

การบททวนวรรณกรรมการใช้งานของผู้ช่วยอัจฉริยะด้วยคำสั่งเสียง และความเป็นไปได้ในปัจจุบัน

Voice User Interface (VUI): A Review of Present and Potential Voice Assistant (VA) applications

កណ្ឌិតា អនុរោមជី* សកម្ម ទីផ្សារវេចខាយ នូវខ័ណ្ឌិត បាន

Krittiva Tanemanee*, Sakol Teerayavrunyou and Nattapit Bujaban

คณิตศาสตร์ทางการเงินและภาษีอากร หน่วยเรียนบัญชีวิชาเอกบัญชีและคอมพิวเตอร์เพื่อประกอบธุรกิจ รายวันเดือน กันยายน จำนวน 10 ชั่วโมง

School of Architecture and Design, Bangkok University, Bangkok 10150 Thailand (nardee@mail.com)

*Corresponding author

Received: December 10, 2020 / Revised: December 28, 2020 / Accepted: December 29, 2020

ABSTRACT: Voice user interface (VUI) could come to replace the graphic user interface (i.e., mobile phone screen and computer). The reason is that users interact with VUI naturally more than touch screen interface. Nevertheless, the VUI with voice assistants still has problems in both acceptance and usability. The factors like privacy issue, voice assistants' personality, the differentiation of age, language impacts the usability. For this review, 30 papers from the database of conference and research are investigated. Many researchers recommended the multimodality of VUI both input and output make an interface ease of use. For user experience, the speech technique and reformulating query and modeling users is the technique that makes machines understand the context of use better.

សំគាល់ និងរួមចូលរឹបចំណុចក្នុងការអនុវត្តន៍ការងារ។

Kazanovic-Vojica, assistant, user experience, part of user

1. INTRODUCTION

Voice user interface (VUI) is an interaction between human and computer by using spoken language. The technology includes speech recognition to understand spoken commands and answer questions, and using text to speech to reply to users. Radio Rex was the first voice-activated toy released in 1911 and it was developed to be the shoebox designed by IBM [1]. It can recognize 16 words spoken into its microphone and convert those sounds into electrical impulses. The mechanism behind the shoebox is the Automated Speech Recognition (ASR) and Natural Language Understanding (NLU). Later, three Bell Labs researchers, Stephen Balashuk, R. Biddulph, and K. H. Davis built a system called "Audrey" for single-speaker digit recognition [2]. Their system located the formants in the power spectrum of each utterance.