Towards a more personalised user experience and better
demographic data on the Digital Doorway public computer terminals

Kim Gush
Meraka Institute, CSIR, South Africa

Abstract: The Digital Doorway project provides rugged multi-terminal computers, enabling users in
underprivileged rural communities in South Africa to teach themselves basic ICT skills without the need of a teacher
present. In addition, the terminals provide a mechanism for educational content delivery and community upliftment
through various means. The paper highlights recent progress made towards answering the question: “How can existing
as well as new Digital Doorway infrastructure be used to increase the user's sense of ownership of the computer and
improve demographic data collected on the Digital Doorways?”

The following topics are addressed:
- Digital Doorway system architecture and software
- Personalised and shared user spaces
- The need for functionality enhancing interventions on the Digital Doorway
- A mechanism for users to transfer files from the Digital Doorway to a personal device
- The collection and analysis of quantitative user data and application usage statistics

The current Digital Doorway configuration is a 3-terminal set-up with a client/fileserver PC and two diskless clients.
A satellite downlink and GPRS backhaul provide a mechanism for content delivery and log file retrieval. A Bluetooth
dongle and GUI application allow a user to save files from the Digital Doorway to a personal device.

Users have the opportunity to create their own user accounts and enter personal information such as age, gender and
home language. These data, together with statistics on application usage, are collected and transmitted to a central
server. User desktops can be personalised using the built in webcam. This provides an incentive for users to create their
own user logins, thus increasing the accuracy of the collected demographic data.

Through the comparison of two hypothetical scenarios, it is proposed that an effective installation requires ongoing
content updates and system maintenance. In addition, a sense of ownership of the equipment needs to be fostered
among the community as a whole and among individual users.

Results from the data collection process indicate a dominance of male users between the ages of ten and twenty-four
years old. Despite the rural location of most sites, the language of choice for most users is English. The twenty most
popular applications on the Digital Doorway are identified.

It is possible to visualise the data collected at each site, establish which sites are well used and identify sites that are
undergoing technical difficulties. Unsuccessful sites may require technical or social intervention, or even relocation.
Improved network visualisation mechanisms will contribute to speedier fault identification and greater system uptime,
in turn, providing a more meaningful ICT education tool.

Keywords: “Digital Doorway”, computer kiosk, ICT literacy, HCI, demographics, South Africa.
INTRODUCTION

The Digital Doorway project commenced in 2002 and sought to bring computer technology within reach of disadvantaged rural communities in South Africa. The project has the altruistic purpose of narrowing the digital divide and is a tangible implementation of one of the aims mentioned in former President Thabo Mbeki’s State of the Nation address delivered in 2002, “a critical and pervasive element in economic development in the current age is the optimum utilisation of information and communications technology” (Mbeki 2002).

The primary aim is to allow users within walking distance of the device to teach themselves basic computer literacy skills through experimentation and peer learning, without the intervention of an official teacher or instructor. Secondary aims include delivery of educational content, provision of community-specific content and uplifting the community through mesh networking services and ICT enabling technologies. Each of the currently installed 150+ sites in the country consists of a rugged 3 or 4-terminal computer system running open source software and content, in a publicly accessible location such as a recreation hall, community centre or school. Motivation for the choice of open source software is covered in a separate paper (GXXXX 2004); further background to the Digital Doorway project is available in (XXXX et al. 2004) and (Smith, R 2005).

A brief time line of the project is given in Table 1.

