

# