

# Information Technology and the Welsh Language: A Strategy Document





# Information Technology and the Welsh Language

## A STRATEGY DOCUMENT



V 1.0 - 2006

# 1 Contents

1	Contents .....	4
2	Glossary of Terms .....	6
3	Foreword.....	7
4	Why is this strategy needed? .....	8
4.1	<i>Licensing and distribution of products we commission</i> .....	9
4.2	<i>Timetable and current actions</i> .....	9
5	How to contribute to further drafts of this strategy.....	10
6	The Way Forward for Welsh in IT.....	11
6.1	<i>IT Terminology &amp; Lexicographical resources</i> .....	11
6.2	<i>Localisation</i> .....	12
6.2.1	Localisation of Open Source Applications.....	12
6.2.2	Localisation of other applications .....	13
6.2.3	Switchability of interface and language attributes .....	14
6.2.4	Bilingual Website Design.....	14
6.3	<i>Content Creation Aids</i> .....	15
6.3.1	Spellcheckers, Grammar Checkers, and Computerised Dictionaries .....	15
6.3.2	Welsh Language Keyboards, diacritic marks & fonts .....	16
6.3.3	Diacritic Marks and Web pages in Welsh.....	17
6.4	<i>Translation and terminology: how can IT help Welsh?</i> .....	17
6.4.1	Computer Aided Translation.....	17
6.4.2	Terminology Management.....	18
6.4.3	Alignment technology: the creation of parallel corpora of English-Welsh for Translation Memories .....	19
6.4.4	Translation workflow and Bilingual Document Management .....	19
6.4.5	Translation workflow and simultaneous auto updating of multilingual Content/Document Management Systems.....	19
6.4.6	Machine Translation .....	20
6.5	<i>Language technology issues and corpora</i> .....	21
6.6	<i>Speech Technology</i> .....	21
6.6.1	Speech Synthesis.....	22
6.6.2	Speech Recognition.....	22
6.6.3	Adaptive technology .....	22
6.6.4	MT, SR, TTS (Integration of Machine Translation and Speech Technology).....	23
6.6.5	Embedding of speech technology into other applications .....	23
6.7	<i>Capacity building</i> .....	23
6.7.1	End Users: IT Training in Welsh.....	24
6.7.2	Developers: enlarging the base of language technology specialists.....	24
6.8	<i>E-learning &amp; The Welsh Language</i> .....	25
6.8.1	Welsh Learners, <i>Gloywi Iaith</i> & E-learning.....	25
6.9	<i>Call centre software</i> .....	25
6.10	<i>Back end databases and bilingualism</i> .....	26
6.11	<i>Metadata</i> .....	26
6.12	<i>Mobile devices</i> .....	27

6.13	<i>Embedded software</i> .....	28
6.14	<i>Optical character recognition and Handwriting recognition</i> .....	28
6.15	<i>Sharing International Experiences in the field of IT</i> .....	28
6.16	<i>Future-proofing for English-Welsh bilingualism in IT</i> .....	29
7	Further reading and selected web references .....	30
7.1.1	The Welsh Language Board Website .....	30
7.1.2	International Standards .....	30
7.1.3	Galicia .....	31
7.1.4	Catalonia .....	34
8	Summary of targets and policy statements .....	36
9	List of respondents to the consultations .....	40

## 2 Glossary of Terms

CAT	Computer-assisted translation
CMS	Content management system
EBMT	Example-based machine translation
ESRC	Economic and Social Research Council
GPS	Global Positioning System
IPSV	Integrated public service vocabulary
IT	Information technology
LIP	Language interface pack
LT	Language Technology
MT	Machine Translation
OCR	Optical Character Recognition
PDA	Personal Digital Assistant
RML	Regional or Minority Language
SAPI	Speech Application Programming Interface
SL	Source language
SMS	Short Message Service
SR	Speech Recognition
TL	Target language
TM	Translation memory
TTS	Text to speech
WISPR	Welsh and Irish Speech Resources
XPSP2	Windows XP Service Pack 2

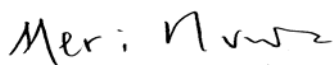
### 3 Foreword

The Welsh Assembly Government's strategic policy document for the Welsh Language, *laith Pawb*, notes the aspiration to create a "truly bilingual Wales [...] where people can choose to live their lives through the medium of either or both Welsh or English and where the presence of the two languages is a source of pride and strength to us all." The need for an IT Strategy for the Welsh language, called for in *laith Pawb*, led the Board to open an initial public consultation during 2003 and a second consultation in 2005. Our thanks go to those who responded to those consultations; this strategy includes the original themes along with further submissions and research.

In relatively few years, our lives have come to be influenced directly, and indirectly, by Information Technology. In order to realise the Government's bold vision of a bilingual Wales, Welsh must have its rightful place in technology, hence the need for a strategic, long-term approach. This document, then, aspires to the normalisation of the Welsh language in the world of Information Technology.

In addition to the input of stakeholders via the consultation process, this strategy is intended to be a live document. Information technology is constantly evolving, and we envisage that the document will be updated at regular intervals. We aspire that the strategy will result in a series of concrete actions over the coming years, and that it will contribute to the growing public and private status, and use, of our language.

Although drafted by the Welsh Language Board the strategy's remit extends well beyond the duties and responsibilities of that institution alone. We must all, as individuals and as institutions, take responsibility for the future of Welsh – in every sphere.



**MERI HUWS**

Chair

Welsh Language Board

## 4 Why is this strategy needed?

Improving the situation of the Welsh language in information technology will provide a boost to the language's status, one of the many prerequisites noted by language planners for a sustainable future for a language. *Iaith Pawb* emphasises the importance of mainstreaming consideration of the Welsh language into all areas of policy development and service delivery. In particular, the document identified the need for a strategy to mainstream the Welsh language in the field of IT.

Developing coherent strategy for this policy area is essential, given the central part information technology now plays in the everyday lives of citizens: behind the scenes of organisations providing services to the public, via websites, appliances of all types in the home, the workplace and beyond. This IT strategy and the accompanying standards document, available on our website, aim to provide a comprehensive policy statement and list of what needs to happen to ensure the language is mainstreamed into all relevant areas of IT and keeps pace with technological developments.

The Board, by its very nature, takes a keen interest in the promotion of Welsh in all fields. Although this document has been drafted by the Welsh Language Board after public consultation, the Board will be happy to commission, facilitate and coordinate IT materials to the best extent of its abilities and resources, but should not be looked upon as a provider for every element mentioned in this strategy document.

We have also examined steps taken and mistakes made in other countries with similar sociolinguistic situations to Wales. The Board will avail itself of all appropriate opportunities for trans-national cooperation in this field, and to build on existing applications. Our membership of the European Network of Language Boards and the Regional or Minority Language Working Group of the British Irish Council (see section 6.15 below) has taught us that in the field of promotion of regional or lesser-used languages it is not always necessary to reinvent the wheel. Welsh can also, evidently, benefit from research done in major language technology.

There may be cost implications to many of the elements of this strategy. In those cases where the Board is identified as the appropriate provider or funding agency, we will tailor our action plans accordingly. Where it is more appropriate for others to be responsible, we will draw their attention to this and, where appropriate, engage in dialogue with them. Where other providers are referred to, it should be noted that all third party registered trademarks are the property of their respective owners.

One of the main ways to put this philosophy to work is via wide, free dissemination of Welsh language IT resources. The first policy statement of this strategy therefore deals with this (targets and policy statements are grouped together at the end of each relevant section and reproduced as a list at the end of this document).

#### **4.1 Licensing and distribution of products we commission**

In the process of achieving these aims, wherever the Board gives grant in aid for the creation of IT products, or commissions IT products, we will strongly encourage or require adherence to international and open standards. We will also ensure that, where appropriate and feasible, products and component parts are made available under a suitable free non-commercial license agreement, subject to Crown Copyright (where applicable) or (where part of a commercial product which offers multilingual functionality) at no additional cost to the end user.

#### **Targets/Policy Statements**

- In commissioning or giving grant in aid to IT projects, we will, where appropriate, encourage or require adherence to international and open IT standards
- In commissioning or giving grant in aid to IT projects, we will ensure that the products, and where appropriate and feasible, component parts, created will be made available, subject to Crown Copyright (where applicable), under a suitable free, non-commercial licence agreement or (where part of a commercial product which offers multilingual functionality) at no additional cost to the end user.

#### **4.2 Timetable and current actions**

An implementation plan will be prepared to put this strategy into action. Regular progress updates will be drawn up in order to report on the targets of the plan.

## 5 How to contribute to further drafts of this strategy

**A**s already noted, in addition to its initial consultations, this document will be revised at regular intervals. We therefore invite individuals and/or organisations from the public, private, and voluntary sectors to send in constructive suggestions to improve this document

Please send your comments to the address below. Unless otherwise requested, comments on the strategy will be available, on request, for public viewing.

**DR JEREMY EVAS**

*Leader, Research, Language Technology & Grants Unit*

[Jeremy.evas@welsh-language-board.org.uk](mailto:Jeremy.evas@welsh-language-board.org.uk)

Welsh Language Board,  
Market Chambers,  
5-7 St Mary St,  
Cardiff,  
CF10 2AT.

<http://www.welsh-language-board.org.uk>

## 6 The Way Forward for Welsh in IT

This document outlines those steps and policies we believe necessary to drive forward the situation of the Welsh Language in Information Technology. It will be made available on our website. Comments on it will be welcome at any time.

In any document, consistency of the use of terminology is of paramount importance. We therefore begin our needs analysis with the field of terminology to describe the field of Information Technology itself.