Table 1: Digital Doorway Time line

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>Dr Sugata Mitra of NIIT, India trials a mechanism to observe “unassisted learning” of a computer system and coins the phrase “minimally invasive education” (MIE) Error! Reference source not found. in his ‘Hole In The Wall’ project</td>
</tr>
<tr>
<td>2000-2001</td>
<td>Mitra’s MIE concept proven to be successful in India (Mitra 2000)</td>
</tr>
<tr>
<td>2002</td>
<td>Digital Doorway project commences in South Africa (GXXXX et al. 2004)</td>
</tr>
<tr>
<td>2002-2004</td>
<td>Similar findings on the success of MIE validated in South Africa</td>
</tr>
<tr>
<td>2003</td>
<td>Migration to fully Open Source software</td>
</tr>
<tr>
<td>2004</td>
<td>4-terminal Digital Doorway housings introduced</td>
</tr>
<tr>
<td>2005</td>
<td>Project expanded to 24 diverse sites around South Africa for comparison purposes</td>
</tr>
<tr>
<td>2006</td>
<td>Xubuntu diskless fat client solution developed</td>
</tr>
<tr>
<td>2007</td>
<td>Project expansion to 100+ three-terminal sites, software refinement, initial MESH network integration prototypes</td>
</tr>
<tr>
<td>2008</td>
<td>Further massification and system refinement</td>
</tr>
<tr>
<td>2009 --&gt;</td>
<td>Greater focus on content and community enabling</td>
</tr>
</tbody>
</table>
Unlike the OLPC\(^1\) project (Negroponte 2006), the Digital Doorway terminals are designed to be social entities where a number of users congregate in one place and participate both in peer learning as well as individual interaction with the system. The design is based around a rugged, centralised hub of activity rather than individually distributed laptops.

System refinement is an iterative process, relying on creative thinking, observation of users and user feedback. At each juncture of the project the following question needs to be asked: “How can we best improve the impact that the Digital Doorway is having on a community?” This question cannot be properly answered without a clear understanding of the community involved and accurate demographic data of users of the Digital Doorways. Thus, a further question is raised that this paper partly addresses, namely: “What interventions can be introduced on the Digital Doorway to help improve collected demographic data?”

Community impact in the context of this project has, as its overarching goal, the provision of effective ICT literacy education to the neediest communities in South Africa. The interventions mentioned here further enhance the impact of each Digital Doorway by preventing the content from appearing stale and unchanging.

This paper highlights some recent progress made towards personalising user experience and providing a mechanism to increase the accuracy of user demographics at each site. The following topics are addressed:

- Digital Doorway system architecture and software
- Personalised and shared user spaces
- A mechanism for users to transfer files from the Digital Doorway to a personal device
- The collection and analysis of demographic data and application usage statistics.

In addition to technical innovations introduced into the project, we provide feedback from actual user interactions with the system.

**SYSTEM ARCHITECTURE AND SOFTWARE**

In order to provide a technology platform that was robust, connected to a network, able to be monitored, resilient to system crashes and power failures and capable of delivering a content-rich, fully multimedia PC environment, a number of technical challenges had to be overcome.

The result is the current configuration, which consists of a 3-terminal setup with a client/fileserver PC and two diskless clients (“fat clients”) connected via an Ethernet switch. The system includes a separate Mindset server (containing 100 Gigabytes of additional educational content), a satellite dish (for downloads only) and GPRS modem backhaul (for status reports and log file uploads to a central server). Each Digital Doorway is represented by the items within the rectangle in Figure 2. The additional connectivity components are also displayed.

The system runs open source software, employing a modified Ubuntu Linux OS with the Xfce display manager and a distributed client system where each client becomes an independent terminal, but content is shared among all terminals. The inherent security of Linux is employed to prevent unauthorised modification of system-critical files. User generated content saved within a home directory is only modifiable by the owner of that directory.

Content on the Digital Doorway is accessed either via icons on the desktop (high visibility) or via a customised Xfce XML menu structure / taskbar menu system (lower visibility). The menu is divided into two chief groupings, namely, Programmes and Resources. The Programmes menu is subdivided into the following categories: Edutainment, Office, Internet, Games, Sounds and Misc. The Resources menu is subdivided into the following categories: Simulations, Wikipedia, Agriculture, AudioBooks, Computers, Crafts, Fun, Heath and Safety, Literature, Realworld and Science. Content comprises binary applications, PDF documents, html, Java, audio and video content.

---

\(^1\) One Laptop Per Child
PERSONALISED AND SHARED USER SPACES

To enable first-time users of the computers to get started quickly, various guest account logins have been created. Although the guest accounts could have been made login-less, the login procedure (entry of a simple user name and password) was maintained to teach users this particular skill. Instructions for logging in are clearly displayed on the login screen itself in multiple languages.

