Welsh Government
Consultation Document

National policy on Welsh linguistic infrastructure

Date of issue: 16 March 2021
Action required: Responses by 6 July 2021

Mae'r ddogfen yma hefyd ar gael yn Gymraeg.
This document is also available in Welsh.

Overview

National policy on Welsh linguistic infrastructure

In July 2017, the Welsh Government published its national strategy *Cymraeg 2050: A million Welsh speakers*. As the strategy's title suggests, we want to see the number of people who can speak Welsh increasing to 1,000,000 by 2050 (from the number recorded in the 2011 Census, which was 562,000).

The strategy also anticipates that the percentage of people speaking Welsh daily, and who can speak more than a few words of Welsh, will increase from 10 per cent (in 2013-15) to 20 per cent by 2050.

To increase the number of Welsh speakers, and ensure they're able to use Welsh in a variety of situations, a modern and robust infrastructure is essential.

To support this, we've developed this draft national policy which proposes ways of improving how the area works for the user.

How to respond

Responses to this consultation should be e-mailed/posted to the address below to arrive by **6 July 2021** at the latest.

The consultation includes a set of questions. We'd welcome your responses to them.

We welcome responses in English or Welsh.

You can respond in one of the following ways:

Online:
Fill in the response form online by following the following link: gov.wales/consultations

Email:
Fill in the response form and email it to UnedIaithGymraegWelshLanguageUnit@llyw.cymru. (Remember to include 'National policy on Welsh linguistic infrastructure' in the subject line)

Post:
Fill in the response form at the end of this document and post it to:

Welsh Language Division
Welsh Government
Cathays Park
Cardiff
CF10 3NQ

Further information and related documents	**Large print, Braille and alternative language versions of this document are available on request.**
	The consultation documents can be accessed from the Welsh Government's website at gov.wales/consultations
Contact details	For further information:
	Welsh Language Division Welsh Government Cathays Park Cardiff CF10 3NQ
	email: UnedIaithGymraegWelshLanguageUnit@llyw.cymru

General Data Protection Regulation (GDPR)

The Welsh Government will be data controller for any personal data you provide as part of your response to the consultation. Welsh Ministers have statutory powers they will rely on to process this personal data which will enable them to make informed decisions about how they exercise their public functions. Any response you send us will be seen in full by Welsh Government staff dealing with the issues which this consultation is about or planning future consultations. Where the Welsh Government undertakes further analysis of consultation responses then this work may be commissioned to be carried out by an accredited third party (e.g. a research organisation or a consultancy company). Any such work will only be undertaken under contract. Welsh Government's standard terms and conditions for such contracts set out strict requirements for the processing and safekeeping of personal data.

In order to show that the consultation was carried out properly, the Welsh Government intends to publish a summary of the responses to this document. We may also publish responses in full. Normally, the name and address (or part of the address) of the person or organisation who sent the response are published with the response. If you do not want your name or address published, please tell us this in writing when you send your response. We will then redact them before publishing.

You should also be aware of our responsibilities under Freedom of Information legislation

If your details are published as part of the consultation response then these published reports will be retained indefinitely. Any of your data held otherwise by Welsh Government will be kept for no more than three years.

Your rights

Under the data protection legislation, you have the right:
- to be informed of the personal data held about you and to access it
- to require us to rectify inaccuracies in that data
- to (in certain circumstances) object to or restrict processing
- for (in certain circumstances) your data to be 'erased'
- to (in certain circumstances) data portability
- to lodge a complaint with the Information Commissioner's Office (ICO) who is our independent regulator for data protection.

For further details about the information the Welsh Government holds and its use, or if you want to exercise your rights under the GDPR, please see contact details below:
Data Protection Officer:
Welsh Government
Cathays Park
Cardiff
CF10 3NQ

e-mail: Data.ProtectionOfficer@gov.wales

The contact details for the Information Commissioner's Office are:
Wycliffe House
Water Lane
Wilmslow
Cheshire
SK9 5AF

Tel: 01625 545 745 or
0303 123 1113
Website: https://ico.org.uk/

Contents

Ministerial Foreword

Where do you look when you want to find out what a word means, or what it is in another language? On one level it's quite a simple question, but if you speak Welsh the answer can be more complicated.

I know there are many resources available – be they dictionaries or terminology databases – but sometimes that can be a part of the problem. So when I'm writing a speech or helping my children with their homework, we'll frequently find ourselves hopping from one website to another to find a word or term. And then not every resource will have the answer to every query, or an answer in one place will be different to one in another. I know that some people are increasingly turning to websites like Bing and Google Translate, but of course the answers we get from these sources aren't always reliable.

We have many projects of the highest standard here in Wales, which can help all of us use Welsh. For instance, specialists have been developing Geiriadur Prifysgol Cymru (the standard historical Welsh dictionary) since 1921, and it's become the foundation for all other work on words and terms in Wales. Other resources, such as Bangor University's Porth Termau, and the Welsh Government Translation Service's BydTermCymru, have appeared more recently in response to gaps in the provision and the demand for contemporary terms. These databases have established themselves and are now extremely influential. But the field remains reactive, with terms commissioned if there is a great demand for them, and no one keeping a high level, strategic overview of future needs.

The question arises, therefore – is this good enough? When we speak about increasing the number of Welsh speakers, and increasing use of the language, we speak about making sure that everyone can use Welsh without having to overcome any extra barriers. That means being able to find information through Welsh, and about Welsh, without having to search for it, and without having to take extra steps compared to people who are trying to do the same thing in English. In that respect, linguistic infrastructure – as the name suggests – is essential for providing the foundation for Welsh to thrive in future. Without it, the language turns into a *Wild West* without rules and without authoritative sources of information (what's the Welsh for *Wild West* again?!).

For several years, there's been a great deal of thought and discussion surrounding the future of linguistic infrastructure in Wales, and in this document we attempt to build on foundations laid by the Welsh Language Board and the Welsh Language Commissioner. The main issue that needs attention has remained pretty consistent: confusion for the user as a result of a number of different sources for words and terms (a number of you have said that you don't always know which one to use; others aren't aware that some of them exist at all), and that those sources can offer suggestions that contradict each other. We want to help to change that.

At present, there is no single body or team responsible for having an overview Welsh linguistic infrastructure. The Welsh Government provides funding for a number of different projects, but there's no formal coordination.

It's important to note that Welsh isn't the only language facing such challenges, but it's fair to say that the main gap in Wales is this element of coordination of work and resources, so that we can also more strategically disseminate and give access to those resources. By doing that better, we're confident that there's potential to improve the provision for everyone who wants to use Welsh.

