Use of Voice Recognition and Speech Command Technology as an Assistive Interface for ICT in Public Spaces.

A whitepaper published by Peter W Jarvis (Senior Executive VP, Storm Interface) and Nicky Shaw (Operations Manager North America).

The Use of Voice Recognition and Speech Command Technology as an Assistive Interface for ICT in Public Spaces.

A whitepaper published by Peter W Jarvis (Senior Executive VP, Storm Interface) and Nicky Shaw (Operations Manager North America).

FINAL DRAFT. For consultation, comment and public review.

September 2018.

Introduction.
The emergence and **increasing use of smart speakers (AI) in the home environment has delivered significant benefits for those with mobility, sensory, cognitive or dexterity impairment.** For millions of disabled people voice recognition and speech command technology, allied with audible confirmation and presentation of requested information, permits more **informed decision making and personal control of their immediate environment.**
This improved access to information and control opens a new world of communication, entertainment, education and opportunity for those who are unable to see, read or interact with content presented on a display screen and for those who lack the mobility or dexterity to manipulate tactile system interface devices (such as keyboards, trackballs or touch screens etc.). **Speech Command Technology creates significant new opportunities for independent living.**

This improved accessibility also creates unique challenges for system designers, legislating authorities and those concerned about privacy and misuse of personal data. As Voice Recognition and Speech Command technology moves **beyond the domestic environment into public spaces and the urban infrastructure** we will need new guidelines to increase public awareness and new regulation to protect the general population against the misuse of recorded information.

This whitepaper explores the implementation and integration of Speech Command technology within ICT kiosks and self-service applications. It is intended to provide a framework for a proposed Code-of-Practice. This CoP to be drafted for public consultation and possible adoption by the Kiosk Manufacturer Association (KMA) as an addendum to its Accessibility Guidelines.
To illustrate certain devices or technologies there are some references in this document to products manufactured by Storm Interface. These are intended as exemplars only. Other brands and products are available.

1. **Who's Listening**
   1.1 When a private citizen purchases a connected smart speaker device for home use, he/she makes an informed decision to install that device into their home environment. Before connecting their new device to the manufacturer's cloud-based AI applications new customers are required to agree and accept many terms and conditions of service. By doing so they make a decision to accept a listening device into their home; albeit with an option to mute that device or switch it off at any time. The customer knows where the device is located, what its connected status is and how to switch it off.
   1.2 However, to overcome the latency (delay) inherent in delivering cloud-based AI services to a device that has just been switched on, these devices (by default) usually remain in a powered and connected configuration. Amazon have referred to this default configuration as "Always on, always ready". This configuration is sometimes referred to by more cynical commentators as "Always on, always listening". The device needs to be configured in this way to operate as an effective 'hands free' Voice Recognition and Speech Commanded information system.

2. **In a Public Environment.**
   2.1 Speech Command and Voice Recognition technology will provide an effective and valuable improvement in accessibility to public ICT systems. Applications such as public transport ticketing and airline check-in terminals would be typical examples.
   2.2 As part of a multi-modal approach to accessibility, Speech Command will provide an additional option for those with disabilities (and those without) to confirm their biometric identity and to interface with the kiosk's application software. The kiosk user will be able to choose from a combination of tactile, audible or visual interface devices to best meet their specific accessibility needs.
   2.3 However, it will be essential that all kiosk users and those members of the public in proximity to the kiosk be made aware that the terminal includes Voice Recognition and/or Speech Command technology and that the Speech Command facility is "**on and listening**". This awareness is essential for two reasons:
      2.3.1 To inform the kiosk user that Speech Command / Voice Recognition technology is available for their use and convenience.
      2.3.2 To warn members of the public (in proximity) that their conversations may/will be picked up by the Speech Command / Voice Recognition facility and may be transmitted to a remote server for analysis, processing and possible retention.
   2.4 This awareness must be provided for members of the public who are sighted, partially sighted, non-sighted or hearing impaired.

