future versions need to integrate shared activity and multi-user characteristics into the interaction design, rather than treating student screens as a single user's perspective.

In conclusion, these lessons led me to view interface design no longer as the surface of a finished system, but as one of many contributions that only function effectively in relation to other contributions. I also gained a clearer understanding that my responsibility lay in the structure of the interaction, not just its appearance. This was more humble and accurate than I initially envisioned.

7. Relevant Products and Documentation

My individual work produced the following artefacts, which served as both design decisions and research documentation:

- High-fidelity student interface (onboarding, avatar selection, gesture-based data collection)
- Teacher-facing dashboard and classroom display design
- Visual design system and front-end components, built in React with Chakra UI on top of the open-source ml-trainer codebase (layout, components, states, information hierarchy)
- UEQ analysis and interpretation, using the User Experience Questionnaire (Laugwitz et al., 2008), read against its benchmark dataset (Schrepp et al., 2017); I treated the UEQ as supplementary evidence alongside observation, since some students found the format difficult
- Evaluation insights drawn from observation and post-activity discussion
- Animal Buddies wristband concept
- Major Contributions to the Design Process and Evaluation sections of the team report.

Reference

Bilstrup, K.-E. K., Kaspersen, M. H., Lunding, M. S., Petersen, M. G., Schaper, M.-M., Van Mechelen, M., Tamashiro, M. A., Smith, R. C., & Iversen, O. S. (2022a). Supporting critical data literacy in K-9 education: Three principles for enriching pupils' relationship to data. *Proceedings of the 21st Annual ACM Interaction Design and Children Conference (IDC '22)*. https://doi.org/10.1145/3501712.3530783

Provided the argument that children should develop a critical relationship to the data behind ML systems. I used it to frame students connecting the biased model back to the class's own data as evidence of emerging data literacy rather than a one-off observation.

micro: bit Foundation. (2022). *CreateAI for micro: bit.* https://createai.microbit.org/

Provided the embodied micro: bit ML tool that MoveML builds on, since MoveML extends the same ml-trainer codebase. I used it to position MoveML within an existing line of work rather than as a standalone invention.

Buechley, L., Eisenberg, M., Catchen, J., & Crockett, A. (2008). The LilyPad Arduino: Using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 423–432. https://doi.org/10.1145/1357054.1357123

Provided precedent for treating a wearable computing device as an object that children personalise rather than as bare electronics. I used it to justify reframing the wristband as the Animal Buddies wearable.

Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., Jongejan, J., Pitaru, A., & Chen, A. (2020). Teachable Machine: Approachable web-based tool for exploring machine learning classification. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*, 1–8. https://doi.org/10.1145/3334480.3382839

Provided an example of a novice-friendly tool where users build and test their own ML classifiers. I used it to locate MoveML among existing build-and-test ML learning tools.

Druga, S., Christoph, F. L., & Ko, A. J. (2022). Family as a third space for AI literacies: How do children and parents learn about AI together? *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*, 1–17. https://doi.org/10.1145/3491102.3502031

Provided child-facing AI-literacy research treating a child's relationship to the system as part of the core learning interaction. I used it to support my argument that personalisation should be built into the activity rather than added as surface decoration.

Dwivedi, U., Gandhi, J., Parikh, R., Coenraad, M., Bonsignore, E., & Kacorri, H. (2021). Exploring machine teaching with children. *2021 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 1–11. https://doi.org/10.1109/VL/HCC51201.2021.9576171

Provided design guidance that ML metrics and outcomes should be visible so children can experiment with them. I used it to back my decision to keep prediction outcomes inspectable on the teacher dashboard.