Over the following pages, we set out a general background to the different components of Welsh linguistic infrastructure (dictionaries, terminology, corpora, standardisation), and outline proposals for the way forward.

It's important to emphasise that the aim of the proposals in this document isn't to start from scratch, but to build on foundations that are already in place. A fact which emphasises how deeply these foundations go in Wales is that Geiriadur Prifysgol Cymru will celebrate its centenary this year. This not only shows how much work has taken place on this valuable resource, but also the importance of ensuring that work continues uninterrupted. The centenary will therefore be an opportunity to look back, celebrate and congratulate, but also to look to the future.

Geiriadur Prifysgol Cymru, as is the case with linguistic infrastructure more generally, has adapted to new digital platforms, and I'm sure that they'll continue to adapt over the coming years. That being so, a number of our proposals tie in with our *Welsh Language Technology Action Plan*, and our aim is to provide, free of charge, digital versions of authoritative resources for everyone who uses Welsh in any way. In future, we will place even greater emphasis on cooperation and new kinds of partnership work, which will include all the organisations with a part to play in planning a better future for the language.

We've already set up a Welsh Language Standardisation Panel to begin standardising how some words are spelled. That's a positive start to work that really needs to be done. The aim of these proposals is to make it easier to use Welsh, in order to help with the journey towards a million speakers. What comes next will be guided by your responses to this consultation.

Eluned Morgan MS
Minister for Mental Health, Wellbeing and Welsh Language

What is linguistic infrastructure?

When we speak about 'linguistic infrastructure', we're speaking about those things that help us use Welsh from day to day. The bricks and mortar such as dictionaries and terminology resources, and all the research and standardisation that happens behind the scenes to enable these resources to grow and develop. These days, many of us use these resources online.

Glossary

Corpus (plural: corpora)
A large collection of linguistic data (e.g. words, terms, phrases) assembled following certain principles. Corpora are used to create digital resources and to conduct linguistic research in order to develop terminology databases or dictionaries. They can be based on sound recordings, written texts or human gestures (i.e. sign language).

Dictionary
Traditional dictionaries usually describe the language and record its vocabulary. They can also suggest synonyms, and their editors don't legislate as to which words should be used in different contexts. This is different to terminology, where a term corresponds to one concept only.

Term
Terms are words or phrases which have specific meanings in specific contexts. These meanings can be different to the meaning these words have in everyday language.

Terminology
A collection of terms relevant to specific subject areas (e.g. health, education). A terminology shows which terms are the standard ones to use in a specific context e.g. when writing for a specific sector or customer. Terminology work is different to that on a dictionary, as it means studying concepts, conceptual systems and their labels, rather than studying words and their meanings as happens when developing a dictionary.

Section 1 – Executive summary

During the summer of 2017, we published a strategy called *Cymraeg 2050: A million Welsh speakers* which introduced a plan for doubling the number of Welsh speakers to a million. The strategy entailed a significant change to the Welsh Government's policy for the language.

In the strategy we stated that, in order to reach the target of a million speakers, we need more people to speak Welsh and more people to use it. But we also said that we need to create favourable conditions – things like linguistic resources, a prosperous economy, or ensuring people's support for the language. Resources such as dictionaries, terminologies and corpora (things we refer to here collectively as 'linguistic infrastructure') are an extremely important part of this.

As increasing the number of speakers means including everyone, from all backgrounds, we want to see these different elements of linguistic infrastructure being coordinated effectively. We want to do this so that everyone can use them easily – be they new speakers, school-children and their parents, students, people who use Welsh in the workplace, professional translators, social care and health professionals, lawyers, the police and emergency services, journalists, teachers, and many more.

Our priorities

With so many excellent resources available to people who want to use Welsh, it can be frustrating when people say they don't know where to look for a word or term. The truth is that these resources are scattered around (some in digital form, some in the form of books), and people don't always know where to find answers or guidance. We therefore want to create a new way of coordinating the main resources and providing them in one place, so that everyone can easily find answers. We propose doing this by creating a new unit in the Welsh Government to help the different elements work together better. The unit will also be responsible for running an easy-to-find, easy-to-use website for everyone who wants to use Welsh, giving easy access to resources such as the main dictionaries and terminology databases, as well as those such as corpora, all in one place.

With this in mind, this document outlines our proposals under the following six areas of work:

1. **Dictionaries**
2. **Terminology**
3. **Corpora**
4. **Standardisation** (e.g. orthography/spelling, place-names, terms)
5. **Creating one interface to make it easier to search through all resources**
6. **Setting up a new unit to coordinate Welsh linguistic infrastructure**

Our aim is to make sure that all the elements we've grouped together under the title 'linguistic infrastructure' work together smoothly, in order to make it easier for more people than ever to use Welsh in their day to day lives.

Section 2 − Introduction

Why have the responsibility for coordinating these elements in one place?

Making it easier for the user: People who want to use Welsh express a number of concerns. One of them is that they find it difficult to find resources. Another is that they're often confused due to the number of terminology resources and dictionaries available, and find it difficult to know which one has priority. Even though these concerns contradict each other to a degree, they suggest that people are unclear about where to find answers to linguistic questions. Creating one unit to be responsible for making the different resources work together better, and providing access to them through one website, is a way of getting to grips with this.

Ensuring that resources are better coordinated would also lead to more clarity for the user, especially if all their output (be they words or terms, corpora or linguistic advice) was presented through one searchable portal, which could with time be tailored to users' needs.

Creating a consistent experience for the user: When we're searching for a word in English, a number of us will turn to the same resources. For instance, when searching for a word, you can turn to a dictionary such as the *Oxford English Dictionary* and be confident that it contains authoritative answers about a word's meaning, spelling or origin; or if you're searching for more practical explanations and examples of how to use the word in a sentence, to the *Cambridge Dictionary*. This helps to create a unity of experience where a historic, national institution gives the user the feeling that they're drawing from one standard source – this in turn gives them confidence that they're using the right word.

A number of nations and languages have official bodies or academies of different models, which are responsible for coordinating work on linguistic infrastructure, and which act as an obvious starting point to anyone with a question about the language. In the case of Welsh, the experience can depend on what the user already knows about the resources which are available. For instance, a professional translator will know which source, from among many, to choose; some Welsh speakers will have some knowledge, but will often be confused or unaware of a resource which fits the bill; and often new speakers or people who don't speak Welsh won't know where to start.

We believe that everyone who uses Welsh to be on the same journey with the language, and that by working together we can reach a million speakers. Offering one consistent experience to everyone, be they new speakers or more experienced with the language, is therefore extremely important when it comes to linguistic resources. The unit mentioned above will aim to make it easier for people to use the language by commissioning and coordinating resources, and by marketing them much more effectively to give users confidence that the answers they find are authoritative.