### 6.1 IT Terminology & Lexicographical resources

We recognise the field of standardisation of terminology and lexicographical resources as one of our priorities for the normalisation of Welsh. The Welsh Language Board has a Corpus Planning Unit which is dedicated to this work. Whilst it is not within the remit of the present document to describe that work in detail, it should be noted that a standardised dictionary of Information Technology terms is already available in order that Welsh speakers be able to discuss this domain. This dictionary is available via the Board's website, as part of the Board's national database of standardised terms.<sup>1</sup> This database will collate all previously published and term lists we have commissioned and present them via an easy to use interface.

The database is accessible free of charge, to all. It should nevertheless be noted that this strategy document *does* deal with IT facilities for managing lists of terminology and the standardised use thereof in section 6.4.2 below.

#### Targets/Policy Statements

- Term lists and gazetteers we commission/fund will be available in a computerised format that follows internationally recognised standards (e.g. ISO).
- Term lists and gazetteers we commission/fund will be able to be adapted for import into the terminology management systems described in section 6.4.2.
- We will examine the possibility of making our term lists available, free of charge, as internet updates to other relevant electronic dictionaries.
- We will examine the possibility, subject to strict standardisation quality control procedures, of a 'suggest a term' facility to our national database of standardised terms.

Correct, standardised Terminology in the field of IT is a precursor for any cogent discussion of this domain in Welsh, and necessary for the next element we shall deal with: localisation.

---

<sup>1</sup> <http://www.e-gymraeg.co.uk/bwrdd-yr-iaith/termau/Default.aspx>

## 6.2 Localisation

Localisation is not the same as translation, although translation does form a substantial part of it. At a general level, it can involve addressing issues such as:

- ensuring date and time formats are appropriate for the user's culture;
- ensuring data and lists are sorted using the appropriate alphabet;
- ensuring number formats are correct for the current culture – is ten Euros and fifty cents “€10.50”, “€10,50” or “10,50 €”?
- ensuring the user interface makes sense and is flexible enough to support a specific language;
- ensuring data storage can accurately store data in different scripts and character encodings;
- translation of text – labels, help files, alerts, etc – into a specific language

The standards document gives detailed, technical advice on localisation which it is not appropriate to publish here. The document is available on the Board's website.<sup>2</sup>

### Targets/Policy Statements

- We will encourage all institutions to offer software/IT provision on a basis of equality between English and Welsh.
- For those organisations with a statutory language scheme under the Welsh Language Act 1993, and who operate software whose interface is available in Welsh, we will, where reasonable and practicable, require them to offer both English and Welsh Language versions on a basis of equality.
- We will continue, where appropriate, to issue circulars regarding mainstreaming bilingual considerations into IT
- We will cooperate with and provide locale and language data to institutions who wish to embed the Welsh language in their applications.
- We will discuss with and provide advice to professional societies and institutions in the field of IT in order to publicise the needs and standards for bilingual IT provision in Wales. Where we consider necessary and/or appropriate, we will organise training and dissemination activities for the members of those organisations. In institutions that have a statutory language scheme, we will do this in partnership with the official responsible for that scheme.
- In this dialogue, we will emphasise the need for bilingual provision to be mainstreamed into all areas of policy and performance indicators
- We will circulate the associated standards document and organise suitable training/publicity events for Board staff and language scheme officers

### 6.2.1 Localisation of Open Source Applications

Welsh has, for some time, benefited from a dedicated virtual community of volunteers who have given freely of their own time to translate many computer applications into Welsh - see, amongst others, the following:

---

<sup>2</sup> <http://www.welsh-language-board.org.uk/en/cynnwys.php?cid=6&pid=109&search=2&nclD=126&nclD2=151&snclD=151>

- the kyfieithu website, which collects together many Welsh language open source resources<sup>3</sup>
- Meddal - An extensive catalogue of Welsh language open source software. Meddal has released a CD of quality controlled Welsh Language Open Source applications).<sup>4</sup>
- Cymrux, a group which has produced a Welsh language version of the Linux Open Source operating system.<sup>5</sup>
- A voluntary translation of GNOME, one of the main Linux desktop applications has also been available since 2002

We are grateful for the dedication and commitment of these volunteers, and the terminologically pioneering work which they undertook. As an example, one of the largest single voluntary localisations of software into Welsh was that of OpenOffice,<sup>6</sup> an open-source<sup>7</sup> suite of Office applications, undertaken by a virtual group of volunteers. From the 2003-4 financial year, the Board has given a £35,000, grant spread over three-years, to the Mercator Centre of the University of Wales, Aberystwyth, to promote the use of OpenOffice, including the translation of all help files and manuals, and the creation of a version of the suite which is switchable between Welsh and English. The grant will also enable the “Agored” scheme to trial the suite in a number of institutions.

### 6.2.2 Localisation of other applications

We are also eager for software companies to localise software applications into Welsh. Another of the significant localisations undertaken thus far is the Microsoft Language Interface Pack programme<sup>8</sup>, in partnership with the Welsh Language Board. This programme, launched in Wales in December 2004, and funded in entirety by Microsoft, enables users, free of charge, to use Welsh Language Interface Packs for Microsoft Office 2003 and Windows XP. Welsh forms only one part of this worldwide, multi language programme. As in the case of the Open Source localisation detailed in section 6.2.1 above, the programme has also led to skills development in the localisation industry in Wales, from the point of creation and use of standardised terminology in the field, and capacity building of all those involved in the project (Cymen translation company and Draig Technologies were the main implementation partners).<sup>9</sup> As with all such localisations, the attendant high profile publicity has provided a significant boost for the image of Welsh.

### Targets/Policy Statements

- We will consider requests for help in localisation from the open source community
- We will engage in dialogue with software producers to explore possible synergies for localising other products, applications and suites into Welsh
- We will contact retailers of home and office PC's to discuss the possibility of offering Welsh language versions of available software as a preinstalled option or as an option in 'self build' PC internet sites.

<sup>3</sup> <http://www.kyfieithu.co.uk>

<sup>4</sup> <http://www.meddal.com>

<sup>5</sup> <http://www.cymrux.co.uk/>

<sup>6</sup> <http://www.openoffice.org/>

<sup>7</sup> See <http://www.opensource.org/docs/definition.php> here for general information regarding open source programs.

<sup>8</sup> <http://www.microsoft.com/industry/government/LocalLanguage.msp>

<sup>9</sup> <http://www.cymen.co.uk> , <http://www.draig.co.uk>



### 6.2.3 Switchability of interface and language attributes

Many Welsh speakers live in bilingual households (it is significant to note that only 21% of 3-15 year old Welsh speakers, and 45% of all Welsh speakers live in entirely Welsh-speaking households<sup>10</sup>), and work in bilingual or predominantly English language settings. It is therefore of the utmost importance that the interface and language attributes of a given system are switchable between English and Welsh in order for people with different language abilities to be able to share the same machines or applications.

It is also useful for language preference to be linked to user profiles, in order, for example, for a non-Welsh speaking parent to be able to share a home PC with his/her Welsh-speaking child.

In Microsoft Office 2003, KDE and GNOME (and in the Welsh language Board sponsored Welsh language version of OpenOffice), the interface language can be easily changed.

Microsoft has also released a supplementary language control console which enables users to alter their language interface, keyboard and locale settings.<sup>11</sup>

Moreover, it is possible, for example, on Linux based system to simultaneously run a screen-based calculator with a French interface, together with a text editor with a German interface with all other screen items in Welsh. While such a simultaneous multilingual situation may be rare, the need for easy switching between Welsh and English is an everyday occurrence, and of paramount importance for a bilingual society.

#### Targets/Policy Statements

- In all our product development discussions, calls for tender and dealings with external organisations, we will emphasise the need for easy switchability between Welsh and English.

### 6.2.4 Bilingual Website Design

The standards document gives detailed technical advice on creation of bilingual websites and other applications. As part of its monitoring programme of statutory language schemes under the Welsh Language Act 1993<sup>12</sup>, the Board has commissioned two independent surveys of the websites of public sector organisations with a statutory language scheme. The first in 2001,<sup>13</sup> was a baseline analysis, which was repeated in the 2003 survey<sup>14</sup>. These surveys informed us in specifying the Standards Document which accompanies this strategy document. It gives detailed advice on the creation of bilingual applications and websites. We will disseminate this document widely and use its standards as the basis for all forthcoming snapshot surveys of websites.

<sup>10</sup> Source: 2001 Census, standard table S143 and commissioned table C0055.

<sup>11</sup> <http://www.microsoft.com/downloads/details.aspx?FamilyID=CC37E08C-1019-4AE8-A258-0D5412876DEC&displaylang=en>.

<sup>12</sup> [http://www.opsi.gov.uk/acts/acts1993/Ukpga\\_19930038\\_en\\_1.htm](http://www.opsi.gov.uk/acts/acts1993/Ukpga_19930038_en_1.htm)

<sup>13</sup> <http://www.welsh-language-board.org.uk/en/cynnwys.php?clD=6&plD=109&nID=62>, together with a detailed list of sites surveyed <http://www.welsh-language-board.org.uk/en/cynnwys.php?clD=6&plD=109&nID=79>

<sup>14</sup> <http://www.welsh-language-board.org.uk/en/cynnwys.php?clD=6&plD=109&nID=214> together with a detailed list of sites surveyed <http://www.welsh-language-board.org.uk/en/cynnwys.php?clD=6&plD=109&nID=215>



## Target/Policy Statement

- We will continue to carry out occasional snapshot surveys of public sector websites as part of the monitoring process of statutory language schemes, using the Standards Document as a basis
- We will examine possibilities for ensuring standard forms of compliance in this field
- We will examine the possibilities for an accreditation scheme for compliance with the standards document
- We will publish guidance (and examine the possibilities for creation of automatic tools) on how to enable a non platform specific browser language locale setting which will ensure that readers automatically receive Welsh language versions of pages where they are available.

