Interface Digital Twins: Rendering Physical Devices Accessible to People who are Blind

Nicholas Caporusso
Department of Computer Science
Northern Kentucky University
Highland Heights, United States
caporusson1@nku.edu

Nicholas Elleman
Department of Computer Science
Northern Kentucky University
Highland Heights, United States
ellemann1@nku.edu

Samuel Sungmin Cho
Department of Computer Science
Northern Kentucky University
Highland Heights, United States
chos5@nku.edu

***Abstract*— An Interface Digital Twin (IDT) is a digital replica of a physical user interface that uses mobile device accessibility features (e.g., text-to-speech, vibration, and gesture control) to make the interface more accessible, help users understand its components and learn how to use it, and provide them with additional information for completing their tasks on the actual device. Although the concept of IDT has been designed to render physical devices and appliances accessible to individuals who are blind, it can be applied to different sensory and cognitive conditions. In this paper, we detail the key components and implementation roadmap of IDTs and we discuss how they can enhance the usability and accessibility of the User Interface UI of physical devices and appliances.**

Keywords- accessibility; universal design; user interface, Assistive Technology.

I. Introduction

Approximately 12 million people 40 years and over in the United States have vision impairment, including 1 million who are blind [1]. As a result of poor accessibility of user interfaces of physical devices, people with disabilities, and especially individuals who are blind, are unable to independently use several devices, including basic home appliances. Designing the user interface of novel devices results in the challenging trade-off between the advantages of using modern components and controls that are appealing to most users (e.g., touchscreens) and the degree to which the interface can be utilized by a minority of individuals who are blind or have other conditions.

Improving product usability and, particularly, reducing access barriers is crucial for serving an increasing number of users with diverse needs and abilities, such as the elderly and people with cognitive disorders or physical impairments. The introduction of Universal Access principles resulted in improvements in several aspects of Information Technology (IT), over the last decade. However, there still exists a large divide between the digital and physical world in terms of equitable access to devices and products. Previous research has demonstrated that the lack of tactile feedback and affordances can significantly increase the complexity of an interface [2] [3] [4]. In some cases, being able to independently use a product for the first time might be difficult or impossible. Some challenges include labels and messages that are not translated into tactile or auditory form (e.g., credit card readers) or lacking touch or sound cues. Despite existing accessibility guidelines, regulatory gaps and the cost of updating existing technology make it difficult to rapidly implement new solutions that are applicable to systems currently on the market [5] [6] [7].

In this paper, we present the concept of Interface Digital Twin (IDT) and we discuss how it can be applied to create accessible replicas of the User Interface of a physical device that can be navigated with a smartphone using touch and audio. We describe the architecture of IDTs and we detail its integration in different types of existing technology using affordable resources that are already available, widespread, and easy to use.

II. System design

The objective of our work is to develop an affordable and immediately actionable solution to the lack of accessibility of many products currently on the market. Specifically, our goal is to empower users and organizations that aim at enhancing their accessibility compliance to directly implement improvements without requiring any intervention from the manufacturer or modification to the equipment. To this end, in a previous work [8] we introduced the concept of IDT, that is, the digital replica of the UI of a physical device or appliance. Specifically, IDTs can be designed to reproduce the visual and functional components of the UI of a device, so that they can be visualized on a smartphone, so that users can navigate it. Also, additional information such as descriptive labels can be added to provide additional guidance to the user. By doing this, people who are blind can navigate the interface with the help of the integrated text-to-speech engine, which has the purpose of translating the location and arrangement of the visual components of the interface, as well as further instructions, into auditory formats.

To this end, as described in Figure 1, the typical interaction workflow with an IDT would be as follows:

1. the user approaches the device and recognizes the presence of a tactile tag that indicates the availability of an IDT
2. the user uses their smartphone to scan the tactile tag, which contains a QR code with the link to the URL where the IDT is located
3. the smartphone downloads the IDT from the server and shows it to the user in the preferred format
4. the user navigates the interface on their smartphone, listen to the audio guide, and get additional information as text or in auditory format
5. after they understand how to operate the interface, they continue their task on the actual device.

This would enable individuals with different abilities and needs to access the information they require for completing the task. This, in turn, is especially relevant in the case of individuals who are blind. Additionally, IDTs can serve a more general audience, because they can be utilized as a digital manual by people who have never used the device before.

