E-publishing in Welsh

a Report for the Welsh Books Council

by the Language Technologies Unit, Canolfan Bedwyr, Bangor University

PRIFYSGOL
BANGOR
UNIVERSITY

Delyth Prys, Gruffudd Prys, Dewi Bryn Jones a David Chan

October 2011

Contents

1. Background

The strategy document *A Living Language: A language for living* noted that *"*Reading, writing, watching and listening to all forms of media in that language [Welsh] are crucial for it to thrive..... But there is potential to reach more readers through electronic and digital media. **We will explore the potential for Welsh language publishers to benefit from e-book technology."** The Welsh Government therefore sees the publication of eBooks as part of its strategy to promote the Welsh language and to enable Welsh speakers to lead holistic lives through it.

In order to promote the eBook market in Wales, the Welsh Book Council intends to offer e-books to the public through its Gwales website, and to encourage booksellers who have commercial websites to do likewise. To this end it has nominated Gardners as its e-book supplier. Gardners will receive files in the EPUB format from publishers so that they may be published as e-books using the Adobe DRM system to ensure that digital rights are protected.

As there are some specific technical problems with supporting Welsh in e-book format, the Welsh Book Council commissioned the Language Technologies Unit at Canolfan Bedwyr, Bangor University, to write this report, which includes recommendations for developing and supporting Welsh eBooks. In a separate document, prepared in cooperation with the Book Council's Design Department, technical guidelines will be published to aid publishers to create EPUB files.

2. Defining an eBook

An eBook is a book produced in a digital format. As with a paper book, the purpose of an eBook is to present the reader with content, mainly in the form of text. As in the case of a paper book, eBooks can include images, both in conjunction with text or even without (although this would be unusual). It is true that eBooks can include additional features which facilitate the reader's experience (for example the ability to jump from the index to a

specific chapter with a click), but these useful additions are minor evolutions rather than revolutions and do not constitute a fundamental difference between the paper format and the digital format[1].

EBooks also follow the traditional template of the book by being presened as a series of sequential pages placed in a specific order. This is in contrast to a website, for example, which exists as a collection of linked files with no specific prescribed reading order.

There is therefore very little apparent difference between a book and an eBook. Once the content itself has been created (be it a novel, biography, collection of poems or other written material), it may be printed on paper in the traditional manner, or produced digitally ready for transfer to an electronic reading device. In reality, publishers may sell the same books in the two media, and the electronic medium may offer them new, additional, distribution channels.

3 .The international eBook market

The rapid increase in the sale of eBooks worldwide in the last two or three years has been a remarkable phenomenon[2]. The eBook market in the United States of America, a market which often leads on new technological trends, increased 167% in June 2011 alone[3]. In 2011, Amazon, one of the largest booksellers in the world, announced that their eBook sales, for the first time ever, were greater than the sales of their printed books.

The eBook market is one where the technology is still evolving in terms of the formats used to digitizing content, its sales and distribution models, and the devices or ereaders used to enable the reading of the digital content. The situation therefore remains comparatively

---

[1]Sometimes, the term 'eBook' is used to refer to multimedia software which may include audio, fideo and text within an interface that emulates the pages of a book.
See for example thoes published in the *Nici a Cris* series by *Acen*:
http://www.acen.co.uk/cy/siop.html?page=shop.browse&category_id=40 .
However, in this document, the term 'eBook' will be used to refer only to refer to digital books that are reasonably strict emulations of medium of the printed book.

[2] See the statistics at http://www.futureBook.net/content/charting-global-e-book-market-exclusive-data
[3] Quoted from http://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/48646-e-book-sales-jump-in-june-print-plunges.html

unstable, with new formats, distribution channels and ereading software and devices being developed all the time.