3. **A Universal Symbol**
   3.1 It is proposed that a universally recognized symbol for Speech Command functionality be adopted by the Kiosk and Self Service industry.
   3.2 The symbol's purpose is to indicate the presence of Voice Recognition or Speech Command technology.
   3.3 Storm Interface have designed a high contrast, highly visible and tactilely discernible symbol that can be easily applied to the kiosk. During the development of this logo, Storm Interface worked closely with the UK's Royal National Institute of Blind People (RNIB). Feedback received from the RNIB has influenced the logo design. This to aid recognition and ease of use, and to ensure that all contours and edges are rounded to make it comfortable to the touch.
   3.4 As with any new logo, but in particular tactile logos, people will need to learn its meaning. This highlights the importance of introducing a standard logo which can be used across all kiosks and sectors to ensure that blind people need only learn one symbol.
   3.5 When Voice Recognition or Speech Commanded services are activated the symbol will be illuminated with bright white LEDs.
   3.6 The applied symbol should be positioned such that it can be easily seen or tactilely located as a user approaches or addresses the kiosk.
   3.7 When the kiosk is in home screen or screen saver mode, with no detected user activity, an audible signal or statement to indicate the presence of an activated Voice Recognition or Speech Command facility should be played periodically. Alternatively, a proximity sensing device could be used to un-mute a VR or SC device only when a kiosk user approaches the kiosk interface zone.
      3.7.1 Similar audible indicators of a functioning Voice Recognition or Speech Command technology should also be given when such a facility is activated (switched on or un-muted) after a period of non-functionality.
   3.8 A proposed specification for the symbol is reproduced below. Storm Interface and the RNIB propose to make this symbol available as a "free-to-use" graphic device. Storm Interface propose to offer a physical, manufactured version of the graphic device, in the form of an illuminated tile, for sale to and use by kiosk manufacturers, specifiers or operators.
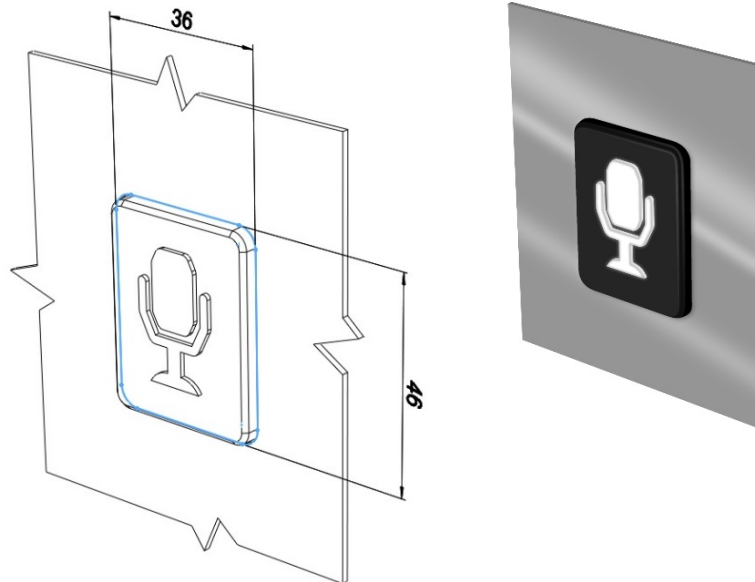
*Figure 1: **Images courtesy of Keymat Technology Ltd. All rights recognized**.*

4. **Hardware**
   4.1 Microphones
   - 4.1.1 Kiosks that offer Speech Command or Voice Recognition technology must support and provide the means for voice input.
   - 4.1.2 This should be by provision of a suitable standard connection point for an audio headset or ear piece (equipped with its own microphone) and by provision of a suitable microphone (or microphone array) permanently installed as a fixture of the kiosk.
   - 4.1.3 In many public kiosk locations or applications it will be necessary to employ advanced noise cancelling and beam focusing technology to enable effective operation of the Speech Command or Voice Recognition technology.
   - 4.1.4 Connection of a headset or assistive hearing device (equipped with its own integrated microphone) should be detected by the host kiosk and the functionality of any permanently installed microphone (or microphone array) should be automatically adjusted to accommodate and allow correct functioning of the headset or hearing aid device
   - 4.1.5 To facilitate reliable and continued functionality, provision and installation of audio device connection points and/or permanently installed microphone devices should accommodate requirements for regular sanitation (wash-down) procedures and should resist the hard use and abuse associated with ICT installations in public spaces. As a minimum requirement, water and dust resistance in accordance with IP54 (or equivalent) should be achieved. A minimum impact resistance of 10J should be achieved.
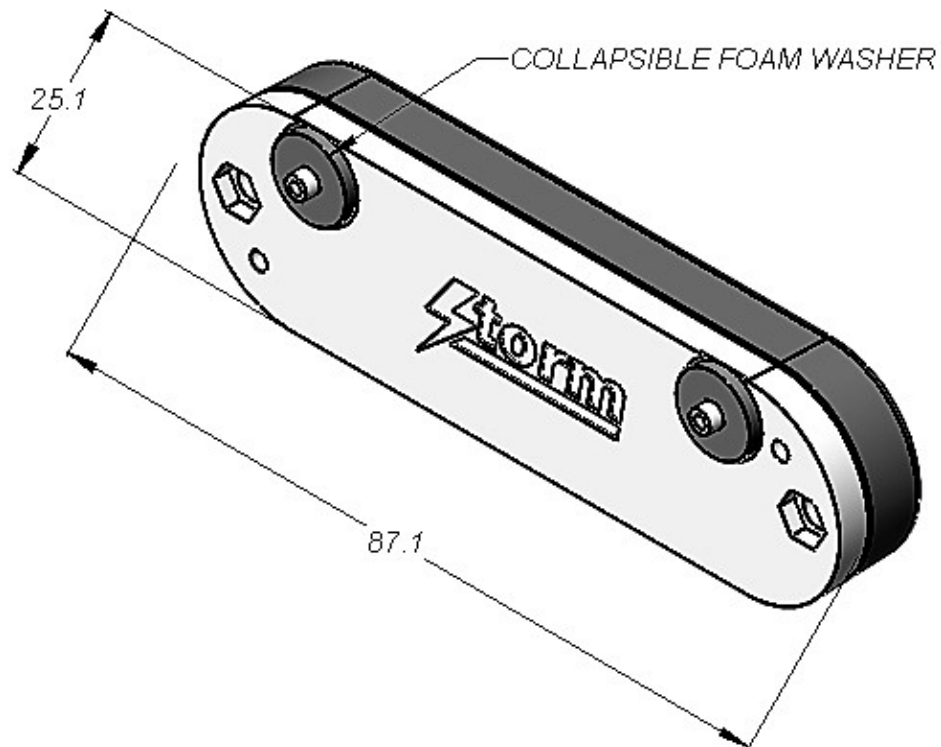
All dimensions are in millimeters (mm)