## 6.3 Content Creation Aids

According to many surveys, substantial proportions of Welsh speakers rate their confidence levels in their Welsh language skills lower than their skills in English. Therefore, one of our initial priorities in the field of Information Technology was the facilitation of content creation in Welsh. The main thrust was on spellcheckers and grammar checkers which would contribute to increasing confidence, and dictionaries which would enable users to find quick and easy translations of individual words, either in Welsh, or English.

### 6.3.1 Spellcheckers, Grammar Checkers, and Computerised Dictionaries

In order to tackle the problems regarding content creation described above, the Welsh Language Board has commissioned or sponsored several spelling and grammar checkers and computerised dictionaries. The first was Cysill, commissioned from the University of Wales, Bangor. Cysill was revolutionary in that it corrected Welsh's initial consonant mutations. This was followed by Cysgair, a computerised dictionary which interfaced with Word processing applications, again funded by a grant from the Board. Both products were updated in 2004, and included on a single CD of Welsh language computer resources, again with the support of the Board, under the name Cysgliad<sup>15</sup>. This package is updateable via the internet.

Amongst the other spellcheckers that already exist in Welsh are the following:

- A free Welsh language spellchecker for Microsoft Office XP<sup>16</sup>
- A free Welsh language spellchecker for OpenOffice<sup>17</sup>
- A free Welsh language spellchecker for Microsoft Office 2003, as part of the free Language Interface Pack for Microsoft Office 2003 described in section 6.2.2 above
- A free Welsh Language Spellchecker for Neo Office (Apple Mac)<sup>18</sup>

## Targets/Policy Statements in the field of spellchecking and grammar checking

<sup>15</sup> <http://www.bangor.ac.uk/ar/cb/meddalwedd.php>

<sup>16</sup> <http://www.microsoft.com/downloads/details.aspx?FamilyID=E02F3714-1FEB-4CA9-9839-497B806B610E&displaylang=CY>

<sup>17</sup> <http://www.e-gymraeg.co.uk/myspell/index.html>

<sup>18</sup> [http://www.meddal.unarddeg.com/index\\_files/Page977.htm](http://www.meddal.unarddeg.com/index_files/Page977.htm)



- We will consider the possibilities of creating Welsh language proofing tools for applications which do not yet benefit from them, for example, applications used in the publishing industry
- We will consider co-operation to facilitate newer versions of existing proofing tools
- We will facilitate the distribution of Welsh language proofing tools

### 6.3.2 Welsh Language Keyboards, diacritic marks & fonts

Another element of content creation in Welsh which caused substantial difficulty for computer users in the past was the absence of the 'ŵ' and 'ŷ' from the standard character set. Initially, this was dealt with by the creation of specialist Welsh language fonts which mimicked the system fonts included as standard on PCs, but which contained the circumflex on 'w' and 'y'. However, both these characters have been included as standard in the Unicode (UTF-8) character set for some time, obviating the need for the purchase or download of specialist fonts for PC users.

However, it has not been clear enough how individual end users should access these diacritics in a standardised way, with individuals and different, or even the same, institutions choosing different keyboard shortcuts, or using character code numbers to insert diacritics into files. This standard diacritic keystroke problem has now been solved, for PC users, in two ways:

- Microsoft Windows XP Service Pack Two (XPSP2)<sup>19</sup> contains a *locale* for Welsh. (*Locale* support is a base for localisation of software, which is described in section 6.2 above and in more detail in the standards document). This Welsh locale includes date format support and an UK Extended Keyboard which gives standard keystrokes for all the accent marks used in Welsh (and other Celtic languages). A technical description of this enablement can be read on the Microsoft website<sup>20</sup>. The free CD available from the Welsh Language Board containing the Language Interface Packs for Windows XP and Office 2003 will, at the user's choice, automatically install XPSP2, and the date and keyboard settings noted above. It should be noted that the locale support contained in XPSP2 is part of the *XP operating system* and is therefore *not* limited to users of any suite of *Office* programs.
- For other PC operating systems, Draig Technology Ltd<sup>21</sup> has created a freeware utility called To Bach<sup>22</sup> (the Welsh for 'circumflex') which enables standard keystrokes for diacritics in Welsh.
- Support for ŵ and ŷ has been included in the Linux Kernel since 1999 and the main Linux, KDE and GNOME desktops are completely compliant with UTF-8. This means that Welsh is fully supported in KDE and GNOME, including interface language, character map, keyboard and fonts. The volunteers working on these systems accept suggestions for improvements to translations which are integrated to subsequent desktop releases, which occur, on average, every two months.

<sup>19</sup> <http://support.microsoft.com/default.aspx?scid=fn;EN-US;windowsxpsp2>

<sup>20</sup> <http://www.microsoft.com/globaldev/handson/user/welsh.mspx>

<sup>21</sup> <http://www.draig.co.uk/>

<sup>22</sup> <http://www.draig.co.uk/ls/tobach/default.asp?sid=01FD8394-816F-45E5-8844-64B4CA737356>

Having said this, we receive frequent enquiries regarding use of diacritics, particularly on the Apple Macintosh system, and we are eager to enable users of this, and other systems, to type Welsh with ease.

### Targets/Policy Statements

- We will, as part of the implementation of this strategy, consider actions needed to enable use of Welsh on other operating systems and packages.
- We will publish simple information on how to set up access to these diacritic marks on our website

### 6.3.3 Diacritic Marks and Web pages in Welsh

Diacritic marks, especially on 'w' and 'y' have proven problematic for many institutions publishing Welsh language web pages. This also no longer need be awkward, and the standards document gives detailed technical instructions on how to overcome these problems via the correct character set encoding.

Having given an overview of the base technology needed to enable content creation in Welsh, and to proof Welsh language documents, we now continue to the contribution that IT can make to a large field of content creation in Welsh: translation.

## 6.4 Translation and terminology: how can IT help Welsh?

The Welsh Language Board already has a Strategy for Translation which has been issued for public consultation. This section, therefore, will not attempt to duplicate any of that strategy's content, and should be seen as dovetailing with it. This section only deals with *technology* and translation, and the contribution that technology can make to the Welsh language translation industry. It commences with a description of IT tools which can increase consistency and speed of translation, whilst simultaneously enabling translators to share work and therefore reduce repetition substantially, Computer Aided Translation.

### 6.4.1 Computer Aided Translation

CAT is the broadest term used to describe an area of computer technology applications that automates or assists the act of translating text from one language to another. They are highly effective in improving translation productivity, especially in facilitating very quick translation of repetitive source texts. Canolfan Bedwyr has published a detailed and useful Welsh language guide to such software<sup>23</sup>, a similarly useful primer is available in English from Language Partners.com<sup>24</sup>. A list of translation memory software providers is provided on the Welsh Language Board's website under 'advice – information technology'.

Translation memory software is already used by a range of public institutions in Wales, amongst them, the Welsh Assembly Government and the Welsh Language Board itself. One of the main virtues of such technologies is that translators can share each other's translation work, regardless of geographic location, via central databases of translation memories served to individual PCs

<sup>23</sup> <http://www.e-gymraeg.co.uk/hyfforddiant/Cymraeg/cofcyfieithu.htm>

<sup>24</sup> <http://www.languagepartners.com/reference-center/whitepapers/catinto.htm>

via corporate networks or even the internet. This means that 100% repetition of source text segments can be deleted, thereby leading to an increase in the consistency and speed of translation (for information on consistency of terminology, see section 6.4.2 below).

Translation memory applications can also offer translations of partial or 'fuzzy' matches in their databases. This also increases the speed of translation. The main caveat of this approach of course, is that the quality of translations shared through centrally stored translation memories depends on the quality and consistency of all those translations fed into them. High-level editorial control is therefore obviously needed on such large-scale projects. Other regions have experimented with regional and national level translation memory databases, and we will examine their experiences via the European Network of Language Boards for which the Welsh Language Board is the secretariat. This network is explained in detail in section 6.15.

### **Targets/Policy Statements**

- We will use our international networks to examine the potential of forming regional or, even, national Welsh language translation memory databases using internet technology which is already available. See, for example, [www.Multilingual.com](http://www.Multilingual.com).

### **6.4.2 Terminology Management**

Again, it should be noted at the start of this section that it is not the place of the present document to deal with the general field of standardisation of terminology *per se*, merely to facilitate the diffusion of standardised terminology via IT.

We aspire to a situation where IT is used to the full extent of existing technology to automatically serve standardised terminology to translators and individuals.

We have commissioned a database to collate various existing lists. The Board intends to make this available during 2005 as a searchable database on its website. We must, however, look beyond the short term and envisage how IT can help further automate and facilitate access to terminology. Technology already exists, as for the translation memory software described in section 6.4.1 above, to automatically pass standardised list of terminology to individual end users. This feeds into terminology management applications linked to those translation memory applications. Some such applications also offer disambiguation facilities which provide extra information to end users, e.g. the Welsh version of Queen's Road in Aberystwyth (Morfa Mawr), may not be the same translation as Queen's Road in any other town in Wales ('Ffordd y Frenhines' is a grammatically and correct translation, but not used for the street name in Aberystwyth). The word "access" has many meanings, including as a noun describing the location where one approaches a building, as well as a verb meaning the actual action of gaining entrance to a building, and indeed gaining access to information; all these meanings could be differentiated by means of disambiguation facilities.

### **Targets/Policy Statements**

- We will use our international and other networks to examine the potential of forming a national Welsh language standardised terminology database using internet Terminology Server technology which is already available.
- We will use internationally recognised standards to facilitate the exchange of terminology

### **6.4.3 Alignment technology: the creation of parallel corpora of English-Welsh for Translation Memories**

As well as more consistent and faster translation and easier access to standardised terminology, CAT also allows legacy translations (stocks of translations previously undertaken) to be imported into translation memory databases, thereby increasing their size and the probability that even more matches will be reached in the future. This is accomplished using a process known as file alignment. It can be a laborious process, again requiring high levels of editorial control, and is particularly suited to stores of documents whose construction and formatting is identical, or very similar in both languages. The alignment matches corresponding pairs of source and target translations, thereby creating a corpus of parallel segments to be imported later into the translation memory database.

#### **Targets/Policy Statements**

- As part of the study of translation memory software noted above, we will consider what parallel text corpora of segmented translation could possibly be incorporated into a national database of translation memories.
- We will use internationally recognised standards to facilitate the exchange of translation memories.

### **6.4.4 Translation workflow and Bilingual Document Management**

The management of many translation projects occurring at the same time can be greatly aided by IT and computerised workflow systems. These can also monitor availability in the capacity of various external or freelance translators to undertake extra work (such as number of words translatable per day per theme, and the cost rate). The more sophisticated of these can even interface with CAT tools in order to reduce repetitive translation which is contracted out. Such technology can also, of course, manage offices, for example, in institutions such as local authorities who have pools of in house translators. It would even be possible for external freelance translators to register for such services in order to receive regular work, and updates to translation memory and consistent terminology.

#### **Targets/Policy Statements**

- As part of the study of translation memory software noted above, we will consider the feasibility of workflow systems for the Welsh language translation industry.

### **6.4.5 Translation workflow and simultaneous auto updating of multilingual Content/Document Management Systems**

The translation memory systems described above are able to be plugged into content/document management systems of all sorts, and linked to workflow solutions for translation units/companies/freelancers. In a world where

information is constantly changing, it is important that both language versions are simultaneously published and updated.

On, for example, a governmental or local authority website serving the public in Wales, parts of pages or paragraphs may be amended on a daily basis. Where the content management system is managed by non-Welsh speakers, this technology can automatically route amended segments/passages to a translator who will then translate them using translation memory (and terminology management software). Once the translator has completed translation, the workflow system will automatically route the translation back to the content management system which will publish it simultaneously with the original English version which has been stored awaiting its Welsh language counterpart. This is an important tool for the further normalisation of the Welsh language and will enable end users to have access to up to date English and Welsh text. (One of the results of the snapshot surveys of websites described in section 6.2.4 above was that the Welsh versions of some public sector websites were not updated as consistently as the English versions).

### Targets/Policy Statements

- As part of the study of translation memory software noted above, we will seek out partner organisations to trial automated content/document management systems in order to further normalise language choice, and consistently updated content for the public.
- We will also ensure that these systems include support for content creation *in* Welsh, and routing for translation into English and onward auto publishing

### 6.4.6 Machine Translation

Machine translation (MT), or automatic translation is a theme often raised when discussing the Welsh language, translation and IT issues. MT has benefited from many years of research, and some major advances have been secured. However, MT technology can not yet offer translation quality to match that of human translators. Nevertheless, usage of free on-line MT systems indicates that for certain types of text, MT can provide a useful and usable level of translation. Furthermore, in combination with controls on the language used for example in technical authoring, MT can provide an excellent quality of first-draft translation needing little revision, and offering great savings in translation costs. It can also provide gist translation from Welsh, enabling access to the international community, and enabling non-Welsh speaking staff to deal with written correspondence from Welsh speaking colleagues/members of the public.

A Welsh Language Board feasibility study<sup>25</sup> into machine translation by Professor Harold Somers of Manchester University, and Editor of the *International Journal of Machine Translation*, states that the state-of-the-art of Machine Translation and the needs of the Welsh-speaking community indicate that it is timely to invest in the development of MT.

---

<sup>25</sup> <http://www.welsh-language-board.org.uk/en/cynnwys.php?cID=6&pID=109&nID=1189>

As a result, the Welsh Language Board has extended an invitation to tender for the first machine translation engine for Welsh. This will have three integral parts: (EBMT [Example Based Machine Translation], RBMT [Rules Based Machine Translation], and SBMT [Statistics Based Machine Translation]. This will enable gist translation between Welsh and English, and, in the future integration with other applications (such as the research pane of MS Office and translation memory systems). This should be very much considered as an aid to translators, editors et al, rather than a means to supplant them.

### **Targets/Policy Statements**

- We will implement the findings of Prof. Somers' feasibility study by funding a free machine translation service between Welsh and English which will be available as a web service and as a free add-in to the most popular office suites.

## **6.5 Language technology issues and corpora**

Sufficiently sized corpora of written Welsh (which could, for example, contain a large number of electronic versions of printed publications) are a prerequisite for further developments in language technology. Such language technology is the basis for many other Welsh language applications, for example the speech technology discussed in section 6.6, and the machine translation discussed in section 6.4.5. At the forefront of this field for many years, Canolfan Bedwyr have developed lemmatisers (which break down given words and tag their grammatical forms), corpora (large databases of written or spoken Welsh), algorithms for sort orders (the Welsh alphabet is different to the English [see also the standards document]), and other language engineering issues. While the majority of these resources *in themselves* will only be of specialist interest, the effect of these very necessary tools is far-reaching and significant, as they feed into other language technology applications. (See also the Open Source Konjugator<sup>26</sup> resources.

### **Targets/Policy Statements**

- We will review the need for updates to these resources and consider related applications for funding/commissioning. Priority will be given to those schemes which have the largest possible impact on the potential target audience.

## **6.6 Speech Technology**

Speech technology may involve production of a synthetic voice or recognition of a human voice by a given IT system. Such technology is already beginning to permeate our everyday lives (Many call centres have automated their processes using speech synthesis [see section 6.6.1], certain mobile phones which can receive e-mail already offer a synthetic voice facility to read e-mail messages aloud to the recipient.). Speech technology can be an asset to any given IT program. It can simplify data access, speed data entry, and allow hands-free control. Significantly, it has enormous repercussions for the visually impaired (see section 6.6.3).

As with many major technical developments, the initial research into speech technology was carried out in majority languages. As indications show that

<sup>26</sup> <http://www.rhedadur.org.uk/index.php?lg=en>  
21 of 40

speech technology, through convergence with other everyday applications, will become a more important part of daily life in the future, it is important that the Welsh language secure a strong foothold in this field. It should therefore be noted that the Welsh Language Board has already co-funded, with INTERREG, the WISPR<sup>27</sup> Project (Welsh and Irish Speech Processing Resources) at Canolfan Bedwyr at the University of Wales, Bangor.

### 6.6.1 Speech Synthesis

The WISPR project has created an initial SAPI compliant (Speech Application Programming Interface) speech synthesis engine for Welsh, which will enable a free Text to speech (TTS) facility. Amongst many other uses, speech synthesis engines can be used to convert words from a computer document (e.g. word processor document, web page), or interface into audible speech spoken through the computer sound system. This would be helpful to people who need or want aural verification of what they are seeing in print. It can also be integrated into many other applications, including those described in sections 6.6, 6.6.4 and 6.6.5.

### 6.6.2 Speech Recognition

Speech recognition, (as defined by the WISPR Project team) or speech-to-text, involves capturing and digitizing the sound waves from a microphone, converting them to basic language units or phonemes, constructing words from phonemes, and contextually analyzing the words to ensure correct spelling for words that sound alike (such as dear and deer). The output is then displayed on the screen as text.

Again, although the technicalities may seem beyond the lay person, the reach and significance of this facility should not be underestimated, as many computer operating systems *already* offer voice recognition facilities as standard. This is only likely to increase with the passage of time. It is not inconceivable that we will order food, do our banking and a whole host of other services via speech recognition in the future. We should therefore also consider “natural user interfaces”, and “natural language commands”, i.e. rather than using a set of *specified* commands, one would speak naturally and the computer system would act appropriately. We should also have as a long-term target “speaker independence”, i.e. to use a strange computer and tell it what to do, without having had to train it previously in the peculiarities of one’s own voice and accent.

### 6.6.3 Adaptive technology

As mentioned above, speech technology has much potential for visually and hearing impaired people via what is known as ‘adaptive technology’. There are 100,000 people in Wales with serious sight loss and 450,000 people who are deaf or hard of hearing in Wales.<sup>28</sup> It is therefore a matter of social inclusion and, for example, the Disability Discrimination Act 1995, that Welsh language adaptive technology is available to this large group of people.

Screen readers are used to replace the visual display traditionally viewed on a monitor for those with visual disabilities. The speech engine described above

<sup>27</sup> <http://www.bangor.ac.uk/ar/cb/wispr.php>

<sup>28</sup> Sources: [www.rnid.org.uk](http://www.rnid.org.uk), [www.rnib.org.uk](http://www.rnib.org.uk)

produces synthesized voice output for text displayed on the computer screen, and/or for keystrokes entered on the keyboard. Talking web browsers use the same technology as screen reading software, but the reading functions are limited to Internet use. It is vitally important that such technologies exist in Welsh.

Although the developmental stage for adaptive technology speech recognition in Welsh is further away than for TTS, it is nevertheless vital for people with movement difficulties, those suffering from RSI (Repetitive strain injury) and other conditions which prevent them from using keyboards, mice and other input devices. This is dealt with in the target under 6.6.5 below.

#### **6.6.4 MT, SR, TTS (Integration of Machine Translation and Speech Technology)**

Any strategy document should provide a clear direction for long-term developments in its subject area, however ambitious. One field which we will consider in the future is the integration of speech technology with the machine translation technology described in 6.4.5 above. The ideal scenario would enable two people, one speaking Welsh, the other English to converse with each other. This would be accomplished by speech recognition, feeding into machine translation engine, and outputting the relevant translation via speech synthesis. However far-fetched this situation may seem at present, it must be emphasised that research is already underway to the field of such integrative technologies in several institutions, for example The Language Technologies Institute of Carnegie Mellon University.<sup>29</sup> The potential for such integration for Welsh language services in an age increasingly controlled by IT is obvious. We will keep a close watch on such integration with a view to availing itself of technical expertise in the field in the future.

#### **6.6.5 Embedding of speech technology into other applications**

Apart from the developments described above, speech technology *already* feeds into other applications in different languages, e.g. child consultation software, telephone weather forecast lines, telephone banking PABX, language learning software and Public Address systems of many types.

#### **Speech technology Targets/Policy Statements**

- We will monitor the WISPR Project and its successors and consider requests to facilitate transfer of the resultant speech engine to other applications.
- We will keep a watching brief on the integration of MT and TTS in other languages, in order to avail ourselves of the most appropriate developments in this field at an appropriate stage in the future.

### **6.7 Capacity building**

This section deals with increasing human potential in the field of IT through the medium of Welsh, and how to ensure continuity and growth of human resources to enable us to develop language technology for the future sustainably.

<sup>29</sup> <http://www.lti.cs.cmu.edu/Research/cmt-projects.html>

### 6.7.1 End Users: IT Training in Welsh

The field of end user training and capacity building through the medium of Welsh in the field of IT requires specific attention (see also section 6.8). Various training suites exist, both in electronic and paper format. The European Computer Driving License is one of the main Welsh language IT training resources available at the moment, part-funded by the Board.<sup>30</sup> As the world of IT moves so quickly, updates to this type of training will regularly be needed, and the Board is eager to cooperate with this, especially to encompass the many computer systems which now offer Welsh language interfaces. To this end, ELWa are carrying out a scoping study of the future avenues for Welsh medium IT training.

#### Targets/Policy Statements

- Together with other relevant bodies, we will discuss the best ways to drive forward IT training through the medium of Welsh.

### 6.7.2 Developers: enlarging the base of language technology specialists

Welsh has benefited substantially in the field of IT from the goodwill and foresight of many enlightened individuals and institutions, and the Board is eager for this to continue, and to facilitate it. We need to ensure that future generations can draw upon sufficient linguistic technological expertise in order for Welsh to sustain and improve its status in this field. To this intent, we will consider the feasibility of funding or facilitating scholarships, apprenticeships, or placements in the various subject areas described in this strategy document.

This strategy document also provides a framework within which Welsh universities and businesses can develop specialist IT skills. There is worldwide demand for bilingual or multilingual IT solutions, meaning that advanced skills, knowledge and research would provide major export opportunities for Welsh companies. Building this capacity will also have beneficial side effects. It will place a higher value on individuals with combined language and IT skills and also build the Welsh economy, creating economic advantages from the use and support of the language.

#### Targets/Policy Statements

- We will consider the feasibility of funding or facilitating scholarships, apprenticeships, or placements in the various subject areas described in this strategy document, e.g. computational linguistics, subject to the priorities and financial constraints mentioned above
- We will contact and examine the potential contributions of professional bodies in the field of IT to development of Welsh language resources and standards
- We will engage in dialogue with relevant bodies in order to further the development of multilingual solutions developed in Wales
- We will scope the possibilities of creating a team of specialists who can help institutions implement multilingual solutions in various institutions

---

<sup>30</sup> <http://hyfforddiantecdl.bangor.ac.uk/>  
24 of 40

## 6.8 E-learning & The Welsh Language

It should again be noted that ELWa and the Welsh Assembly Government already have an e-learning strategy, which this present document does not aim to duplicate.<sup>31</sup> We will co-operate with the implementation of that strategy as appropriate. It is necessary, however, to mention several elements regarding e-learning and the Welsh language.

### 6.8.1 Welsh Learners, *Gloywi iaith* & E-learning

Although the responsibility for Welsh Learning teaching provision lies with other organisations, the Board, by its very nature and remit, takes a keen interest in the success of the Welsh for Adults System and in courses offered to improve the written Welsh of fluent speakers (*Gloywi iaith* classes). It is for this reason that the Board has entered into a partnership with BBC Wales to facilitate the reading of Welsh language web pages by learners and those Welsh speakers who lack confidence to read entire pages of Welsh unassisted. Stage one of this project, BBC Vocab, was first made available on the BBC website.<sup>32</sup>

This innovative system passes a given Welsh language web page through the BBC Online Welsh dictionary, and tags all the words on the web page which are found in the dictionary with a transparent yellow background. On hovering the mouse over one of these yellow backed words, the equivalent English word appears, together with an option to take the reader to the full definition in the BBC online Welsh dictionary.

The Board's involvement with this project increased with its second and third phases. This involved us sponsoring gradation of vocabulary items on the BBC Online Welsh language news website classed by frequency counts. The virtue of the system, whatever level of vocabulary chosen, is its simplicity, and it transcends other methods of vocabulary aid in that it does not impinge at all on the flow of reading.

We have also included the BBC Vocab facility on our own suite of websites and will encourage all institutions and individuals managing websites with Welsh language content to do the same.<sup>33</sup>

### Targets/Policy Statements

- Wherever appropriate within our remit, we will consider co-operation, facilitation and possible funding of IT facilities to aid the learning and boosting of confidence in Welsh
- We will, in partnership with the BBC, encourage the take up of the BBC Vocab service for Welsh by other institutions and individuals managing websites with Welsh language content
- We will continue our partnership with the BBC to implement the gradation scheme for BBC Vocab.

## 6.9 Call centre software

As noted above, the public's interaction with all types of organisations - is increasingly dependent on IT and voice services in particular. The Board is

<sup>31</sup> <http://www.elwa.org.uk/elwaweb/elwa.aspx?pageid=3461>

<sup>32</sup> <http://www.bbc.co.uk/cgi-perl/vocab/show.pl/cy-en/news.bbc.co.uk/welsh/default.stm>

<sup>33</sup> <http://www.bbc.co.uk/cgi-perl/vocab/show.pl/cy-en-x/www.welsh-language-board.org.uk/cy/index.php>



naturally keen to ensure that the call centres' IT systems support an accessible language choice. We will, therefore, develop guidelines on making the best use of the technology in order to promote bilingualism and normalise the use of Welsh.

### **Targets/Policy Statements**

- We will undertake a study of the latest Call Centre software and PABX (Private Automatic Branch Exchange) systems in order to develop a set of guidelines of best practice for the implementation of bilingualism in this field.
- We will enter into discussions with relevant umbrella bodies in the field of call centres regarding implementation of bilingual strategies for provision of service

### **6.10 Back end databases and bilingualism**

Currently, there are over 300 organisations operating a Welsh Language Scheme, with many others offering Welsh Language services through voluntary Welsh Language Schemes or without a scheme. An increase in the development of services in many of these organisations, and others, is dependent on ensuring that their back end database systems, such as those which hold UK wide information on individuals in Wales, support the offering and implementation of language choice.

Occasionally, we hear from some organisations that this is not possible without spending substantial amounts of money or subjecting their systems to unacceptable risks. We will therefore undertake a study to measure these statements objectively, and to develop practical guidelines on making the best use of the organisations' IT structures.

The study will:

- Analyse the relationship between the services and the basic technology to assess the possibility of practical improvement of bilingualism without substantial investment, and examine how the standards document can be implemented in such databases

### **Targets/Policy Statements**

- We will commission a study on back end databases and bilingualism and publicise its findings

### **6.11 Metadata**

Metadata describes the content and purpose of other objects. A metadata record is a label, just like the label one might find on an everyday product in a supermarket (see the Govtalk<sup>34</sup> website). The label describes the product, and contains useful information (in particular to the electronic systems that control the store).

Metadata makes content findable by describing it in structured ways that computers can understand. Through metadata, people can browse their way to content instead of having to enter search keywords; or they can carry out

<sup>34</sup> [http://www.govtalk.gov.uk/documents/Laymans\\_guide\\_to\\_metadata.doc](http://www.govtalk.gov.uk/documents/Laymans_guide_to_metadata.doc)  
26 of 40

sophisticated searches, for example to look for all the documents on a given subject, published in a given year.

Metadata can be applied to any kind of digital content - text, data, images, sound recordings and so on.

We need to ensure that guidelines for metadata and all relevant associated materials take due account of bilingualism in Wales as standard, in the same way that the *The Illustrated handbook for web management teams* produced by the UK Government's Cabinet Office did after the Board responded to its consultation document in 2002.<sup>35</sup> This is also dealt with in the Welsh context in the standards document.

One interesting avenue is the IPSV (Integrated Public Service Vocabulary) controlled vocabulary which enables standardised searching and interoperability of public sector website.

### Targets/Policy Statements

- We will enter into dialogue with relevant IT Standards sections of Government in order to ensure that bilingualism in Wales is mainstreamed as a standard to be followed.
- We will facilitate, with other relevant organisations, the creation of a standardised version of the IPSV in Welsh, and assist with its dissemination.