EBooks may be read on a number of different devices, including ordinary PCs and laptops, smartphones, and new multipurpose devices in tablet form. However, it is likely that one of the reasons for recent increase in the popularity of eBooks has been the arrival of dedicated eBook readers, which, together with tight integration with online stores, make it easy to buy eBooks, download and read them. Lifehacker gives us a detailed picture of the relative popularity of different ereaders:

http://lifehacker.com/5826280/five-best-eBook-readers
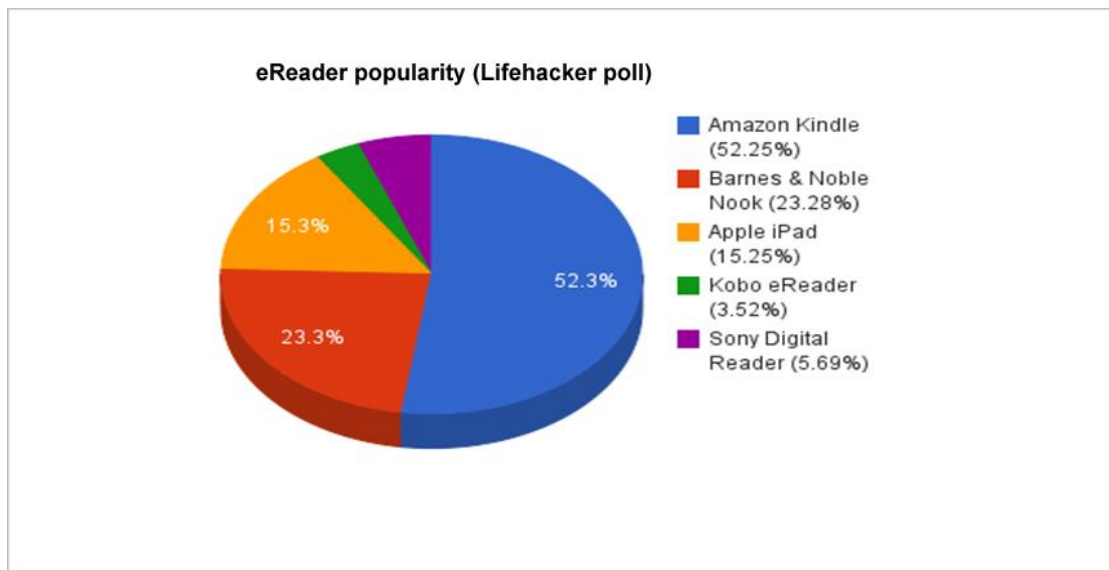
A comprehensive list of various eBook readers may be found on the *mobileread* website:

http://wiki.mobileread.com/wiki/E-book_Reader_Matrix