Avoiding duplication and gaps in provision: The duplication we refer to above can not only cause confusion, but also lead to inefficient use of resources and specialist skills.

At present, there are several dictionaries and terminology databases available with different people and organisations responsible for them, with very little strategic coordination between them. This can lead to resources whose results contradict each other, and inconsistent or no provision at all in other areas. We believe that with better coordination, there will be less duplication of resources, more consistency, and therefore less confusion

to the user. The Welsh Government already provides funding for several dictionary and terminology projects, so we see this as an opportunity to play a coordinating role in order to further develop provision.

Horizon scanning: On one hand there is duplication of resources, and on the other a lack of strategic planning concerning future requirements. For instance, if it becomes evident that work is being planned in a new policy area, or a certain industry will be locating to Wales and Welsh terms will be needed for the workforce, who is responsible for ensuring that terms are standardised in time? We saw something similar happening recently with attention worldwide to the equalities agenda in relation to race and ethnicity resulting in highlighting gaps in standardised terminology in Welsh. This is the result of a systemic failure, as no one body in Wales is responsible for having a strategic oversight of Welsh language terminology. If we look at the two main centres which standardise terms in Wales, the Welsh Government's Translation Service respond to Welsh Ministers' plans and use of terms through the BydTermCymru website, while Bangor University's *Terminology and Language Technology Unit* develops new terminology dictionaries for their Terminology Portal, in tandem with the organisations who commission the work. These are their roles, and they do the work to an extremely high standard.

We propose that one body should be responsible for scanning the horizon for policy developments and large public projects which are in the pipeline, and for ensuring that appropriate experts are responsible for proactively producing and standardising the necessary terminology. This is especially important in the case of technical terms, to ensure that they're conceptually correct and appropriate for specialist use, be that in education, science, medicine or any other field.

As well as planning for the long term, we also need a way of helping with high profile terms that arise without warning, but for which there is a pressing need. We therefore want to reduce the number of situations where individuals and organisations have to urgently come up with new terms without wider consultation or guidance. That can lead to duplication, as different bodies or agencies, because of the pressing nature of their work, set out to create new terms in isolation, often leading to several bodies creating and disseminating a number of terms for the same concept. We believe that helping to coordinate terminology in cases such as these will lead to more consistency, and make it easier for users to recognise, understand and begin to use new terms.

Coordinating for the benefit of all: Although there is some collaboration between the organisations and teams responsible for different elements of Welsh linguistic infrastructure, there is no structure to ensure that developments and lessons learned from one area are shared with others as a matter of course.

Much of the present work on linguistic infrastructure is publicly funded, therefore establishing a new way of coordinating it would be a way of ensuring better value for money – reducing duplication and ensuring that the funding, resources and expertise are used where they're most needed. Ensuring sufficient and stable investment in this area of work, as well as ensuring that projects which have already received public funding remain current and valuable, will also be essential.

In the short term, one possible challenge may be a paucity of individuals with the requisite expertise to do the work – there are only so many people in Wales with experience of developing dictionaries and terminologies, for instance. It will be the unit's responsibility,

therefore, to plan for the longer term, and to explore ways of fostering expertise and expanding the pool of experts who can do the work.

It would also be beneficial for the new unit to work closely with translation services of all kinds, as well as Cymdeithas Cyfieithwyr Cymru (the association of Welsh translators and interpreters) and officials who lead on Welsh language technology at a local level, to ensure that the different areas are joined-up and that any developments in one area are considered in others so that they all benefit.

Section 3 – Areas of development

In this section, we propose six areas to be developed in order to strengthen and better coordinate Welsh language linguistic infrastructure. We believe that doing this will help increase use of the language and with the effort of reaching a million speakers by 2050.

Area 1: Dictionaries

"What's the Welsh word for...?" is a question we hear all the time. On the radio or television, in the classroom, on social media or in the workplace – questions about Welsh words are very common. And often, without anyone realising it, the answer already exists in a dictionary, usually within easy reach online and through apps.

It's easy to take dictionaries for granted. They play an important role in making sure that information about the definition of words, and their translations, are available to their users. In the modern world, different kinds of dictionaries can record the historical use of a language, reflect every day usage, or provide for specific audiences (such as new speakers).

Safeguarding the future of dictionaries like these, and making sure they're easily available to all, would remove barriers and enable more people to use Welsh. Dictionaries are also a strong foundation for developing all sorts of other useful linguistic resources for everyone who wants to use Welsh.

What resources are available now?
There are several Welsh dictionaries available. Below we focus on those that are available free-of-charge online.

Geiriadur Prifysgol Cymru (gpc.cymru)
This is the standard historical Welsh dictionary, originally prepared between 1921 and 2002 by the University of Wales, and which continues to be updated. It is similar in status to the *Oxford English Dictionary*. This is the foundation for all other reference books for the Welsh language e.g. dictionaries, terminology work, grammar books, decisions on orthography (i.e. how words are spelled) and so forth. This resource is also available in the form of an app.

The Welsh Academy English-Welsh Dictionary (geiriaduracademi.org)
An English-Welsh dictionary, edited by Bruce Griffiths and Dafydd Glyn Jones. Before its publication, there was no comprehensive dictionary for use by the translation profession. It has also been an extremely important resource to help people work through the medium of Welsh in bilingual organisations, and in the field of education. The Welsh Language Board, and more recently the Welsh Language Commissioner, have been responsible for a project to convert the dictionary to digital format and publish it online, but the contents were not updated as part of this work. The contents have therefore been static to all intents and purposes since the second print edition in 2003 – a period of swift change for the language – and as a result there are many more recent words which have not been included, or which need to be updated.

University of WalesTrinity Saint David's Welsh Dictionary
(https://geiriadur.uwtsd.ac.uk/)
A comprehensive Welsh-English and English-Welsh dictionary, prepared by the Welsh Language Department at the University of Wales Trinity Saint David in Lampeter.

Geiriadur Prifysgol Bangor (geiriadur.bangor.ac.uk)
Another comprehensive Welsh-English and English-Welsh dictionary prepared by the *Terminology and Language Technology Unit* at Bangor University. It is especially useful to new speakers. It includes links to other dictionaries and technical terminology resources. It is also available free-of-charge on the *ap Geiriaduron app*.

The *Gweiadur* (www.gweiadur.com/en/Pawb)
A Welsh-English and English-Welsh resource developed by D. Geraint Lewis and Nudd Lewis, which is in a beta version at present. It can be used free-of-charge, but users need to register to do so (registrations are currently closed).

We also need to acknowledge that many people use resources like Bing, Google Translate and Microsoft Translator to find words (or terms). Although these can be useful tools, they don't always provide reliable answers.