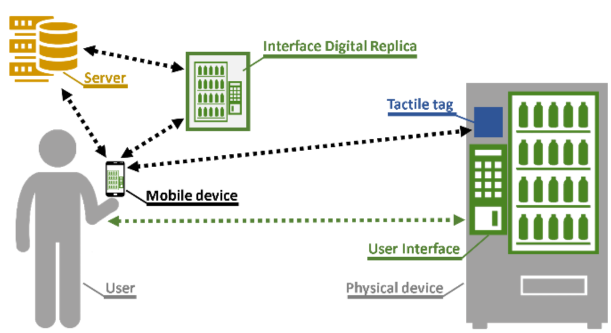

Figure 1. Example workflow of IDT interaction with a vending machine.

In general, IDTs have been conceived to be utilized when the user is in close proximity with the physical device or object the IDT is associated with. Specifically, as shown in Figure 2, we expect individuals to primarily interact with them within in the Intimate (0 – 50cm), Personal (50cm – 1mt), or Social (1 – 4 mt) proxemic spaces. For instance, a user can download and access the IDT of the User Interface of a vending machine with their smartphone, by either scanning a tag or selecting the device among the ones available nearby. However, the proxemic domain of the interaction with an IDT depends on the goal of the user: the IDT can be considered a just-in-time, accessible digital manual that helps the user understand how to accomplish their tasks. Therefore, the user can scan the tactile tag on a device and then leave the intimate and personal space while keep interacting with the IDT and with the device, if enabled, as discussed below.

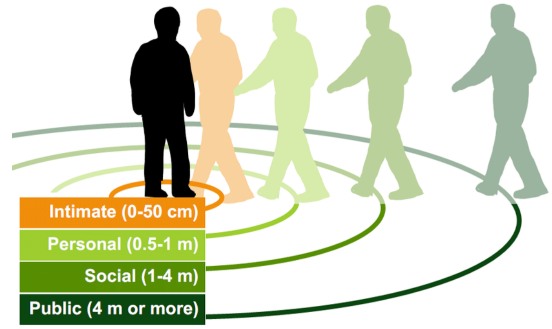

Figure 2. Proxemics and interaction scenarios with IDTs.

Nevertheless, as IDTs are stored in a centralized system, they can be accessed directly from their web address, which results in additional advantages from a proxemics standpoint. By doing this, users could realize their tasks without being in the same intimate or personal space as the device. For instance, with sensory or motor impairments complete a purchase from a vending machine without necessarily visiting it, by asking someone to collect the product on their behalf.

The concept of IDT incorporates that of Interface Description Language, which has been studied and developed in the past to support the description of the interface of software applications independently from their implementation in a specific language or on a particular computer architecture. Many User Interface Description Languages (UIDLs) have been introduced so far to address different aspects of a User Interface. In general, as the role of IDLs is primarily with the purpose of detailing the components of the Graphical User Interfaces (GUIs) of a software application, they incorporate a large dictionary of attributes that are used to describe the type, structure, placement, and appearance of each component of the interface, including its functions. In the case of IDTs, the purpose of the language is not to accurately represent the interface in a digital format, but instead provide a simplified and lower fidelity visualization that simplifies the learning process. To this end, IDTs are defined using a standard language that can be utilized for describing the main characteristics of the user interface of a physical product from an accessibility standpoint. To this end, the language should be concise, interoperable, extensible, and complete. Furthermore, as the overall objective of the project is enabling non-technical users to contribute an IDT, the language should have a low level of complexity. In addition to detailing the structure of the interface and its characteristics, the language should be able to support more advanced features, such as, the operations that can be realized by the device. The development of such language with the goal of limiting the required elements of the language to maintain a good level of simplicity and, simultaneously, enable expanding the syntax to support additional applications and scenarios.

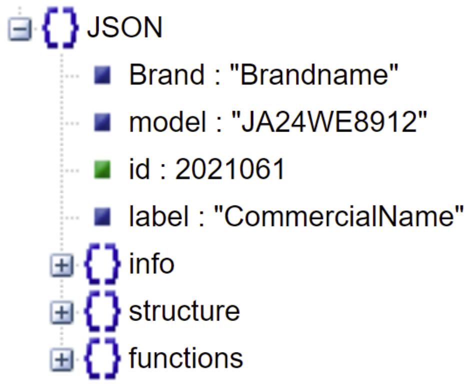

Figure 3. Example of the IDL of an IDT. A JSON structure is utilized to describe the key attributes of the IDT, additional information about its, the structure of the physical interface, and the functions associated with each of its components.

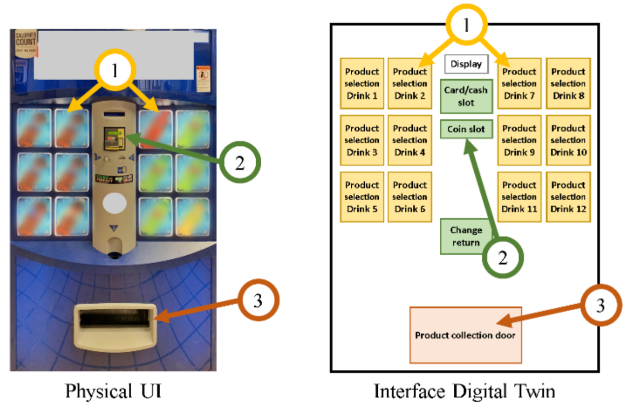

Figure 4. Example of the User Interface of a physical device, i.e., a vending machine (left), and its corresponding IDT implemented in HTML (right).

The second component of the infrastructure for IDTs consists in creating a rendering engine that can interpret the language and translate it into different formats that are instrumental to the goal of IDTs. To this end, the rendering engine is conceived as a multimodal parser that can convert the textual representation of the IDL implemented in JSON into visual and auditory formats, which users can easily access, navigate, understand, navigate, and interact with. Furthermore, by operating on the user's device, the rendering engine can adjust its operation to the characteristics of the user, as specified in their account settings, to adjust to their specific needs. As a result, the engine could activate the text-to-speech system in the case of individuals who are blind, a different coloring system for people who are colorblind, or introduce additional guidance for people who have cognitive conditions.

Several programming languages can be utilized to parse JSON and convert the definition of IDTs into visual and auditory formats on a variety of mobile devices and platforms, from native applications to websites. Specifically, IDTs are conceived to be supported client-side web technologies, such as, the HyperText Mark-up Language (HTML), Cascading Style Sheets (CSS), and JavaScript. The system is designed so that interfaces can be viewed with a standard web browser, in order to maximize the trade-off between software requirements and compatibility, as shown in Figure 4. By doing this most users can access them without installing additional dedicated applications. Nevertheless, as some of the key functions of IDTs (e.g., vibration or touch gestures) are not currently supported by some browsers, the system works best if users access the interface with specific browsers. To this end, the system checks the requirements of the interface and provides users with a list of compatible browsers in case the features required by the interface are not available on the current browser. Alternatively, if none of the browsers available for the device and operating system of the user support the required features, a dedicated application can be utilized to process the content of the interface and render it bypassing the restrictions of the operating system.

As described in Figure 5, the description UIDL and multimodal rendering engine constitute the core components of IDTs, and they are sufficient for providing users with features that enable to navigate the structure of the interface, understand its operations, and subsequently realize the intended task on the actual device. Each IDT can be associated with a unique URL that can be utilized to access their navigation. Moreover, the web address can be incorporated in a tag that enables users to easily retrieve the digital replica when they interact with the physical device. IDTs can be stored and made available to the users via a web platform. This, in turn, acts as a centralized repository that enables developers to create and share IDTs that users can download and access on their device.

Specifically, the platform has the purpose of defining the standards and describe the requirements, specifications, and implementation guidelines for the design of IDTs; by doing this, developers can produce IDTs that adhere to a common set of rules and protocols. In addition, the platform can host wikis, discussion boards, and other communication tools that support user-centered and participatory design practices: members who want to contribute to the project can discuss changes to the standards, promote updates and improvements, and share developers' guides and tutorials. Secondly, the platform acts as a centralized hub for sharing and accessing IDTs.

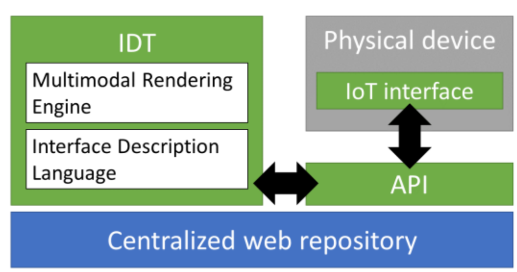

Figure 5. Architecture of IDTs. The Interface Description Language and Multimodal Rendering Engine constitute the core components of IDTs. A centralized web repository, as well as individual repositories, can be utilized to store IDTs and make them available to users. In addition, an IoT interface could be implemented in physical devices and connected to IDTs through a set of APIs.