Although there are so many ereaders available, it is worth noting that apart from the two or three most popular devices the majority of these possess only a small share of the market[4].

4.   Use of eBooks in Wales

In order to understand how the demand for Welsh eBooks and the use of eBook readers in Wales compares with the international picture, a questionnaire was drawn up for this report (see Appendix) asking Welsh readers for their use of eBooks and their attitudes towards them.

A total of 100 people were questioned in the survey, mostly on the Welsh Book Council stand at the National Eisteddfod in Wrexham in August 2011. It was found that 83% wanted to see Welsh books published in electronic form as well as on paper:
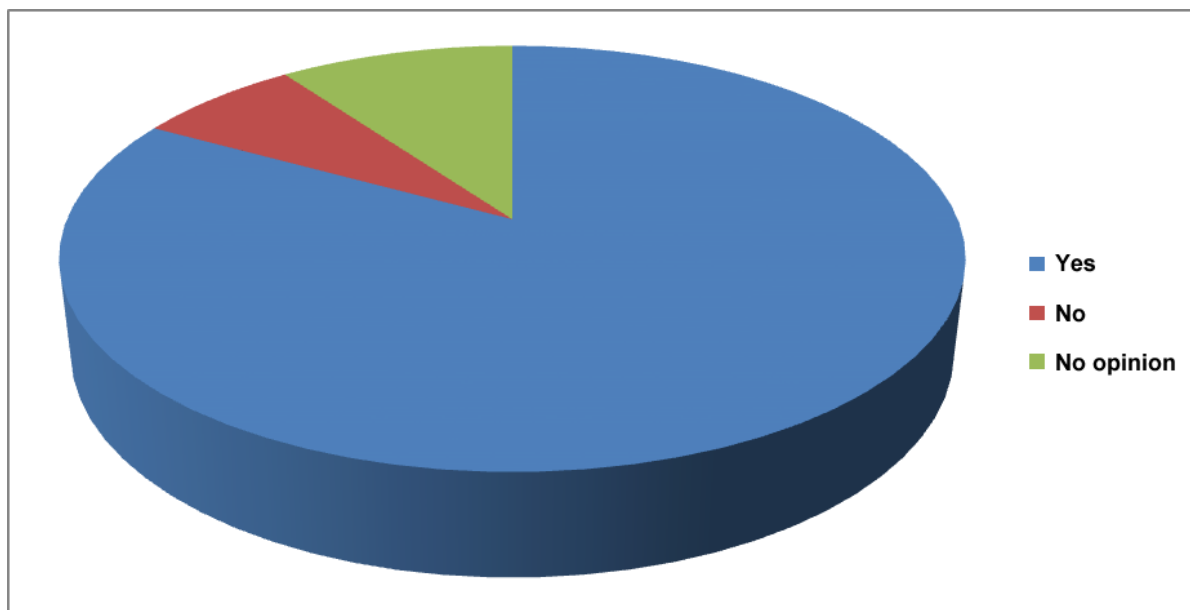
However, only 30% already posessed an eBook reader:

---

[4]It should be noted that there may be a number of different versions of a single device. For example, Amazon's curent Kindle is the 4th generation of the device, and Amazon now sell several current models side by side.
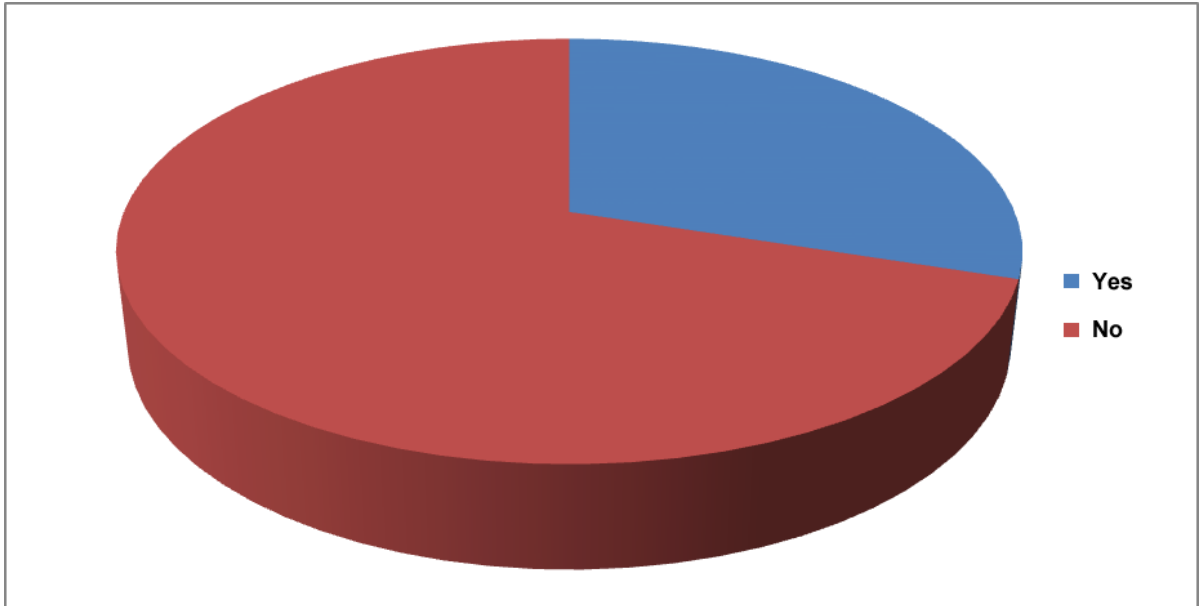
with another 40% likely to buy an ereader in the future, although 28% would only do so if Welsh eBooks were available for download:
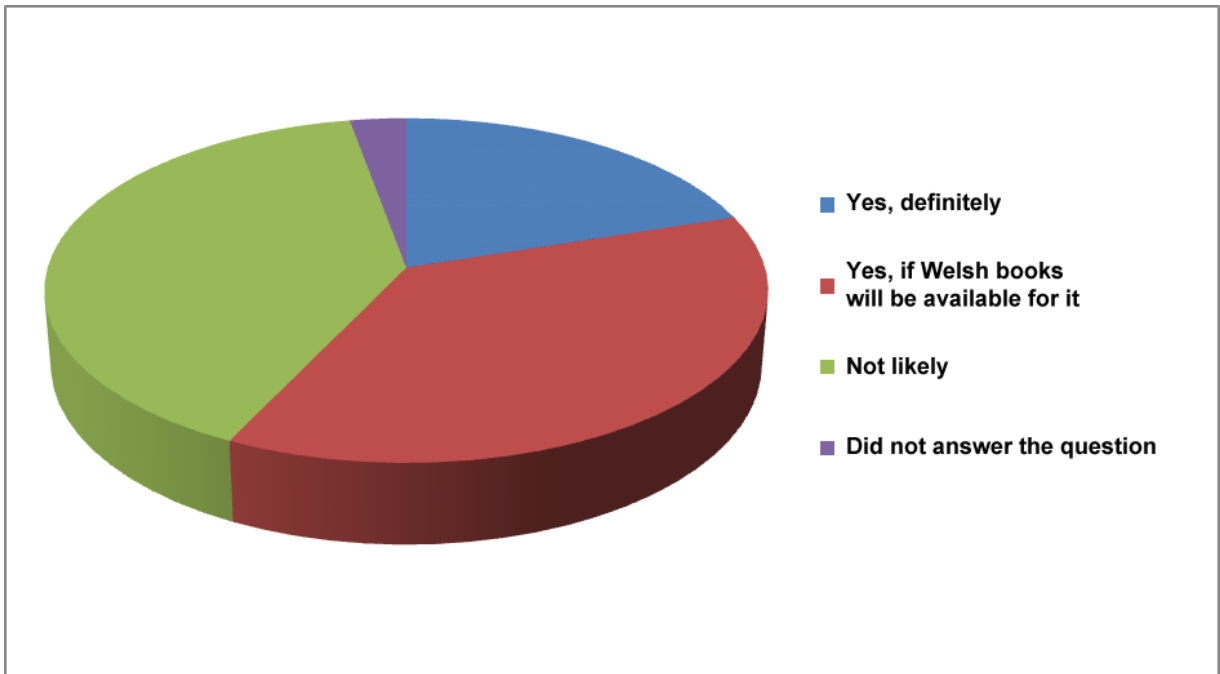
When the 30% who already owned an eBook reader were asked what devices they owned, it confirmed the international picture of the dominance of the Kindle (14%), but multipurpose

devices such as the iPhone (11%), Android phones (6%), and iPad (3%) between them were more popular:
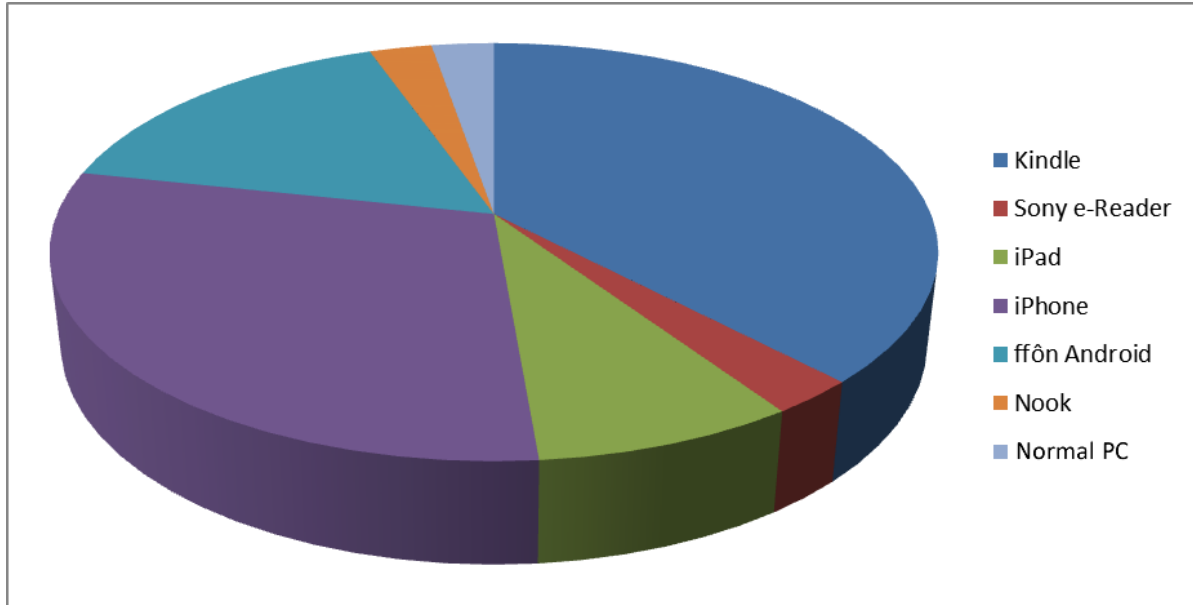
The devices owned by the respondents influence their choice of eBook stores, with Amazon (producers of the Kindle) used by 15, individual booksellers websites used by 5, iBooks used by 7, PDF files used by 11, Gwynedd Library used by 1, the Gutenberg Project used by 1, and Android Market used by 2:
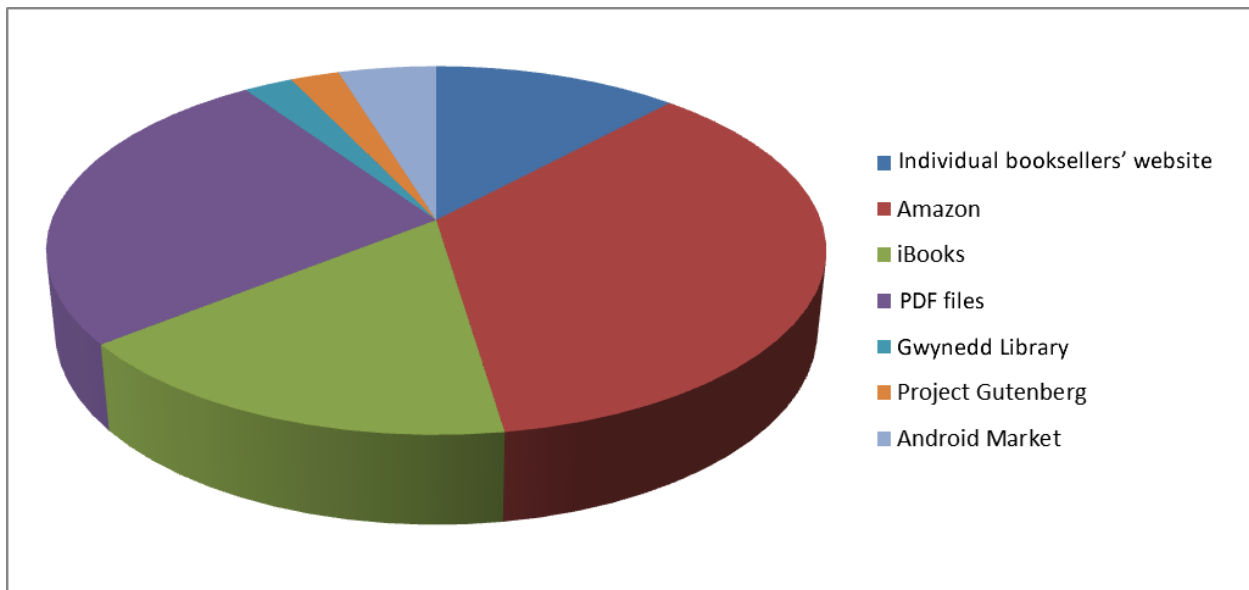
This total is higher than the 30% who reported that they had an eBook reader since many of the respondents had more than one ereader. One respondent had 4 different devices, and another had 3 devices. 10 respondents had 2 different devices, with 5 of the 10 stating that the Kindle was one of these devices. It may be that having more than one device, and expecting to be able to transfer purchased content between different devices (see comment h. below), is a sign of a new trend that may work against the tendency of some booksellers to tie their buyers to a specific eBook reader.

Responders were given the opportunity to provide any comments they wished at the end of the questionnaire, and amongst them were the following comments (here translated from the original Welsh):

i.   It seems likely that this is the way forward – and we need to move with the age!

ii.   I prefer to feel a book in my hand!

iii.   I'm not sure if eBooks are the way forward. As a librarian I worry about the effect this may have on reading ordinary books. Even so I spend most of my time working on computers.

iv.   If developing eBooks promotes sales/how many people read Welsh – all the better.

v.   An eBook reader with text to speech capabilities would be of great benefit to my mother who is partially blind.

vi.   I would like to read magazines and papers on the Kindle also.

vii.   I would like to see Welsh eBooks with English vocabulary included.

viii. The ability to share a book between my different devices is very important to me.

ix. The distribution of Welsh eBooks needs to be facilitated as a matter of urgency!

x. There are many Welsh books that are not likely to be republished – I would be so good to have classics, editions of poetical works etc. as eBooks.

xi. Looking forward to Welsh eBooks and epublications.

xii. I would prefer not to use Amazon! I only read English books on the Kindle at the present time. I'm not even aware of what provision there is in Welsh.

xiii. Remember the blind!

xiv. Any medium that may promote children's interest and skills in Welsh is always a help – on with the developments.

xv. Although I have an iPhone I don't buy eBooks.

xvi. Need for children's books also.

xvii. Need to have Welsh books as eBooks.

These comments show a great deal of desire for the provision of Welsh eBooks, alongside some understandable hesitation about the new technology.

There are also some valuable ideas here about using the new technology to republish Welsh classics, the technology's potential to deliver audio books for visually impaired people, facilitating the inclusion of English vocabulary to help learners, and the importance of the

provision of eBooks for children in addition to adults.

5. Dedicated ereaders

Whilst ereading software comes built-in to dedicated ereaders such as the Kindle and Sony e-Reader, it is also possible to install ereading software on other hardware devices such as computers and smartphones to turn them into ereading devices. There is a wide variety of eBook reading software that can be downloaded and installed onto these devices. Many of these devices now come with ereading software already installed, but users are usually free to choose to install alternative ereading software if they so wish.

Dedicated ereaders on the market vary in terms of size, screen size and resolution, and the latest generation have begun to adopt touch screen technology rather than using buttons to operate them.

Most of these dedicated ereading devices use an eink screen as this type of screen offers many advantages for an ereader. Eink screens are of a very high resolution, much higher than a computer screen, with many more smaller dots being used to create every individual letter. This means that text on an eBook screen appears much clearer than that on an LCD monitor, and more similar in terms of reading experience to reading a paper book. Another advantage of an e-ink screen is that it does not make heavy demands on the device's battery. The battery is used to 'turn' the page rather than continupusly power the device, and its life is therefeore measured according to the number of page turnings. A typical eink screen may last more than 3,000 page turns. Despite this saving in terms of energy used, an eBook screen is much brighter than comparable LCD screens. This, together with the matt (rather than shiny) nature of the screen means that it is much more suitable for reading in strong light (e.g. on a beach) than an LCD screen.

Some of the more innovative models of dedicated ereading devices have the ability to connect to the internet over Wi-Fi or 3G in order to download eBooks from the web. The

procedure with less sophisticated models is to connect the device to a computer by means of a cable and use it to move any eBooks that have been downloaded to the computer over to the ereader.

E-ink screens have some disadvantages. An e-ink screen cannot show moving images successfully, and so is unsuitable for use with multipurpose devices which are also intended to show videos, play games etc. Another disadvantage is that colour e-ink screens are not yet available at a reasonable price, although their price is likely to fall in future. Nor can these colour e-ink screens at present convey the true vividness of the colours. However, this is a fast-developing field, and it is difficult to foretell what technological advances may be acheived even in the comparatively short term.

An example of a prototype ereader with colour screen may be seen at:
http://www.engadget.com/photos/hanvon-e920-color-e-reader-hands-on/#3765207

6. Multipurpose ereaders

Multipurpose ereaders are designed to do more than display eBooks. These devices include smart phones such as the iPhone and the numerous brands that run on the Android system, as well as traditional laptops and computers running Windows, OS X or a distribution of Linux. It is worth noting that an Android smartphone may currently be bought on a pay as you go contract for less than £30, and that ereading software can be installed on the device for free. Ereading devices are therefore well within reach of all the book buying public.

Multipurpose devicese such as these do not tend to have an e-ink screen. Instead, they mostly use a full colour LCD, which is increasingly also a touch screen so that users may control what is on the screen with their fingers. A large variety of sizes and screen resolutions are available, even within the same class of phones (the resolution of the iPhone 4 for example is higher than the iPhone 3, and Android running phones come in many different shapes and screen resolutions).

As with some of the dedicated ereaders, most multipurpose devices can now connect to the web through Wi-Fi, 3G and/or network cable, and so have direct access to online shops selling eBooks.

Because of the multimedia nature of these devices, installed eBooks can use links which refer directly to multimedia content available on the web, from online dictionaries to audio files or videos of interviews with authors on websites such as YouTube.

Below is a table listing the various devices and their operating systems:

Device	Type	OS	Touch screen
iPhone	Smartphone	iOS	Y
Android Phones	Smartphone	Android	N
Windows Phones	Smartphone	Windows	
Nokia Phones	Smartphone	Symbian*	
iPad	Tablet	iOS	Y
Android Tablets	Tablet	Android**	Y
Computer	PC	Win/OS X/Linux***	N
Laptop	PC	Win/OS X/Linux***	N

*Nokia has now signed a contract which sees them abandoning Symbian and developing on Microsoft's Windows Phone platform.

**Many versions of Android now exist, and the Kindle Fire is based on Amazon's Android fork.

***Calibre software, for example, may be installed on computers running Windows, Mac OS X or Linux.

An important difference between multipurpose devices and dedicated ereaders is that the flexibility of multipurpose devices allows them to choose from amongst a number of different

software programs. They are therefore less likely to be tied to one specific eBook format. More than one ereader software programs may be installed on a device, and it is also possible to download the same book on more than one device.

In producing eBooks for the market, it's important to test them on as many devices as possible, using as many ereading software programs as possible in order to ensure that they display correctly despite minor differences between one device and one program and another.

7. Software for reading eBooks

On most multipurpose devices, a choice of software programs is available for reading eBooks. This means that users are free to choose their preferred program. Even so, if a device comes with ereading software already installed (as the iPhone comes with iBooks already installed), many users may not bother to install different software, or may not sufficiently understand the technology to do so without encouragement and guidance.

Most ereading software programs are available free of charge. Usually their developers hope to generate income by directing users to their own or affiliated eBook shops and earn revenue from eBook sales at that shop.

Here is a list of popular ereading software:

Adobe Digital Editions (PC and OS X)
Kindle
Kobo
Aldiko
Stanza
iBooks
Calibre

Google Books

Because of the nature of the different formats and the use of software to manage digital rights, there is no one ereading program that can open every type of eBook.

Information on supported formats in the comprehensive list of ereading software may be found here:

http://wiki.mobileread.com/wiki/E-book_software

8. Supporting the publishing industry

Preparing texts for publishing in the form of eBooks places new demands on the Welsh publishing industry. Although the technology itself is not expensive, (i.e. publishers don't need new equipment if they already have access to a recent computer and software), it requires technical knowledge and know how of a different kind to that required to produce paper-based publications. Because of the various different platforms and standards, and their continuous evolution, it is no easy matter to find one conversion methodology which will work in every case. Even when a piece of software claims to follow a certain standard, its interpretation of some aspects of the standard may cause problems, and difficulties may arise even between different versions of the same software.

The specific requirements of Welsh, including the splitting of words in the right places for hyphenation, and the correct way to deal with accents and special characters (especially ŵ and ŷ) are further issues that need to be resolved when offering support for Welsh publishers. Choosing the EPUB format, therefore, and concentrating on it to begin with, simplifies matters. It should be noted however that there are differences between the various versions of EPUB, as well as differences in its interpretation by different ereading software packages and devices.

In order therefore to help the industry, easy-to-use guidelines are published separately from to this report. The intention is to aid publishers prepare content files for epublishing in EPUB format. It is also recommended that an accompanying training program be established for Welsh publishers to aid their use and understanding of epublishing. A potential source of funding may be the Welsh Government and ESF's Work-based Learning programme.