## 6.12 Mobile devices

Mobile devices, phones, PDAs etc, have, like many other areas of IT, also permeated our everyday lives over the last few years. As they converge with previously separate items, e.g. radios, computerised personal organisers, and music storage devices, they become even more powerful and necessary objects for the smooth running of communication, daily tasks and leisure. This is all the more important for the promotion of Welsh to young people (see also the Board's Youth Strategy).<sup>36</sup> It therefore stands to reason that the Board wishes for them to support Welsh. As mentioned in section 6.6.1, many telephones now offer a synthetic voice solution to read e-mail. Another obvious target is the interface of mobile telephony which the user sees. Although this section of the strategy document is small, we emphasise that we are aware of the large and pervasive nature of mobile computing, and will do all within our power and influence to increase the status of Welsh within this all-important field.<sup>37</sup>

### Targets/Policy Statements

- We will discuss with manufacturers of mobile telephony software in order to ascertain what co-operation can be undertaken to increase the status of Welsh on handsets

---

<sup>35</sup> [http://e-government.cabinetoffice.gov.uk/Resources/WebHandbookIndex1/WebHandbookIndex1Article/fs/en?CONTENT\\_ID=4000079&chk=utLkDp](http://e-government.cabinetoffice.gov.uk/Resources/WebHandbookIndex1/WebHandbookIndex1Article/fs/en?CONTENT_ID=4000079&chk=utLkDp)

<sup>36</sup> <http://www.welsh-language-board.org.uk/en/cynnwys.php?cID=1&pID=109&nID=1426>

<sup>37</sup> Predictive texting solutions are also available in different formats. See, for example, <http://wiki.eatoni.com/wiki/index.php/Products#Languages>

## **6.13 Embedded software**

Embedded operating systems can be seen in nearly every advanced electronic object: televisions, telephones, cash machines, hi-fi systems, GPS, office printers, vending machines, parking meters. The vast scope and variety is a challenge, but many embedded systems are already designed to be multilingual, and in these cases the addition of Welsh as a language option is technically simple - most cash machines and public payphones in Wales already have a Welsh language option.

### **Targets/Policy Statements**

- We will enter into discussion with electronic systems manufacturers with a view to adding Welsh to multilingual systems

## **6.14 Optical character recognition and Handwriting recognition**

These two input methodologies work in a variety of ways to recognise text on a printed page, or to convert human handwriting to text on screen via, for example, a tablet PC using a screen pen or stylus. Documents can be annotated by hand on screen, and those annotations converted into legible text. In order to ascertain a level of correct languages, the symbols recognised are sometimes passed through a dictionary in the target language, and different suggestions passed to the end user.

### **Targets/Policy Statements**

- We will consider the way to further develop forward for input methodologies for Welsh, considering the priorities accorded to each element in this strategy document.

## **6.15 Sharing International Experiences in the field of IT**

Many other areas with similar linguistic situations to Wales have made large advances in the field of IT, e.g. the Basque, Galician and Catalan Autonomous Communities. The interchange of ideas between these regions is therefore important in all spheres of language planning. The Welsh Language Board is a member of the European Network of Language Planning Boards, which was set up in October 2001 to promote regional and minority languages (RMLs) on a European level. It is also member of the Regional or Minority Language Working Group of the British Irish Council.

The Network focuses upon specific issues that affect RMLs across Europe, underpinning the drive to improve language awareness, enhance learning opportunities and develop and support best practice in all aspects of language planning. One of the project's four thematic strands is Information Technology, with Wales and the three autonomous communities above as members of that strand group.

The network will endeavour to gather experience from wider language communities in the fields of machine translation and translation memory software, terminology management, multilingual document management and workflow systems, speech synthesis and recognition, localisation processes, integration of legacy translations into translation memory systems, standardisation of IT terminology technology and, especially, ways in which IT

can be used to provide minority language learning opportunities and help to create new speakers

Having said this, there is no reason why the research and development done on larger languages cannot be used for the benefit of Welsh, and no reason why research and development done on Welsh cannot be of benefit to those languages.

### **Targets/Policy Statements**

- The European Network of Language Planning Boards will discuss and disseminate relevant information in the field of ICT
- The European Network of Language Planning Boards will ascertain how each member can cooperate to influence software companies and take advantage of economies of scale for multilingual developments
- We will cooperate with partners in the British Irish Council in order to push forward the IT agenda for all languages of the UK

### **6.16 Future-proofing for English-Welsh bilingualism in IT**

The world of IT develops at breakneck speed. Only a few decades ago, most would not have believed that computing would have affected so many areas of our lives and that office work, and a large amount of personal correspondence would now be completely dominated by e-mail and the PC - not to mention the all-pervasive influence of the World Wide Web. Just as the typing pool has become a thing of the past, even today's 'modern' technology will quickly evolve – successors are already being proposed to the web itself. It is important that we know the direction that these developments will take and that Welsh be considered at the developmental stage, in order for bilingual services and communication to develop unfettered in the future. We simply do not know what will happen in the long term in the field of IT, therefore, we propose the following target.

### **Targets/Policy Statements**

- We will carry out a study of the likely mid-term technological developments and usage patterns which would be of most benefit to the Welsh language. The results of these studies will then be fed into updated versions of the present strategy document.
- We will, where appropriate, include use of IT as part of language use surveys and research projects that we commission.

## 7 Further reading and selected web references

### 7.1.1 The Welsh Language Board Website

The Board's website is a specialised electronic resource about the Welsh language and contains a comprehensive library of publications. This has enabled us to offer all our current publications electronically, including specialist information on the linguistic results of the 2001 Census. A sophisticated categorisation system has also been finalised which will flag relevant content for the reader from wherever they are on the website.

#### Targets/Policy Statements

- We will build upon the success of our website and associated micro sites
- We will rework the look and feel of the site to make it easier to use
- We will continue its work to make the site accessible to the visually impaired and will incorporate any relevant code for the Welsh language adaptive technology the Board is sponsoring, as soon as it becomes available (see section 6.6.3).
- We will include the BBC Vocab facility on all our websites
- We will transfer all relevant functionality from the Board's website when the Board is incorporated into the Welsh Assembly Government

### 7.1.2 International Standards

These standards, consulted during the drafting of this document, are available from the International Organization for Standardization's website<sup>38</sup>

- ISO 10241 International terminology standards – Preparation and layout
- ISO (INTERNATIONAL STANDARD) 1087-1 Terminology work — Vocabulary —Part 1:Theory and application
- ISO (INTERNATIONAL STANDARD) 1087-2 Terminology work — Vocabulary —Part 2:Computer applications
- ISO (INTERNATIONAL STANDARD) 12199 Alphabetical ordering of multilingual terminological and lexicographical data represented in the Latin alphabet
- ISO (INTERNATIONAL STANDARD) 12200 Computer applications in terminology —Machine-readable terminology interchange format (MARTIF) — Negotiated interchange
- ISO (INTERNATIONAL STANDARD) 12615 Bibliographic references and source identifiers for terminology work
- ISO (INTERNATIONAL STANDARD) 12616 Translation-oriented terminography
- ISO (INTERNATIONAL STANDARD) 12620 Computer applications in terminology —Data categories
- ISO (INTERNATIONAL STANDARD) 15188 Project management guidelines for terminology standardization
- ISO (INTERNATIONAL STANDARD) 15836 Information and documentation — The Dublin Core metadata element set

<sup>38</sup> <http://www.iso.org/iso/en/ISOOnline.frontpage>



- ISO (INTERNATIONAL STANDARD) 1951 Lexicographical symbols and typographical conventions for use in terminography
- ISO (INTERNATIONAL STANDARD) 639-1 Codes for the representation of names of languages — Part 1: Alpha-2 code
- ISO (INTERNATIONAL STANDARD) 704 Terminology work — Principles and methods
- ISO 860 Terminology work – Harmonization of concepts and terms

### 7.1.3 Galicia

Web references used in hyperlinks in the text are not necessarily replicated here. This list is not intended to be exhaustive.

- COTOVÍA: <http://www.cirp.es/prx/proyectos.html>
- CORGA: <http://corpus.cirp.es/corga/>
- <http://www.cirp.es/>
- <http://www.consellodacultura.org>
- <http://www.usc.es/>
- <http://www.uvigo.es/>
- **Ciberlingua**: <http://www.ciberlingua.org/>
- **Crayons, application for Galician teaching**:  
<http://www.galego21.org/ldc/>
- **Galego21**: <http://www.galego21.org/index.shtml>
- **Project Rianxo**: <http://www.geocities.com/Athens/Aegean/7229/>
- **Project Trasno**: <http://trasno.gpul.org/>
- **Project Xis**: <http://www.galego21.org/xis/>
- **Prizes Eganet** (aimed at Galician companies and professionals working in favour of the Information Society): <http://www.premioseganet.org/>
- **Prizes Dorna de informática** (aimed at the creation of web pages and the development of applications for computers in Galician):  
[http://webs.uvigo.es/normalizacion/ani\\_dorna.html](http://webs.uvigo.es/normalizacion/ani_dorna.html)
- **Prizes SIEGA** (aimed at the design of web pages for educational centers, the development of educational material and image galleries in internet, and the creation of pedagogical videos or animations):  
<http://www.edu.xunta.es/ftpserver/portal/CONTIDOS/PREMIOS2004/WEB/>

Arias, X. C. (dir.) (2002): *Información, tecnoloxía e territorio: Galicia ante a sociedade do coñecemento*, Santiago de Compostela, Consello da Cultura Galega.

Web site:

[http://www.consellodacultura.org/mediateca/pubs.pdf/informacion\\_tecnoloxia\\_territorio.pdf](http://www.consellodacultura.org/mediateca/pubs.pdf/informacion_tecnoloxia_territorio.pdf)

Baleirón, R. *et alii* (1998): *O labirinto virtual: sobre as novas tecnoloxías na educación*, Vigo, Xerais.