Objectives:

Ensuring consistency between different dictionaries, where appropriate, for instance by making sure they're updated on the basis of the same research, so that they also enrich each other.

Making sure that everyone knows where to go to find the definition or translation of a Welsh word.

Ensuring that resources that already exist are regularly updated, so that they reflect Welsh as a living language.

Ensuring that a modern Welsh>English dictionary, which is regularly updated, is available for everyone to use digitally.

Making the most of the expertise that already exists, and further develop it.

How:
- Bring some of the main Welsh dictionaries together in one place (where appropriate), in order to provide long-term editorial care and make the most of the expertise that exists in the field. This would also bring value for money.

- Ensure that consistent and complementary dictionary resources are available, so that people who use Welsh have access to comprehensive provision.

- As part of this, take steps to resume the work of amending and updating *The Welsh Academy English-Welsh Dictionary*.

- Prioritise developing and maintaining a modern online Welsh>English dictionary, to fill the gap for such a dictionary that exists at present.

- Ensure that users can access the main dictionaries through a central website (see area 5 below) by marketing and raising awareness of it so that people know where to turn when the question "What's the Welsh word for...?" arises.

- Improve the training available for those who wish to be lexicographers, in order to increase the number of people able to do the work.

Area 2: Terminology

Just like the words in a dictionary, we also depend on terms in our everyday lives – whether we realise it or not.

Terms are used to denote specific concepts in specific fields and, unlike words in more general usage, standardisation processes are required to agree on one term for each concept.

For instance, we can't use a computer without a *mouse* and *keyboard*, or *inputting* a *password* – these are all terms that are connected to the field of information technology. And in collecting terms that are connected with one field together in one place, we create a list of terms or, in other words, a *terminology*.

Mouse has more than one meaning, but by including it in a terminology for the field of technology, we know that it refers to the tool that helps us work on a computer, rather than a furry little mammal with a long tail. In this way, terminology dictionaries are completely different to descriptive dictionaries like the ones we speak about in Area 1 above. But they're just as important. Indeed, it's essential that standardised sources of terminology exist to make it easier to use Welsh in every aspect of day to day life, including in specialist fields like technology, law, health and education.

The work of standardising Welsh terms has been happening since the mid-twentieth century (e.g. the volume *Termau Technegol*, University of Wales Press, 1950). In the field of education, the work of standardising terms started in 1993 when the Schools Curriculum Authority awarded a contract to Bangor University to standardise Welsh terms for schools in Wales. This led to the creation of *Y Termiadur Addysg,* which continues to be an extremely important resource. But the need to create and standardise terms never goes away – recently, for instance, we've seen the need to update terminology in relation to race and ethnicity as a result of the increased attention to this area.

Terminology dictionaries are based on concepts rather than words, and follow a standardisation process set out by the ISO/TC37 family of international standards. This means it can take a little more time to standardise terms and include them in terminology resources. It's important, therefore, that terminology work is coordinated and planned thoroughly to make sure there are no obvious gaps in the provision.

There are two main issues in relation to terminology:

1 – There are so many resources – some standardised, others not so – meaning that people don't always know which ones to use, and which can occasionally lead to inconsistencies between different resources

Although there are many online terminology resources of the highest standard available for Welsh – be they terminology databases produced by terminology experts, or online lists kept for ease of reference by different public bodies – it's evident that people don't always know where to look for a term, or which resource is best for their needs.

As the work of standardising terms takes place in a number of contexts, by individuals and organisations, in order to express a technical concept, this can lead to a situation where a number of different sources suggest terms, some which have been standardised, and

others that haven't. To those who work in a specific field, it can be obvious which source is the most reliable one, and which term to use. But what about the rest of us?

Our objective is to make sure that everyone can use Welsh easily, without any barriers, and terminology is part of this picture. If people don't know where to go to find a term, or if the results differ from source to source, that creates a barrier, and means that people aren't confident in which term to use.

2 – Coordination is needed to prioritise which technical terms to standardise, as well as a better process for creating and mainstreaming new terms which are needed urgently

Broadly speaking, the work of standardising terms can be separated into two activities:
- **filling gaps in the language's vocabulary**, for instance so that words and terms are available for new concepts as they arise.
- **standardising packages of technical terms**, so that those terms are appropriate for specialist use in specific fields e.g. education, science, medicine, food hygiene etc.

Filling gaps in the language's vocabulary: from time to time, a new term will urgently need to be created, as a result of something that's happened in the news, for instance, or because new information needs to be given to the public. In Wales, this usually happens to try to convey in Welsh concepts that have initially been expressed in English. Recent examples have included terms concerning Brexit, race equality, and medical terms as a result of coronavirus.

One scenario – recently, because of the coronavirus pandemic, a period of lockdown was instigated in Wales that was referred to in English as a 'firebreak'. However, in Welsh a number of variations arose, with different media outlets using translations that weren't consistent with that used by Welsh Ministers, who also occasionally found themselves using inconsistent terminology. Each Welsh term used was valid in and of itself, but in a context where clarity was extremely important, ensuring a certain amount of consistency would have been beneficial – not only ensuring people knew what was being referred to, but also increasing people's confidence to use or read Welsh in specific situations (after all, *firebreak* was used by everyone in English, and everyone knew what the term referred to).

Because languages evolve all the time, this work is continuous. However, perhaps because of a lack of resource, to date there has been very little planning to ensure that gaps are filled systematically, nor to ascertain how the new terms that do arise as a result of this take root.

Standardising packages of technical terms: on other occasions, terms aren't needed quite so urgently, but it's still obvious that a package of new terms will be needed in the near future e.g. if the Government intends to develop a new policy in a certain field, if a specific subject garners public attention (e.g. race and equality), or if a new industry comes to Wales and new Welsh terms are needed. But with two large projects creating terms (the *Terminology and Language Technology Unit* at Bangor University, and BydTermCymru in the Welsh Government's Translation Service), who leads on planning the provision when a series of new terms arise, and who decides which terminology packages to prioritise?

In both cases above: if more than one source creates a separate new term for the same concept, which one has priority? How is the user supposed to understand that several Welsh terms refer to just one in English? Which one should they use?

What resources are available now?

There are several sources of Welsh terminology – below we list the main ones:

Welsh National Terminology Portal (termau.org)
This terminology resource was developed by the *Terminology and Language Technology Unit* at Bangor University. The Welsh National Terminology Portal includes collections of specialist terminology which have been standardised in collaboration with a number of bodies, for instance:

- Y Termiadur Addysg (terms for education)
- The Coleg Cymraeg Cenedlaethol's Terms for Higher Education
- Social Care Wales terms
- Justice Wales Network terms
- Edward Llwyd Society nature dictionaries
- Food Standards Agency terms

Some of these resources can also be accessed through Cysgeir (part of the Cysgliad software package) and the Ap Geiriaduron app.

TermCymru (gov.wales/bydtermcymru)
A searchable collection of terms that Welsh Government translators use in their everyday work. It is regularly updated with a view to providing a comprehensive database of standardised terms which reflect current usage in different aspects of the Government's work. As well as being able to search the database, you can also download it in its entirety under open licence, which gives more flexibility than offering it in searchable form only.

The Coleg Cymraeg Cenedlaethol's Terms for Higher Education
(colegcymraeg.ac.uk/cy/adnoddau/termau/)
A database of terms for Welsh-medium Higher Education. This is an online resource for students and staff to facilitate their studies, education and research through the medium of Welsh at University level. This terminology can also be searched through the Welsh National Terminology Portal mentioned above.

Y Termiadur Addysg (www.termiaduraddysg.org)
A Welsh Government-sponsored resource, developed by the *Terminology and Language Technology Unit* at Bangor University, it provides standardised terms for the education sector. These are the standardised terms used in Welsh-medium exam papers and assessments, as well as resources of all sorts for teachers and students. This terminology can also be searched through the Welsh National Terminology Portal mentioned above.

Other terminology collections: Various other bodies also draw up and maintain their own lists of terms, occasionally publishing them online. In most cases, these won't have been standardised using recognised processes, or been fed into the resources above. In the absence of other sources, however, they can be very useful – examples include lists kept by Stonewall and National Museum Wales.

Objective:

Ensuring that everyone can easily find a term, and have confidence that it's the right one for them.

How:

- In order to avoid duplication and help to work towards more consistency, set up a new unit (see area 6 below) to coordinate terminology work (as well as work on dictionaries). The unit would also be responsible for drawing up annual work programmes to prioritise terms to be standardised.

- One long-term objective would be to enable users to search through the main Welsh terminology databases (i.e. BydTermCymru and Welsh National Terminology Portal) in one place, through an easy-to-use website (see area 5 below). This would mean that users don't have to search through a number of different sources on the internet, and give them confidence that the term suggested is approved and standardised.

- Ensure that users know exactly what resource is appropriate for their needs e.g. if they're a school pupil, using Welsh at work, or a new speaker. This will call for user-led research to find out who the different audiences are, and under which circumstances they're likely to search for a term.

- Set up a service to help people find a term, or to decide on an appropriate term when there are several choices. One possible answer would be to create a service to act as a point of contact for queries relating to urgent and high profile terminology, with staff able to consult with a network of subject specialists and members of the Welsh Language Standardisation Panel (see area 4 below) as needed.

- As part of the process of disseminating terms, and to improve our understanding of the kind of terms people are likely to use regularly, monitor and evaluate to what degree terms are accepted, thereby learning lessons and feeding the findings back into the process.

- In order to ensure that everyone knows where to find terms, it will be essential to market the provision effectively, raise awareness of what's available, and offer advice on the best places to find answers, depending on who the user is.

- Increase the number of professional terminologists through a combination of training, marketing the opportunities available in the field, and apprenticeships.

Objectives:

In the case of gaps in the language's vocabulary, resulting from concepts that arise without warning and where a solution is urgently required – establish a quicker process for creating new terms, in a way which responds to events as they arise.

These urgently created terms must have status, and be shared proactively between sources so that as many organisations and agencies as possible recognise and use them. This may mean including representatives from different places to agree on a term that is urgently needed. By getting the main stakeholders, including perhaps members of the press, to

agree on a term from the outset, it follows that there will be more use of the term agreed upon.

In the case of technical terms it can be foreseen will be needed in the near future (e.g. because of a policy in development, or as a large employer wishes to open a new workplace and terms are needed to meet demand in a particular sector) – annual work programmes are needed to coordinate this important work.

In both instances above, we propose that the terms standardised, from whatever source, are also provided through a central easy-to-use website.

Taken together, this will mean better coordination of provision to avoid duplication of effort, thereby reducing confusion for the user and increasing their confidence to use the language (although we accept that occasionally different terms are needed in different contexts e.g. in legislation or for the general public).

How:

- Set up a unit (see area 6 below) to identify gaps in terminology that need to be rectified urgently in the short term, and to scan the horizon for terms that will be needed in the medium-term.

- In the case of urgently needed terms, establish a process for agreeing which term to use – this will help to ensure the term has status, with the aim of leading to more consistency in its use. As a part of this, a method of standardising urgent terms will need to be agreed upon, which follows recognised processes, and of disseminating the term without delay to the people most likely to need it.

- Regarding terms where there is less urgency, explore what methods would be best for standardising terminology packages e.g. convening a group of experts at regular intervals, or asking a centre or organisation to lead on the work.

- Ensure that the terms agreed upon are not only included on the central website (see area 5 below), but are also disseminated and included in other appropriate terminology databases, so that anyone searching for a Welsh term can rest assured they're using the appropriate term.

- Give publicity to new terms which are standardised, so that users are aware of them.

Area 3: Corpora

A 'corpus' (plural: corpora) can mean several things. But what we have in mind in this document is a database of linguistic information collected following specific principles, which is used to conduct linguistic research and create resources. A corpus such as this can be based on written text or recordings of oral language, or a combination of both.

Linguistic work of all kinds is done on the basis of corpora – for example, a corpus can provide evidence for people developing a dictionary, a terminology, or a language technology resource. How the data in a corpus is collected depends on what we want to get out of it e.g. sound and transcription corpora will be needed to train Welsh speech technologies and artificial intelligence.

We can also use them to research how to develop resources for different kinds of Welsh speakers. This includes new speakers and school children (as they can show users how certain words or expressions are used in a sentence, and how often they appear), as well as publishers, lexicographers (i.e. people who develop dictionaries), researchers and the media.

As a result, it's important to make sure that a variety of tagged corpora are available for everyone to use – whether by individuals or technology. If we can also ensure that they're available through open licence, which will mean structured data is available to those who wish to develop different resources e.g. Welsh language systems, speech technologies or artificial intelligence. We will place an emphasis on this in future, in order to ensure that resources, especially those developed using public funds, are available to all.

As this suggests, corpora are also especially important for creating and maintaining a language's digital infrastructure. The aim of our Welsh Language Technology Action Plan[1] is for technology to support the language, so that it's used in as many situations as possible. We're aware that technology develops quickly, and we want to make sure that Welsh moves with it. Our proposals in this new linguistic infrastructure policy will mean, for instance, that corpora are developed (including those that feed dictionaries and terminology databases) that will provide more structured data to support the Welsh Language Technology Action Plan, and enrich the experience of those using technology in Welsh. We need massive Welsh language data sets and relevant corpora to develop speech-to-text capability, new translation systems, and machine learning. To do that, we need to ensure that academic research and development in the field of language technology is supported in the long term. This infrastructure needs to be maintained and updated for future generations.