The other recommendation, if a suitable source of funding may be found, is that a specific software program be created to facilitate publication in Welsh in the EPUB format. This could be done in conjunction with the Welsh Book Council, to be placed on a free website for book publishers to use, on the pattern of *Cysill ar-lein*. This would facilitate the uploading of book texts as Word files, with an accompanying form to be filled, providing metadata details in the correct format, the production of a report on possible errors in the text, and an export of the text in EPUB format ready for publication as an eBook.

9. The way forward

It is clear from the international surveys, and the questionnaire responses in Wales, that eBooks are rapidly increasing in popularity. Good quality eBook provision in Welsh is therefore a matter of urgency, not only in order to provide the same practical equal opportunities for readers of Welsh and English, but also to strengthen the image of Welsh as an attractive, contemporary language.

This is a fast developing field, and neither the technology nor the selling and distribution models have yet reached stability. In such a situation, it is important to try and provide for as many eBook readers and selling and distribution opportunities as possible, rather than limit the provision to only one e-publishing device. The supremacy of 'closed gardens' such as Kindle Amazon, makes it more difficult for the Welsh book market to compete on equal terms, but the EPUB open format and willingness to distribute eBooks though a number of different channels (the website of the Welsh Book Council, websites of individual booksellers, websites of large international sellers) mean that Welsh language readers will have a number of options to entice them to read more in Welsh.

New opportunities may also arise for the Welsh language market through the expansion of the eBook market. Self-publishing may become attractive to some, cutting out the commercial publisher. However Welsh book publishers offer editing, proofreading and marketing services to their authors, as well as the printing of paper books. These services may gain additional importance as the eBook market develops, possibly enabling more books to be published as the cost of printing and distributing paper copies becomes less important.

Another topic worthy of consideration is the epublishing of older books now out of print (see the response in 3.x above). Old books that are out of copyright and distributed free of charge on some of the ereaders are very popular in English. Such provision would be very attractive to readers of Welsh eBooks, and establishing such a scheme, possibly in conjunction with the National Library and in partnership with their digitization programme, would attract readers to visit, for example, the Gwales website, if free eBooks were placed there side by side with modern books for sale.

Publishers also have books which are still within copyright but out of print. Very often printing costs militate against republishing them, but once again their publication as eBooks would attract new readers and offer a valuable service. These books include a number of classics, as do the books that are out of copyright. Publishing them as eBooks would be a valuable service that would prove attractive to many.

Accessibility for visually impaired people is an additional issue, once again brought up in the questionnaire responses. Text-to-speech technology is not detailed in this report, but Bangor University has developed this technology for Welsh, and it is now available from a number of commercial sellers to be incorporated in eBooks, together with dictionaries for Welsh learners, and similar resources to enrich the experience of ereading Welsh to a wider audience.

It is important therefore to view the arrival of eBooks as a golden opportunity to widen the provision of materials in Welsh, and not as a threat. Providing a technical service and advice to publishers to correctly prepare Welsh books for epublishing is an important and timely priority.

Appendix: August 2011 Questionnaire [Note: this translation for information purposes, as the original questionnaire was delivered only in Welsh].

Welsh e-books Questionnaire

The Welsh Book Council has commissioned the Language Technologies Unit at Canolfan Bedwyr to write a report on publishing Welsh e-books. We would be grateful if you could help us by answering the questions below. Many thanks!

1. Would you like to see Welsh books published in electronic form as well as on paper?

    ☐ Yes   ☐ No   ☐ No opinion
    _____

2. Do you have an eBook reader at present? (If the answer is no, go to question 5).

    ☐ Yes   ☐ No
    _____

3. If yes, which one is it?

    ☐ Kindle   ☐ Sony e-Reader   ☐ iPad   ☐ iPhone
    ☐ Android phone   ☐ tablet PC   ☐ Other (state what) ....
    _____

4. If you buy eBooks, what shops/software do you use (tick every one that applies)

    ☐ individual booksellers websites   ☐ Kindle Amazon   ☐ iBooks
    ☐ PDF files   ☐ Other (state what) ....
    _____

5. If you do not have an eBook reader, are you likely to buy such a device in the future?

    ☐ Yes definitely   ☐ Yes, if Welsh books are available for it
    ☐ Not likely
    _____

6. Any other comments you would like to make

    ....,