Although an IDT increases the accessibility and availability of information for the interface of a physical device, communication between the appliance and the phone of the user is one-directional, only. Indeed, as the IDT is downloaded from a centralized web repository, the actual device is not involved in any interaction: the exchange of information is realized between the platform hosting the IDT and the device of the user. The IDL that constitutes the IDT specification and the Multimodal Rendering Engine enable users to access a browsable digital interface that does not incorporate any control capabilities, unless the physical device is equipped with an Internet of Things (IoT) module that supports receiving commands issued on the IDT via standard Application Programming Interface (API) that can be accessed through Internet protocols such as HTTP. By doing this, the proposed system can be utilized to establish bi-directional communication between the user and the device, thus, supporting browsing the interface via its IDT and simultaneously controlling the functionality of the device.

Therefore, the IDL of the proposed system should include components that extend the functionality of the IDT beyond just representing the structure of the interface. To this end, the IDL schema enables specifying the functions that can be accessed through the IDT (see Figure 3). They include the Internet URL of the API.

As a result, IDTs could implement all the necessary features for achieving a complete interaction with a physical device and, ultimately, they could entirely substitute the need for its physical UI. For instance, several models of vending machines are adjusting to the transition to a cashless society by incorporating contactless payment systems that enable users to complete the purchase with their phones. Eventually, switching from the concurrent use of physical and digital interfaces to exclusively using the latter, would result in more accessible and customizable UI where interaction can be enhanced with additional information, engagement, and support. Simultaneously with improving the overall user experience, this would result in removing several electromechanical components from the physical device and, consequently, reduce system production and maintenance costs.

III. Conclusions and Future work

Individuals with disabilities, with specific regard to vision impairments, are unable to independently operate a variety of equipment, ranging from simple domestic appliances to more sophisticated devices, including publicly available vending machines. In this paper, we introduced IDTs, a novel approach that aims at providing users with a system that can instantly and inexpensively make any interface accessible. We proposed the concept of IDTs, which are digital copies of physical user interfaces that a user may traverse using their smartphone's accessibility capabilities. By navigating an IDT before using the actual devices, users could be able to navigate the UI and understand the location and function of the interface's components, as well as learn how to use the device and complete their tasks. The concept of IDTs originates from the idea that physical interfaces can have an entirely digital replica that provides users with increased accessibility and facilitates understanding the dynamics and functioning of the interface. To this end, IDTs are not simply a digital representation of the interface, which could be realized with virtual reality or other types of technologies. Conversely, IDTs are designed to incorporate the information necessary to enable users to effectively understand and learn how to use a device, considering their different abilities and needs. Moreover, IDTs could be tightly coupled to the physical interface, to provide users with control capabilities.

In this paper, we detailed the architecture of IDTs and the overall implementation strategy and roadmap. In our future work, we will introduce a complete specification of the IDL, the rendering engine, and we will present a fully working prototype of the system.

References

[1] R. Varma, T.S. Vajaranant, B. Burkemper, S. Wu, M. Torres, C. Hsu, F. Choudhury, and R. McKean-Cowdin, “Visual impairment and blindness in adults in the United States: demographic and geographic variations from 2015 to 2050”. JAMA ophthalmology, 134(7), pp.802-809. 2016.

[2] N. Caporusso, K. Udenze, A. Imaji, Y. Cui, Y. Li, and S. Romeiser, ” Accessibility Evaluation of Automated Vending Machines.” In International Conference on Applied Human Factors and Ergonomics (pp. 47-56). Springer, Cham. 2019.

[3] A. Bhowmick, and S.M. Hazarika, “An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends”. Journal on Multimodal User Interfaces, 11(2), pp.149-172. 2017.

[4] M.S. Baldwin, R. Khurana, D. McIsaac, Y. Sun, T. Tran, X. Zhang, J. Fogarty, G.R. Hayes, and J. Mankoff. “Tangible Interfaces”. In Web Accessibility (pp. 715-735). Springer, London. 2019.

[5] A.F. Siu, “Advancing Accessible 3D Design for the Blind and Visually-Impaired via Tactile Shape Displays”. In The Adjunct Publication of the 32nd Annual ACM Symposium on User Interface Software and Technology (pp. 146-149). 2019.

[6] M. Buzzi, B. Leporini, and C. Meattini, “Simple Smart Homes Web Interfaces for Blind People”. In WEBIST (pp. 223-230). 2018.

[7] A. Khan, S. Khusro, and I. Alam, “Blindsense: An accessibility-inclusive universal user interface for blind people”. Engineering, Technology & Applied Science Research, 8(2), pp.2775-2784. 2018.

[8] Caporusso, N., Elleman, N. and Cho, S.S., 2020. Interface Digital Twins: Rendering Physical Devices Accessible to People who are Blind. In 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 982-986). IEEE.