In addition to the guest accounts, users are able to create individual user spaces, with a personal password. This user creation procedure is initiated from the gdm login screen. Users are then presented with a screen where they specify the following characteristics:

- User login and password
- Name
- Age
- Gender
- Home language
- Preferred language

Preferred language refers to the language the user chooses in which to interact with the computer, provided the application supports that language.

The information above is saved to a log file and transmitted back to a central server via GPRS once a day at a set time. The data is then used to help gain an understanding of user demographics at each site.
INTERVENTIONS

In terms of the original research goals, an effective system could be described as one where users with limited or no prior exposure to computers were able to navigate their way around the Digital Doorway and perform basic computer operations after a few weeks. Furthermore, the Digital Doorway would continue to stimulate users in the community in terms of computer literacy development as well as improving information literacy on an ongoing basis.

As the Figures 3 and 4 illustrate, experience has shown that the Digital Doorway site becomes an ineffective system over time, due to two chief reasons: 1) Hardware failure (due to a lack of adequate system maintenance and a lack of community “ownership” of the unit) and 2) Lack of new and stimulating content. On the other hand, where community ownership, proper system maintenance and relevant content updates are present, the site grows in popularity and becomes effective in terms of ICT literacy and uplifting of a community.

In the Digital Doorway project, due to the extremely remote locations of many sites, proper system maintenance is possible only where the status of the system can be ascertained. This is achieved through two mechanisms: 1) a community champion providing feedback to the maintainers, and 2) a technical mechanism for automatic site status reporting. The interventions described below are aimed at improving user content, and increasing a user’s sense of ownership of the system. Thus, crucial factors necessary for a successful site are provided. These, together with good system maintenance, will ensure that the Digital Doorway site meets its goals.
Desktop Personalisation

In order to provide a more enjoyable and personalised user experience, as well as providing an incentive for more users to create their own user spaces (in turn, providing more valuable data for subsequent research), a mechanism was designed for Desktop personalisation. The mechanism employs the built in webcam and works as follows:

Once a user has logged in, he or she selects an icon on the desktop which, in turn, calls two customised scripts: grabddcam and setbackground. The grabddcam script uses the installed webcam application to capture an image of the user and then to resize, add a frame and the user name to the image and merge it with a custom background. The setddbackground script then takes this newly created image and applies it to the current user’s background. In this way, a user’s desktop can be personalised with both his /her name and photograph.

Natural curiosity and exploration are the primary drivers to ICT learning in a minimally invasive educational setting. Content with a high visual appeal and/or catchy music tends to attract the user’s attention and provides the catalyst for long term interaction. Desktop personalisation through the use of the webcam has the goals of providing an incentive for user self-registration through a user-created expressive interface (Preece 2002), as well as developing a sense of ownership between the user and the Digital Doorway unit. This latter aspect helps to ensure that the system is both accepted and looked after by the community.

Besides instructions on how to login as a guest user or to register as a new user, the text on the login page should indicate the benefits of registering a personal user name and logging in with that user name to motivate the entry of demographic data and associated application usage.

A mechanism for allowing users to transfer files from the digital doorway to a personal device

Numerous requests (both via the built in feedback form, and verbal requests) have been made for users to be save their work externally, as well as for printing facilities. Users would be more likely to register and sign in with their own user names if they knew that they could work on documents or other products over
more than one day, initially saving them in their user space on the Digital Doorway and then transferring them to their own device. The remote and often rural locations of most Digital Doorway sites, combined with the fact that the machines remain unattended for most of the time, make it essential that the housing be made as robust as possible. Openings, such as CD-ROM drives, USB ports and the like, had by necessity to be eliminated to prevent users either breaking off the CD tray or sticking objects or chewing gum into the holes.

Currently it is not practical to install a printer at each site. The logistics of managing ink and paper remotely are too complicated. An alternative proposal is to combine each Digital Doorway with a staffed print facility, where a member of the community charges for printing services and uses the cash to sustain the business.