Cuíñas Rodríguez, J.: “A nosa lingua nas tecnoloxías da información e da comunicación”, *Pontenorga* 0 (1998), 13-16. (No CDSG)

Docampo Amoedo, D. (2002): “As tecnoloxías da información e as comunicacións”, en López, X. (coord.), *A comunicación en Galicia 2002*, Santiago de Compostela, Consello da Cultura Galega, 127-138.  
Web site: <http://www.consellodacultura.org/mediateca/pubs.pdf/informe02.pdf>

Docío-Fernández, L., Garcia-Mateo, C. (2000): “Automatic Segmentation of Speech based on Hidden Markov Models and Acoustic Features”. *Proceedings of ICSLP2000* (International Conference on Spoken Language Processing). Beijing (China). October 2000.

Flores Seijas, J. R. (2003): “Software en galego: mellor libre”, resumo da intervención nos VII Encontros para a Normalización Lingüística, celebrados en Santiago de Compostela os días 23 e 24 de outubro de 2003.  
Web site: <http://www.consellodacultura.org/arquivos/cdsg/docs/enl7jrflores.pdf>

Gómez Guinovart, J. e Lorenzo Suárez, A. M. (eds.) (1996): *Lingüística e informática*, Santiago de Compostela, Tórculo.

Gómez Guinovart, X. (2003): “A lingua galega en Internet”, en Bringas, A. e Martín B. (eds.), *Nacionalismo e globalización: lingua, cultura e identidade*, Servizo de Publicacións da Universidade de Vigo, Vigo, 2003, 71-88.

González González, Manuel y Fernández Rei, Elisa (1998): “Un sintetizador de voz para el gallego” (ponencia en el VIIème Colloque de Linguistique Hispanique), *Travaux de linguistique hispanique* (sous la direction de Gilles Luquet), Paris, Presses de la Sorbonne Nouvelle, 1998, 65-76.

González González, M. *et alii* (1999): “O galego e as tecnoloxías da fala: o caso dun sintetizador de voz”, en Kremer D. (ed.), *Actas do V Congreso Internacional de Estudos Galegos: un século de estudos galegos*, Sada, Edicións do Castro.

González González, Manuel (2001): “A lingua galega ante os desafíos do século XXI”, *Actas do Congreso “A cultura no século XXI”* V.1, Santiago de Compostela, Xunta de Galicia, 2001, 73-84.

González González, Manuel; García Mateo, C.; Rodríguez Banga, E.; Fernández Rei, E. (2002): *Diccionario de verbos galegos Laverca* (contén CD-ROM co programa Laverca 1.0), Vigo, Edicións Xerais de Galicia.

González González, M. (2002): “*Laverca*: diccionario de verbos gallegos con voz sintetizada” en Díaz García, J. (ed.): *Actas del II Congreso de fonética experimental*, Sevilla, Universidad de Sevilla, 209-214.

Labraña Barrero, S. (2000): “A fonética galega: da articulatoria tradicional ás tecnoloxías máis actuais”, en Kremer D. (ed.), *Actas do VI Congreso Internacional de Estudos Galegos: un século de estudos galegos*, Sada, Edicións do Castro.

López, X e Túñez, M. (2002): *Xornalismo en internet: actitudes profesionais e condicións laborais dos periodistas en liña*, Santiago de Compostela, Consello

da Cultura Galega. Web site:

[http://www.consellodacultura.org/mediateca/pubs.pdf/xorn\\_internet.pdf](http://www.consellodacultura.org/mediateca/pubs.pdf/xorn_internet.pdf)

López, X., Gago, M. e Pereira J. (2002): “Internet e os medios de comunicación electrónicos”, en López, X. (coord.), *A comunicación en Galicia 2002*, Santiago de Compostela, Consello da Cultura Galega, 115-126.

Pode consultalo en:

<http://www.consellodacultura.org/mediateca/pubs.pdf/informe02.pdf>

Lorenzo Suárez, A. M. (1994): “O tratamento informático do material sociolingüístico: algunhas suxestións a partir dos protocolos de investigación”, en Gómez Guinovart, X. (ed.), *Aplicaciones lingüísticas de la informática*, Tórculo Edicións, Santiago de Compostela, 97-112.

Lorenzo Suárez, A. M. e Gómez Guinovart, J.: “Terminoloxía, informática e lingua galega”, en *Cadernos de Lingua* 13 (1996), 5-33. (No CDSG)

Monteagudo, H. e Bouzada, X. (2004): *A lingua galega nos medios de comunicación*.

Web

site: [http://www.consellodacultura.org/mediateca/pubs.pdf/c\\_lingua\\_medios.pdf](http://www.consellodacultura.org/mediateca/pubs.pdf/c_lingua_medios.pdf)

Moreno Aguirre, J. L. *et alii*: “Aspectos ortográficos, léxicos y morfosintácticos del etiquetado lingüístico de un corpus de informática en lengua gallega”, en *Revista de la Sociedad Española para el Procesamiento del Lenguaje Natural* 27 (2001), 13-19.

*Redada* ( electronic publications on the introduction of Galician language in the internet): <http://www.galego21.org/nos/redada/artigosereleatorios.htm>

Rodríguez-Linares, L., Garcia-Mateo, C. (2000): “A Speaker Authentication Module in TelCorreo. Proceedings of TSD 2000” (Third International Workshop on Text, Speech and Dialogue) pp. 375-380. in *Lecture Notes in Artificial Intelligence 1902*. Springer Verlag. Brno (Czech Republic). September 2000.

Rodríguez-Linares, L., Cardenal Lopez, A., Garcia Mateo, C., Perez-Pinar Lopez, D., Rodríguez Banga, E., Fernandez Salgado, X. (2000): “TelCorreo: A Bilingual E-mail Client over the Telephone”. Proceedings of TSD 2000 (Third International Workshop on Text, Speech and Dialogue) pp. 381-386. in *Lecture Notes in Artificial Intelligence 1902*. Springer. Brno (Czech Republic). September 2000.

Rodríguez-Liñares, L., Garcia-Mateo, C. (2000): “Application of Speaker Authentication Technology to a Telephone Dialogue System”, *Proceedings of ICSLP2000* (International Conference on Spoken Language Processing). Beijing (China). October 2000.

Romero, D. e Vaquero, I. (2000): “Lingua e contidos nos sitios web das administracións públicas galegas”, en *Actas dos III Encontros para a Normalización Lingüística*, Santiago de Compostela, Consello da Cultura Galega. (No CDSG)

Romero, D. e Vaquero, I. (2001): *Da periferia á rede: Internet en Galicia. Lingua e contidos*, Vigo, Xerais. (No CDSG)

Romero, D. e Vaquero, I.: “O galego na rede: nova realidade, nova política”, en *Tempos Novos* 36 (2000), 40-43.

Vaquero Quintela, I. (1999): “A utilización da Internet nun SNL / SL como ferramenta de información, formación e dinamización”, en *Actas dos II Encontros para a Normalización Lingüística*, Santiago de Compostela, Consello da Cultura Galega. (No CDSG)

Valverde Juncal, M. (2004): “Primeira análise das webs municipais galegas”, comunicación lida nos IV Encontros Cultura e Concellos, organizados polo Consello da Cultura Galega en xuño de 2004. Web site:

[http://www.consellodacultura.org/mediateca/pubs.pdf/webs\\_municipais.pdf](http://www.consellodacultura.org/mediateca/pubs.pdf/webs_municipais.pdf)

#### 7.1.4 Catalonia

- <http://www.uoc.edu/in3/atlantis/cat/index.html>
- Vegeu projecte Linmiter < <http://www.linmiter.net/>>
- Secretaria de Política Lingüística < <http://www6.gencat.net/llengcat/informat/index.htm>>
- Vegeu Termcat, Centre de Terminologia <<http://www.termcat.net>>
- Xarxa Telemàtica Educativa de Catalunya < <http://www.xtec.es/>>
- Vegeu La Farga (foment i ús del programari lliure) <<http://www.lafarga.org/>>
- Vegeu Softcatalà <http://www.softcatala.org>
- Centre de Tecnologies i Aplicacions del Llenguatge i la Parla (TALP). Universitat Politècnica de Catalunya (UPC)
- < <http://www.talp.upc.es/>>
- Departament de Teoria del Senyal i Comunicacions. Universitat Politècnica de Catalunya < <http://www.tsc.upc.es/>>
- Centre de Llenguatge i Computació (CLiC). Universitat de Barcelona
- < <http://clic.fil.ub.es/>>
- Grup de Fonètica. Departament de Filologia Espanyola. Universitat Autònoma de Barcelona
- < <http://liceu.uab.es/~joaquim/home.html>>
- Institut Universitari de Lingüística Aplicada (IULA). Universitat Pompeu Fabra
- < <http://www.iula.upf.es/projec.htm>>
- Universitat Oberta de Catalunya
- < <http://www.uoc.edu/in3/cat/index.html>>
- < <http://www.uoc.edu/in3/interlingua/>>
- Universitat de Lleida < <http://www.udl.es/rectorat/vi/recerca/GRUDL/consulta.cgi#Resultats>>
- Secretaria de Telecomunicacions i Societat de la Informació <<http://www10.gencat.net/dursi/AppJava/home.jsp?area=3>>
- Secretaria de Política Lingüística <http://www6.gencat.net/llengcat/informat/index.htm>
- Termcat, Centre de Terminologia <<http://www.termcat.net>>
- Softcatalà <http://www.softcatala.org>

- WICCAC: Webmàsters Independents en Català, de Cultura i d'Àmbits Cívics < <http://wiccac.org/>>
- Observatori per a la Cibersocietat < [http://www.cibersociedad.net/congres2004/index\\_f.html](http://www.cibersociedad.net/congres2004/index_f.html)>
- Localret, Consorci de l'administració local per al desenvolupament de les xarxes de telecomunicacions i de les noves tecnologies (782 municipis adherits) <http://www.localret.es/localret/qelr.htm>
- Vilaweb Tecnologia. Diari Electrònic Independent < <http://www.vilaweb.com/tecnologia/index.html>>