What resources are available now?

Sometimes corpora projects are planned in advance in order to satisfy specific needs, and sometimes they're created as a by-product of other activities (e.g. developing dictionaries or terminology databases). There are a number of Welsh or bilingual corpora, which are the result of different research projects over the years. Indeed, any large collection of words, terms or phrases that have been tagged is a corpus. Behind any dictionary or terminology, there will be a corpus, but these aren't always available for everyone to use because of intellectual property constraints.

---

[1] https://gov.wales/sites/default/files/publications/2018-12/welsh-language-technology-and-digital-media-action-plan.pdf

The main corpus resources available to people who use Welsh are listed below (not all of them are necessarily available to download through open licence).

The National Corpus of Contemporary Welsh (CorCenCC: Corpws Cenedlaethol Cymraeg Cyfoes (corcencc.cymru)

The CorCenCC project involves Cardiff, Swansea, Lancaster and Bangor Universities. It includes samples of contemporary Welsh which happens naturally in real-life communication, in the form of oral, written and electronic data (e.g. Welsh used on social media), presented in the form of an online, searchable text corpus. It is accompanied by an online teaching and learning toolkit, which draws on data from the corpus to offer resources for learning Welsh to people of all ages and abilities.

Welsh National Corpora Portal (corpws.cymru)
A collection of online written Welsh and bilingual corpora, maintained by the *Terminology and Language Technology Unit* at Bangor University. It provides an interface which can be used to search for words, terms and phrases, which are displayed in their context. Also, if a translation of the text is available, it can show it side by side with the original. The interface includes software called a *lemmatizer*, which can find Welsh words in any form (e.g. if they've been mutated, or are shown in plural form). The Corpora Portal also includes links to other corpora that could be of interest to researchers. Some of these were produced by Bangor University, and some by other organisations.

Objectives:

Ensuring that extensive and high quality corpus resources are available for everyone to use through open licence. It's possible that this will happen organically as a by-product of activities which are already taking place, but coordinating the field and maintaining dialogue with the main providers will help ensure that appropriate corpora are available for any new developments in relation to the Welsh language.

In order to support new speakers, ensuring that a corpus of common forms and mistakes is available for use by organisations such as The National Centre for Learning Welsh, to help with their learning programmes.

Keeping an eye on the gaps that arise in the kinds of corpora that exist, and commission, develop and curate as needed. As a rule, collecting and creating data takes time, therefore we need to make sure that these activities, as well as efforts to process, enrich, annotate and label the data, are focussed.

On a central website, creating and maintaining links to authoritative corpora, including: lexicons, Welsh/English sentences which have been checked and aligned, place-names and street names, lists of named entities, sound corpora which have been transcribed, and so on, in order to enable dictionary and terminology projects to continue to develop on the basis of evidence.

How:

- Give the new unit (described in area 6 below) the task of overseeing the field and ensuring that appropriate corpora are available to use for various needs, facilitating provision where needed.

- Explore how to safeguard corpora that already exist, for instance by providing a facility to keep back-ups of them, to reduce the risk that resources developed with public funds are lost.

- Maintain links to corpus resources on one website (see area 5 below) in order to help peope find corpora that are appropriate for their needs.

- Take steps to help expand the corpora that already exist, for instance by ensuring that public documents are readily available to enlarge them.

- Take steps to enrich corpora by coordinating the relationship between different public providers, and ensure that opportunities to do so are not lost.

- Publish guidance on how to use different corpora for different purposes.

Area 4: Standardisation

Several different kinds of standardisation take place in relation to the Welsh language, but in general they all aim for consistency in the way Welsh is used, by ensuring rules exist where they're needed.

The different kinds of standardisation include:

- standardising orthography (i.e. how we spell words)
- standardising grammar (e.g. whether words are masculine or feminine, what their plural forms are, etc.) and style
- standardising place-names (the Welsh Language Commissioner leads in this area)
- standardising terms (we address terminology under area 2 above).

Standardising orthography

Byrger or *byrgyr* (for 'burger')? *Cydbwyllgor* or *Cyd-bwyllgor* (for 'joint-committee')? *Microffon* or *meicroffon* (for 'microphone')? Which one should we use?

At present, the spellings for a number of Welsh words can vary, and this leads to inconsistency between different sources, as well as uncertainty for the user. Making sure that the language's orthography (the way Welsh words are spelled) is continuously standardised is important if we want to see more consistency, for instance between different dictionaries and terminologies, and to give more certainty to people who want to use the language. And it's essential if we want to create a website to be a portal to Welsh language linguistic resources, as described in area 5 below.

To get to grips with this, we've set up a Welsh Language Standardisation Panel. The Panel includes academics, terminologists, lexicographers, translators and linguists who will standardise and modernise Welsh orthography.

The Standardisation Panel will have a key role in making sure that infrastructure providers (for instance, those who maintain dictionaries and terminology databases) recognise decisions concerning orthography.

Standardising grammar and style

After the initial work of standardising the most pressing variations in orthography, the Welsh Language Standardisation Panel will consider the next set of priorities. These may include standardising grammar or producing style-guides. This could lead to further advice and guidance being available on using the language, thereby giving clarity to anyone who wants to use Welsh.

With time, this standardisation work could lead to developing easy-to-use language guidance that would be available online. This could include guidance, advice and answers concerning, for instance, conjugating verbs, mutations, appropriate use of register in certain contexts, and so on, on a central website (see area 5 below). The coordination unit (proposed in area 6 below) would be able to develop this aspect of the work in the long term.

Standardising place-names

The Welsh Language Commissioner leads on place-name standardisation, by convening a Place-name Standardisation Panel which focusses on the spellings of villages, towns and cities in Wales.

The Panel includes specialists on place-names and Welsh orthography, and follows specific guidance in order to suggest standard forms for place-names. One part of the work is to maintain a List of Standardised Welsh Place-names, which is an online resource that can be searched or downloaded (due to technical issues, the List is currently unavailable online; if members of the public have a query about a place-name, they can contact the Commissioner's office to discuss).

We recognise that standardising place-names is different, in terms of process and status, to other kinds of standardisation mentioned in this paper. Place-names are included here as they're part of the wider standardisation picture. But we believe that the work which we propose in relation to orthography, grammar and style, in tandem with work on place-names, will raise the status of standardisation in general. In addition, in collaboration with the Welsh Language Commissioner, it makes sense that the unit proposed in Area 6 below would have a role in emphasising the status of standardised place-names, and further raising awareness of the List.

What resources are available now?

The lack of an authoritative body to lead on standardising different aspects of the Welsh language is a gap that has been recognised for some years. At present, no organisation or body leads on coordinating standardisation. Enabling the newly-established Welsh Language Standardisation Panel to continue its work regarding orthography (as well as, in time, other linguistic matters) would go a long way towards creating a consistent experience for the user, and make it easier for them to use Welsh. It would also be an initial step towards providing as many resources as possible through one website (see area 5 below).

The Welsh Government Translation Service's BydTermCymru includes a useful style-guide[2] which aims to help ensure consistency across the Government's translation work – although the advice can also be of use to other translators and indeed anyone drafting Welsh texts. This style-guide is based on the advice in Gramadeg y Gymraeg (1996) by Peter Wynn Thomas, Orgraff yr Iaith Gymraeg (1987), and Geiriadur Prifysgol Cymru.

Objectives:

Raising the status of different kinds of standardisation in Wales, and coordinate where necessary.

Ensuring that the Welsh Language Standardisation Panel meets the demand for an authoritative voice on orthographical matters.

Ensuring that variations in spelling between different resources are ironed out over time as providers adhere to the Welsh Language Standardisation Panel's decisions.

Ensuring that recognised, modern style-guides are available to everyone who uses Welsh.

---

[2] https://llyw.cymru/bydtermcymru/style-guide (Welsh only)

How:

- Enable the new linguistic infrastructure unit (see area 6 below) to coordinate and have oversight of standardisation needs, as appropriate.

- Support the Welsh Language Standardisation Panel to continue to tackle linguistic issues e.g. orthography/spelling.

- Ensure that the Panel's membership includes a variety of experts who can resolve issues which cause confusion.

- Ensure that the Panel's decisions are reflected and mainstreamed through the main Welsh dictionaries and terminologies.

Area 5: Creating one interface to make it easier to search through all resources

What are we trying to solve?

From what we've described above, it can be seen that one of the main issues we're trying to resolve is that so many resources are located in different places, whether they be dictionaries, terminology databases, or corpora. The everyday user can't be expected to know what status these resources have, nor which is most likely to suggest an answer appropriate to their needs. This can also mean that many people can't find relevant resources, or that they don't know about their existence in the first place – this can be detrimental to both the user and the resource itself.

What are we proposing?

Whether trying to find a word or a term, it makes sense that being able to find all the relevant information in one place would make it easier to find answers in relation to Welsh.

We therefore propose creating one interface or website to act as a portal to everything associated with Welsh linguistic infrastructure. This website could enable you to search easily in several online dictionaries and terminologies at the same time, by using a variety of tools from digital personal assistants to mobile phones to laptops.

Each dictionary or terminology whose resources would be searchable via the website would remain independent projects with their own websites, but the words and terms included in their resources would also be available to search through the website. This would mean that nobody would have to hop from one website to the next searching for an answer. With time, we would also aim to provide this resource through apps to be used on mobile phones or tablets.

Perhaps the first steps towards creating a website would be the initial work to standardise orthography, and considering how best to tag words and terms from different projects in a way that would allow them to be searched through a central interface. Having done that, we would like to be in a position where you can note what kind of user you are (e.g. someone using Welsh at work, a professional translator, school pupil, parent - whether able to speak Welsh or not, teacher, university student, and so forth), or what field you'd like to search, and that the results would be arranged according to relevance. This would give you more certainty that you've found the appropriate answer.

This work would take time, of course, therefore at first we would want the website to offer advice about which sources are most appropriate for different needs, and direct people to the right places. Ultimately, our aim is that you'll be able to search, view and compare results from a number of sources simultaneously.

More than a search engine

But this central website would be more than just a portal for searching words or terms. It could also be a springboard for all kinds of activities in relation to Welsh.

When it's up and running, we could develop the website to help all kinds of Welsh speakers or users. For instance, by drawing people's attention to the fact that they can search for a word or term in a way that recognises mutated words, or that shows how words mutate or

are conjugated in other ways. This could be complemented by the inclusion of various style-guides, as suggested in area 4 above. Materials such as these would be especially useful to new speakers and those using the language every day at work.

The website could also keep a record of the terms and words users search for, in order to identify gaps in provision and help to prioritise what new terms need to be standardised in future. Another way of doing this would be to include a form on the website, for individuals or organisations to request a specific term or word, if it isn't already available on the website. Officials working for the proposed unit (see area 6 below) could assess the request and refer it, if appropriate, to be considered and standardised.

The website would also be an ideal place to publish decisions of the Welsh Language Standardisation Panel in relation to orthography, grammar and style (see area 4 above).

In the case of corpora, it would be relatively simple for the central website to include links to all the appropriate corpora on one interface, or even host some of them if that was desirable, in order to help people find corpora that are appropriate for them. The aim would be to share different corpora so that people can develop new resources. It would be good practice, and ensure value for public money, to work towards a situation where the most up-to-date versions of resources could be downloaded from the website under open licence. Indeed, we propose that the principle of providing resources through open licence is relevant to as many of the resources mentioned in this document as possible.

What resources are available now?

At present, the closest we have to providing every resource on one website is Bangor University Language Technology Unit's website (techiaith.bangor.ac.uk/) which provides access to the Bangor Welsh-English Dictionary, as well as the University's Terminology Portal and Corpora Portal. This includes an extensive and searchable collection of terminology and corpus resources across a number of fields. It doesn't include BydTermCymru, the other main terminology database for the Welsh language, and doesn't enable the user to search all the main online dictionaries, e.g. Geiriadur Prifysgol Cymru.

One lesser-known example that acts as a window to a large number of resources is the European Dictionary Portal (www.dictionaryportal.eu/cy/ctlg/?objlang=cy), which is curated by experts from the *European Network of e-Lexicography*. It offers a way of searching a number of sources in one place, but rather than providing all the results side by side, the user must go to each individual website to see the results. This website also doesn't provide the user with advice about which resource could be most appropriate for them.

Objectives:

Creating one website to enable users to search the main Welsh dictionaries and terminologies simultaneously.

Developing the website to be a hub for all kinds of work in relation to Welsh linguistic infrastructure.

How:
- Appoint language technicians (as part of the process of setting up the unit described in area 6 below) to be responsible for creating the website and running it from day to day.

- Come to an agreement with different projects so that we're able to include search results from their databases through the website, in addition to their own websites.
- Take initial steps to prepare the way, including work to ensure that certain word spellings are consistent, and that words and terms in different sources have been tagged consistently, so that they show up in searches.
- Establish a process for receiving and responding to enquiries submitted via the website.
- Establish a link between the search patterns and enquiries that are submitted through the website, and the standardisation processes outlined in area 4 above.
- Explore the website's potential from the perspective of corpora, standardisation, sharing best practice, tracking how words and terms are used, and so on.