A mechanism has been designed which employs a Bluetooth dongle and customised software to enable users to save their work to a personal device such as a cellphone or PDA. An icon on each user's desktop calls a script which, in turn, starts a Bluetooth object exchange (OBEX) server application and displays a custom graphical user interface (GUI) with which the user can select a file to download. The GUI has purposefully been kept extremely simple in design and easy to use, consisting of just two buttons on the main interface (select file, download file). The specific Bluetooth device needs to be selected from a list of discovered devices, and the transfer is established. The user confirms on the personal device that he/she wishes to accept the file and the file is downloaded. This file can then be either viewed on the personal device, sent to other users or transferred in a similar fashion to a Bluetooth-enabled computer linked to printing facilities where the file can be printed. An example of the application of such a mechanism would be the printing of a particular article by teachers for use in the classroom.

**User-paid GPRS internet surfing**

Analysis of the user feedback obtained via the feedback form mechanism brought a number of requests for Internet access to light ("We want to now about internet "); “I dont feel gaining some skills bcos we don’t surf the internet”; “I would like you to please enable your internet service please ”; “Can u pls conect me to the internet ”; “I need internet browser a.s.p ’; “The digital doorway is de best cus we get 2 c other parts of de world .we only need 2 know on de internet dat how does it work-we tried a lot to open it but it doesnt work-how can we solve dis kind of situation ”; “Activate the internet please ”). (Language usage and spelling of quotations are unchanged).

Due to the high cost of providing Internet access to the Digital Doorways and the observed reduction in the use of the other content on the system when Internet was made available, we did not want the Digital Doorway to become a free Internet café. A compromise was to develop a system where a second GPRS modem is attached to the server and one of the three terminals allows the user to surf the Internet at his/her own expense. Each user has a credit balance for Internet that can be topped up using pre-paid cellular top-up vouchers. After a user's individual Internet credit has expired, surfing is no longer possible.

Guest users can also have their Internet credit topped up; however, this credit becomes available to any user on that terminal who logs in with that guest account.

Through the above mechanism, more users are encouraged to create individual user logins for themselves, potentially increasing the amount of supplied demographic data input during this user login creation procedure.
THE COLLECTION AND ANALYSIS OF DEMOGRAPHIC DATA AND APPLICATION USAGE STATISTICS

In order to be able to analyse user demographics and usage at the sites accurately, a mechanism for collecting this data and transmitting data back to a central server had to be designed. The following sequence of events plays out in this process:

- User chooses between guest account login, own login or new user creation
- If new user creation is chosen, user takes action to create own user space
- User enters details (age, gender etc.) on form
- Details are added to user registration file on DD server
- File is updated once a day to Mindset server
- Applications on DD are launched from menu using a wrapper script (tracker)
- Wrapper script writes start time and application name and arguments to user-specific log file
- Log files are tarred together and copied to Mindset server once a day
- GPRS modem is used to transmit tarred log files and registration files back to central ftp server for download and analysis.

Application usage data are logged for each user. The more users who create their own user logins, the more accurate the demographic and application usage data will be. Consequently it is important to create incentives for users to create and use their own user logins. An incentive is the permanence of files saved in the user’s home directory and the ability to protect the permissions of those files. The customised desktop mentioned earlier provides a second incentive, thereby enriching the computer experience for the user.

SOME RESULTS FROM THE DATA COLLECTION PROCESS

The results below are based on data gathered from 84 sites around South Africa, up to the beginning of January 2008. Given that data is being collected from so many sites, each site consisting of three Digital Doorway terminals and multiple users per terminal, the amount of data is extensive. Numerous social insights and conclusions can be drawn from the analysis of this data. Some broad results have emerged from this process and are highlighted in the ensuing discussion.

Registered users

Sites analysed: 75
Total self-registered users: 3,896
Average number of self-registered users per site: 52

Registered user versus guest user application launches

Registered users: 52,409 recorded application launches
Guest users: 468,433 recorded application launches

Thus, only 11.2% of activity is generated by logged in, self-registered users. Most users have not registered their own user names or they do not log in with nor use them. The reason for this is the added level of complexity and computer skill required to create one's own user name. Logging in as a guest user is a one-step process (entering the guest login as indicated on the screen); logging in as a registered user for the first time is a three-step process (typing “new” to reach the user creation screen, entering details, logging in with the newly created user name and password). No benefits of creating a user name are indicated on the login screen, further encouraging the user merely to log in with a guest account.
If the interventions are successful, this percentage should increase in future. This will be a result to track after the interventions have been in the field for some months.

**Registered user demographics**

**Gender:**

Sixty-seven percent of all registered users indicated their gender; 33% did not. Of this 67% who specified gender, the following results emerged: Male: 75%, Female: 25%

![Figure 6: Proportion of Male Users to Female Users](image)

The overall percentage of female to male users was higher than observed at Mamelodi (Grobler 2004); however, male users are still the clear majority. This could indicate a tendency of male users to be more forceful in their commandeering of the Digital Doorway terminal. A glance at the photograph at the beginning of this paper shows how the system is being used by males; female users remain on the periphery.

Further research questions are raised regarding gender, for example, “Do more boys or girls create and use individual user accounts rather than guest accounts?; How does this relate to behaviour differences between the genders?”; and “What are the differences in the most popular applications accessed between genders?”

**Language:**

On self-registration, users were given the option of specifying their home language and their preferred language of interaction. It is interesting to note that English as preferred language choice showed a markedly higher percentage than home language choice.
Age:

Users could also specify their age. This yielded the following statistics tabulated in Figures 9 and 10.
Although there was a high degree (16.7%) of spoilt or meaningless entries, this is to be expected in an environment such as the Digital Doorway where users are still coming to grips with ICT basics. The high percentage of users in the 10 – 25 age range can be attributed to the following factors:

- a number of Digital Doorway are located in schools and higher education facilities.
- young people are not afraid of new technology
- young people have more spare time to spend at the Digital Doorway
- details of the self-registration process (the source of these results) are shared between youngsters more than adults.

While Mitra's Hole-In-The-Wall project is aimed at children, the Digital Doorway project seeks to enable both children and adults. Nevertheless, the above results show that the greatest impact is still felt by children and young adults; older adults seemingly do not benefit that much from the devices.

**Site information for all users**

**Application Usage:**

The following results are aggregates of both guest user application launches and registered user application launches:

**Table 2: Top 20 applications**

<table>
<thead>
<tr>
<th>Total Hits</th>
<th>App</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>31,342</td>
<td>Thunar file manager</td>
<td>10,078</td>
</tr>
<tr>
<td>30,973</td>
<td>Webcam launcher</td>
<td>8,399</td>
</tr>
<tr>
<td>29,550</td>
<td>DD homepage</td>
<td>7,857</td>
</tr>
<tr>
<td>28,140</td>
<td>Gnibbles game</td>
<td>7,077</td>
</tr>
<tr>
<td>24,964</td>
<td>Ktron game</td>
<td>7,376</td>
</tr>
<tr>
<td>23,538</td>
<td>Fun movie 1</td>
<td>7,367</td>
</tr>
<tr>
<td>17,377</td>
<td>Tuxmaths game</td>
<td>7,236</td>
</tr>
<tr>
<td>16,861</td>
<td>Tuxpaint application</td>
<td>6,922</td>
</tr>
<tr>
<td>16,319</td>
<td>Educational suite</td>
<td>6,783</td>
</tr>
<tr>
<td>15,968</td>
<td>Xfce run application</td>
<td>6,301</td>
</tr>
<tr>
<td>15,675</td>
<td>Gnometris game</td>
<td>6,144</td>
</tr>
<tr>
<td>11,742</td>
<td>Potato guy</td>
<td>6,046</td>
</tr>
<tr>
<td>11,596</td>
<td>Mindset</td>
<td>5,924</td>
</tr>
</tbody>
</table>

Total recorded application launches: 529,484

**SITE CATEGORISATION AND ACTIVITY**

Sites can be grouped into the following four basic categories as depicted in Figure 11:

**Operationally successful** (site is up and running, no hardware failures, GPRS and satellite communicating);

**Socially successful** (much usage of the system, many users self-registered);

**Operationally unsuccessful** (electricity down, site switched off at night, poor GPRS coverage, hardware failure etc.);

**Socially unsuccessful** (site not visited much, users not self registering user names).
It is possible, using the available data, to ascertain the date of the first and the last logged activity, as well as days where activity was recorded, days where no activity was recorded, and days where no records were received. In this last case, typically the machine was switched off, or the GPRS communication was down. Plotting a graph of site uptime, downtime and recorded activity for a number of different sites gives useful overview information for the project maintainers, a sample of which is included in Figure 12.