## 8 Summary of targets and policy statements

- In commissioning or giving grant in aid to IT projects, we will, where appropriate, encourage or require adherence to international and open IT standards
- In commissioning or giving grant in aid to IT projects, we will ensure that the products, and where appropriate and feasible, component parts, created will be made available, subject to Crown Copyright (where applicable), under a suitable free, non-commercial licence agreement or (where part of a commercial product which offers multilingual functionality) at no additional cost to the end user.
- Term lists and gazetteers we commission/fund will be available in a computerised format that follows internationally recognised standards (e.g. ISO).
- Term lists and gazetteers we commission/fund will be able to be adapted for import into the terminology management systems described in section 6.4.2.
- We will examine the possibility of making our term lists available, free of charge, as internet updates to other relevant electronic dictionaries.
- We will examine the possibility, subject to strict standardisation quality control procedures, of a 'suggest a term' facility to our national database of standardised terms.
- We will encourage all institutions to offer software/IT provision on a basis of equality between English and Welsh.
- For those organisations with a statutory language scheme under the Welsh Language Act 1993, and who operate software whose interface is available in Welsh, we will, where reasonable and practicable, require them to offer both English and Welsh Language versions on a basis of equality.
- We will continue, where appropriate, to issue circulars regarding mainstreaming bilingual considerations into IT
- We will cooperate with and provide locale and language data to institutions who wish to embed the Welsh language in their applications.
- We will discuss with and provide advice to professional societies and institutions in the field of IT in order to publicise the needs and standards for bilingual IT provision in Wales. Where we consider necessary and/or appropriate, we will organise training and dissemination activities for the members of those organisations. In institutions that have a statutory language scheme, we will do this in partnership with the official responsible for that scheme.
- In this dialogue, we will emphasise the need for bilingual provision to be mainstreamed into all areas of policy and performance indicators
- We will circulate the associated standards document and organise suitable training/publicity events for Board staff and language scheme officers
- We will consider requests for help in localisation from the open source community
- We will engage in dialogue with software producers to explore possible synergies for localising other products, applications and suites into Welsh
- We will contact retailers of home and office PC's to discuss the possibility of offering Welsh language versions of available software as a preinstalled option or as an option in 'self build' PC internet sites.

- In all our product development discussions, calls for tender and dealings with external organisations, we will emphasise the need for easy switchability between Welsh and English.
- We will continue to carry out occasional snapshot surveys of public sector websites as part of the monitoring process of statutory language schemes, using the Standards Document as a basis
- We will examine possibilities for ensuring standard forms of compliance in this field
- We will examine the possibilities for an accreditation scheme for compliance with the standards document
- We will publish guidance (and examine the possibilities for creation of automatic tools) on how to enable a non platform specific browser language locale setting which will ensure that readers automatically receive Welsh language versions of pages where they are available.
- We will consider the possibilities of creating Welsh language proofing tools for applications which do not yet benefit from them, for example, applications used in the publishing industry
- We will consider co-operation to facilitate newer versions of existing proofing tools
- We will facilitate the distribution of Welsh language proofing tools
- We will, as part of the implementation of this strategy, consider actions needed to enable use of Welsh on other operating systems and packages.
- We will publish simple information on how to set up access to these diacritic marks on our website
- We will use our international networks to examine the potential of forming regional or, even, national Welsh language translation memory databases using internet technology which is already available. See, for example, [www.Multilingual.com](http://www.Multilingual.com).
- We will use our international and other networks to examine the potential of forming a national Welsh language standardised terminology database using internet Terminology Server technology which is already available.
- We will use internationally recognised standards to facilitate the exchange of terminology
- As part of the study of translation memory software noted above, we will consider what parallel text corpora of segmented translation could possibly be incorporated into a national database of translation memories.
- We will use internationally recognised standards to facilitate the exchange of translation memories.
- As part of the study of translation memory software noted above, we will consider the feasibility of workflow systems for the Welsh language translation industry.
- As part of the study of translation memory software noted above, we will seek out partner organisations to trial automated content/document management systems in order to further normalise language choice, and consistently updated content for the public.
- We will also ensure that these systems include support for content creation *in* Welsh, and routing for translation into English and onward auto publishing
- We will implement the findings of Prof. Somers' feasibility study by funding a free machine translation service between Welsh and English which will be available as a web service and as a free add-in to the most popular office suites.

- We will review the need for updates to these resources and consider related applications for funding/commissioning. Priority will be given to those schemes which have the largest possible impact on the potential target audience.
- We will monitor the WISPR Project and its successors and consider requests to facilitate transfer of the resultant speech engine to other applications.
- We will keep a watching brief on the integration of MT and TTS in other languages, in order to avail ourselves of the most appropriate developments in this field at an appropriate stage in the future.
- Together with other relevant bodies, we will discuss the best ways to drive forward IT training through the medium of Welsh.
- We will consider the feasibility of funding or facilitating scholarships, apprenticeships, or placements in the various subject areas described in this strategy document, e.g. computational linguistics, subject to the priorities and financial constraints mentioned above
- We will contact and examine the potential contributions of professional bodies in the field of IT to development of Welsh language resources and standards
- We will engage in dialogue with relevant bodies in order to further the development of multilingual solutions developed in Wales
- We will scope the possibilities of creating a team of specialists who can help institutions implement multilingual solutions in various institutions
- Wherever appropriate within our remit, we will consider co-operation, facilitation and possible funding of IT facilities to aid the learning and boosting of confidence in Welsh
- We will, in partnership with the BBC, encourage the take up of the BBC Vocab service for Welsh by other institutions and individuals managing websites with Welsh language content
- We will continue our partnership with the BBC to implement the gradation scheme for BBC Vocab.
- We will undertake a study of the latest Call Centre software and PABX (Private Automatic Branch Exchange) systems in order to develop a set of guidelines of best practice for the implementation of bilingualism in this field.
- We will enter into discussions with relevant umbrella bodies in the field of call centres regarding implementation of bilingual strategies for provision of service
- We will commission a study on back end databases and bilingualism and publicise its findings
- We will enter into dialogue with relevant IT Standards sections of Government in order to ensure that bilingualism in Wales is mainstreamed as a standard to be followed.
- We will facilitate, with other relevant organisations, the creation of a standardised version of the IPSV in Welsh, and assist with its dissemination.
- We will discuss with manufacturers of mobile telephony software in order to ascertain what co-operation can be undertaken to increase the status of Welsh on handsets
- We will enter into discussion with electronic systems manufacturers with a view to adding Welsh to multilingual systems
- We will consider the way to further develop forward for input methodologies for Welsh, considering the priorities accorded to each element in this strategy document.

- The European Network of Language Planning Boards will discuss and disseminate relevant information in the field of ICT
- The European Network of Language Planning Boards will ascertain how each member can cooperate to influence software companies and take advantage of economies of scale for multilingual developments
- We will cooperate with partners in the British Irish Council in order to push forward the IT agenda for all languages of the UK
- We will carry out a study of the likely mid-term technological developments and usage patterns which would be of most benefit to the Welsh language. The results of these studies will then be fed into updated versions of the present strategy document.
- We will, where appropriate, include use of IT as part of language use surveys and research projects that we commission.
- We will build upon the success of our website and associated micro sites
- We will rework the look and feel of the site to make it easier to use
- We will continue its work to make the site accessible to the visually impaired and will incorporate any relevant code for the Welsh language adaptive technology the Board is sponsoring, as soon as it becomes available (see section 6.6.3).
- We will include the BBC Vocab facility on all our websites
- We will transfer all relevant functionality from the Board's website when the Board is incorporated into the Welsh Assembly Government

## 9 List of respondents to the consultations

- Adam van Doornick
- Aled Owain Richards
- Alun Thomas
- Andrew Joscelyne
- Arwyn Roderick
- Association of Welsh Translators
- Canolfan Bedwyr E-Welsh Unit
- Caoimhín O Donnaile
- Carmarthenshire County Council
- Cwmni Pioden
- Dai Rees Jones
- Daniel Cunliffe
- Dean Baker
- Delyth Murphy
- Department of Work and Pensions
- Dewi Jones
- Dewi W Evans
- Dr Del Morgan
- Draig Technologies
- Dylan Jones
- Einion Gruffudd
- Elen Jones
- Eleri Haf Davies
- Enid Roberts
- Euryngwen Williams
- Fflintshire County Council
- Gill Richards
- Grahame Davies
- Gwenllïan Carr
- Gwydion Gruffudd
- Gwynedd County Council
- Hugh Middleton
- Hywel Marshall
- John Nicholson
- Joyce Chatterton
- Ken Sawyers
- Kevin Donnelly
- Lisa Francis AM
- Llŷr Griffiths-Davies
- Marcus Davage
- Mark Jones
- Mary Gowers
- Mary Lynn Jones
- Meurig Hughes
- Muiris Mag Ualghairg
- Myfanwy Alexander
- Neil Shadrach
- Office Québécois de la Langue Française
- Owen Thomas
- Paul Edwards
- Paul Squires
- Penri Williams
- Peter Williams
- Philippe Moseley
- Powys County Council
- Rhoslyn Prys
- Rhys Dafis
- Rhys Jones
- Richard A Wenner
- Richard Sheppard
- RNIB Cymru
- Roger Matthews
- Ruth Harries
- Steffan Cravos
- Stuart Hodges
- University of Wales, Swansea
- Yamaguti Daigaku  
Jinbungakubu