Area 6: Setting up a new unit to coordinate Welsh linguistic infrastructure

What are we trying to solve?

When describing the different resources above, it's obvious that work of an extremely high, professional standard is taking place across all areas of linguistic infrastructure in Wales. What we also see is that there are gaps that could be filled by maintaining more oversight over the field, and by coordinating more effectively what already exists.

For years, decades even, there has been talk about the need to better coordinate Welsh linguistic infrastructure. A number of experts in the field have also agreed with this.

In other countries with languages in similar situations to Welsh, governments and government agencies play a prominent role in planning and coordinating provision in terms of dictionaries, terminologies, corpora, and helping to standardise the language.

We consider it to be appropriate, therefore, to propose establishing a central unit to coordinate dictionaries, terminology, corpora, standardisation work, and to set up a website to provide resources together in one place. This unit could also work closely with a variety of stakeholders in fields such as translation and information technology.

What are we proposing?

In order to maintain an overview of this field, and help to plan it strategically, we propose setting up a unit to coordinate different aspects of Welsh linguistic infrastructure, and to ensure that the provision makes it easier for users to find definitive answers. Such a unit would lend prestige to the field and promote the status of words and terms in Wales, so that there is more consistency of use across different organisations and the media. In turn, this would help people use their Welsh skills more often in more situations.

We consider that the unit would undertake activities such as those described in areas 1-5 in this document.

We propose locating this unit in the Welsh Government. This would enable the unit to benefit from holistic planning, thereby linking in with other relevant policy areas, for instance the Government's Welsh Language Technology Action Plan. It would also mean the unit would be well placed to be aware of developments in different policy areas (e.g. education, health, the economy) in order to scan the horizon for terminology needs.

On a practical level, locating it in the Welsh Government would also keep the costs and overheads associated with setting up such a unit lower than establishing it outside the Government. At a time of significant financial challenges in all aspects of public life, ensuring that any investment is focused on the end-product and specialist staff, rather than overheads, Human Resource services and offices, makes a lot of sense.

The proposed unit's structure

In order to ensure that the unit is well-located to respond to policy developments, to scan the horizon for terminology needs, and respond to enquiries of a linguistic nature, it will be essential that its staff are professional workers, some with a background in linguistic infrastructure. This would enable them, for instance, to:

- Coordinate a strategic work programme for dictionaries, terminology standardisation, corpus work, and standardising orthography.
- Recognise terms which need to be standardised quickly, in response to national and international developments, and foster a relationship with relevant experts and the media to ensure that these new terms are standardised quickly and used widely and consistently from the outset.
- Scan the horizon to ascertain which terms and packages of terms will be needed in the near future, and plan the unit's work programme and coordinate relevant infrastructure projects accordingly (the unit would not be responsible for the work itself, only for coordinating it – the standardisation of terms, for instance, would continue to be led by experts and take place in the relevant centres).
- Deal with enquiries from members of the public and organisations.
- Convene and provide secretariat for the work of the Welsh Language Standardisation Panel, and ensure the Panel's decisions are mainstreamed into relevant dictionaries and terminology databases.
- Implement the linguistic infrastructure policy and monitor its progress.
- Maintain a national network of dictionary and terminology experts, as well as other linguists.
- Develop an educational and outreach element, and ensure that it's able to discuss and identify needs, find solutions, and communicate them widely.

It would also be important for the unit to have a strong marketing role, to ensure that target audiences are aware of resources, understand what they're for, and use them. That being so, ensuring that the unit has status, and fostering awareness of what it does, would help lead to more consistency in the use of words and terms, and therefore encourage more people to use them.

As one of its objectives would be to develop the long term robustness of Welsh linguistic infrastructure, we would also aim to take advantage of apprenticeship schemes to ensure that trainees are given opportunities to learn about this area of work.

Although the unit would be based in the Welsh Government, we propose that it should have its own identity and brand, in order to be visible and accessible to the public.

Funding

It isn't the Welsh Government's intention to reduce the funding which at present is distributed to various linguistic infrastructure projects. Any funding provided to create the proposed unit, and to develop the field more generally (e.g. developing the offer in relation to dictionaries, creating a website) would be new funding for this area.

What resources are available now?

At present, there is no one body or team responsible for coordinating Welsh linguistic provision in its entirety. The Welsh Government provides funding for a number of separate projects, but there is currently no formal coordination, and individual projects operate independently of each other.

Objectives:

To set up a unit with specialised staff to coordinate Welsh linguistic infrastructure, and provide strategic leadership so that the different components work together better.

The unit would be responsible for the activities and objectives outlined in areas 1-5 in this document.

How:

- Set up a unit in the Welsh Government, beginning by appointing a Head and senior-official to begin the work.
- Ensure that the unit includes staff with expertise in an aspect of linguistic infrastructure e.g. dictionaries, terminology, standardisation, corpora.
- Create a strong identity and ensure that the unit and resources connected with it are marketed effectively.
- Maintain strong relationships with the main stakeholders.
- Increase skills and expertise through apprenticeships and raising awareness of the field as a career.
- Begin to develop packages of work under the proposals described in areas 1-5 above.

Section 4 − Epilogue

There's no doubt that the way Welsh linguistic infrastructure currently works needs to be improved. There has been talk for decades of the need to better organise the field. The aim of our proposals in this paper would be to coordinate the different elements of linguistic infrastructure, commission resources such as terminology as needed, make sure that the fruits of this labour are easily available to all, and market the provision effectively. We believe this would make it easier for everyone to use the language confidently, be they new speakers, the parents of school children, or people using Welsh at work in a professional capacity.

Needless to say, additional funding will be needed to recruit staff and undertake new activities. The purpose of this consultation is to ask for your opinion, in order to help us weigh up what exactly the new way of doing things will entail, and what responsibilities the new unit should have.

It must be emphasised that we're not at the beginning of the journey. For decades, we've recognised that different elements of Welsh language infrastructure are inter-connected. Most of the components already exist – in this document we propose how to coordinate and give access to them in a more coherent and effective way, whilst adding the proactive element of scanning the horizon for developments and planning strategically.

People who use these services need a simpler way of accessing them, so that everyone is able to find simple, clear answers to their questions. As things stand, using dictionaries and terminologies can be quite confusing, and entail having a number of different resources open online at the same time in order to find an answer. We need to make sure that the resources are easy to use for everyone, not only professional translators. Getting to grips with this would ensure that everyone knows where to turn for support, so that we can all use Welsh effortlessly.