Sites with a high level of user activity from the day of installation (e.g., Albinigirlshigh) can be regarded as ideal sites. Sites where very little record of either activity or inactivity is recorded after the installation date (e.g., Alice library) are examples of operationally unsuccessful installations. Reasons for the latter are that either the Digital Doorway itself is being switched off / not functioning correctly, or the Mindset server and gprs back haul are not functioning correctly. A site, such as the Bloemfontein fablab (High level of records, but no user activity), is an example of a socially unsuccessful site.

Sites experiencing a high level of operational problems will negatively impact on the social success of the site; however, an operationally successful site will not necessarily lead to a socially successful site. Factors, such as community need, perceived value, initial enthusiasm and community champion involvement, influence the success of a site. Sites that are not socially successful become immediately obvious after comparing the number of logged days with activity to the number of logged days without activity. Where the latter is high compared to the former, the site is not being well used and relocation or intervention should be
The gaps in the graph in Figure 12 also highlight how certain assumptions on the use of the Digital Doorway can negatively affect the technical data acquired. The units were designed to remain powered 24 hours a day, and the log files transmitted at the end of each day (i.e., at midnight) at a certain time. However, champions at certain sites turned off the Digital Doorway each evening and turned it on the next morning. Consequently, the unit was never on in the middle of the night and thus no logs files were sent. A more robust system could transmit the log files the following morning if they were not transmitted the night before.

Is there a correlation between a successful site and the number of self-registered users at that site? Plotting the totals of application launches versus the number of registered users (on a site by site basis) reveals some interesting results (cf. Table 3). Why is there such a variation in number of registered users between one site and the next? Clearly, some popular sites have very few registered self-users. These discrepancies can be explained by two influencing factors: the nature of peer learning and user demographics. Firstly, at sites where the notion of creating own user logins has been understood, this practice is transmitted from one user to the next, resulting in a high self-registration count. At other sites, the advantages of self-user creation have not been grasped and users are content to log in using the guest accounts. Secondly, certain sites (e.g., Ntshongweni) are located at primary schools where young children, although encouraged to use the Digital Doorway by teachers, are less capable of grasping the benefits of creating their own user accounts and are possibly unable to read and understand the on-screen instructions for doing so.

Considering that both a level of ICT awareness and a degree of literacy is required for users to register their own user logins, we can thus state that the number of self-registered users at a site is a reflection of the following: activity levels at that site; the effectiveness of the site in achieving its goal of ICT and information literacy training; and the degree to which peer learning has taken place (where users have learned from each other to create their own user logins).

**Table 3: Application launches versus no. registered users, per site**
THE FUTURE

AUDIO PERSONALISATIONS

Work is currently under way to further personalise the user’s computer experience by providing customised audio feedback to the user after login. A system will identify the user’s name and his/her preferred language and use TTS (text to speech) technology merged with pre-recorded audio clips to welcome the user by name in his/her own language and present him/her with additional introductory information on the Digital Doorway.

LITERACY-DEPENDANT DESKTOP LAYOUTS

A further development will be to allow user responses to direct which applications are brought to prominence on the desktop. Based on the age, language and literacy level of the user, the Digital Doorway could modify desktop icons and menu entries to best cater for the expected needs of that particular user.

IMPROVED NETWORK MONITORING TOOLS

In order to ensure optimum uptime and efficient maintenance of each Digital Doorway, a real-time visual Digital Doorway network monitoring tool is being developed for immediate and accessible reporting of each site in the network, including individual component monitoring and connectivity status.
CONCLUSIONS

Despite the low percentage of self-registered users to guest users and the corresponding limit to the amount of demographic data available for analysis, valuable insights can already be drawn from the collected data. By implementing the interventions mentioned in this paper and thus motivating more users to provide demographic data, we hope to gain further insight into the user base of the Digital Doorway computer terminals. Using this improved quality of data we will be better equipped to provide an ICT service that will improve the lives of some of the neediest people in South Africa and on the continent.
References


http://www.hole-in-the-all.com/docs/Paper01.pdf