Figure 2. ***Beam array microphone for outdoor or unsupervised public environments. Other brands and products are available.***

4.2 Speakers

4.2.1 Kiosks that offer Speech Command or Voice Recognition technology must support and provide the means for audible reproduction of sound or speech.

4.2.2 This should be by provision of a suitable connection point for an audio headset or earpiece and by provision of a suitable amplified speaker system permanently installed as a fixture of the kiosk.

4.2.3 In many public kiosk locations or applications it will be necessary to employ sound directing or sound focusing technology to prevent noise pollution or irritation to those in the local vicinity of the kiosk.

4.2.4 Connection of a headset or assistive hearing device (equipped with its own integrated speakers) should be detected by the host kiosk and the functionality of any permanently installed amplified speakers should be automatically adjusted to accommodate and allow correct functioning of the headset or hearing aid device.

4.2.5 Tactile discernable sound volume controls must be easily accessible to those using assistive headsets, earpieces or hearing aid devices. Tactile sound volume controls should be accessible and functioning throughout the kiosk user session. Wherever possible tactile discernible controls should be suitably shaped to enable function with headsticks or assistive easy grip styli.

*Figure 3. **Tactile discernable sound volume controls must be easily accessible to those using assistive headsets, earpieces or hearing aid devices and those using headsticks or easy-grip styli.***

4.2.6 To facilitate reliable and continued functionality, provision and installation of audio device connection points and/or permanently installed amplified speakers should accommodate requirements for regular sanitation (wash-down) procedures and should resist the hard use and abuse associated with ICT installations in public spaces. A minimum requirement for water and dust resistance in accordance with IP54 (or equivalent) should be achieved. A minimum impact resistance of 10J should be achieved.

4.3 Wireless Devices

4.3.1 For those kiosk users who prefer to use wireless headsets, earbuds or implants in preference to wired devices with a cable and jack-plug connector, it should be possible to connect a personal wireless transponder (powered by a button cell battery) into the jack-plug socket. These personal devices provide encrypted communication between the transponder and a paired personal headset. The transponder would be removed and retained by the kiosk user when the kiosk session is completed.

*Figure 4: **Compact wireless transponder. These devices can be paired with a wireless headset or earpiece to provide a private listening capability. The transponder can be plugged directly in to the kiosk's audio jack socket. Other brands and types of transponder are available.***

5. Conclusions:

The emergence of Voice Recognition as a means of biometric confirmation of identity, coinciding with the profound impact of AI on speech commanded ICT, will drive adoption of speech command technology in public spaces and applications. Whereas this presents many challenges and risks to privacy and protection of personal data, it will lead to a new era of equality in access to information, freedom and independence for those with disabilities. It will be necessary for accessibility mandates, regulation and standards to be adapted in support of this revolutionary change in the way humans interface with the digital world. **Speech Command Technology creates significant new opportunities for independent living.**

Ends.

Contact: Peter Jarvis: peterJ@storm-interface.com
Nicky Shaw: nickys@storm-interface.com

APPENDIX A: Feedback and Comments from industry, disability support groups and public consultation.

28th September:
Comments from: Pat Pound. Open Doors Organization.
Subject: Re: FINAL DRAFT Whitepaper. The Use of Voice Recognition and Speech Command Technology as an Assistive Interface for ICT in Public Spaces.docx

This looks quite good and comprehensive, I can tell a lot of work has gone into it. Only have a couple of thoughts.
I would like to see people with speech impairments included in the list of impairments, and in the design. For example, they're probably needs to be a requirement so that the input session does not time out if people speak slowly. Also, a feature for making corrections needs to be identified and easily used.   also, are Bluetooth input devices going to be allowable? Another thought is dictating passwords is always a challenging area. Any thoughts about ways to do this? Perhaps there can be letters or characters on the screen and you react as you see them come up rather than saying it to the letters that someone could copy. Just a thought. I also prefer voice input systems give me a choice of either speaking certain words or pressing numbers or letters. Sometimes people with speech impairments can say some things more clearly than others.   Ends.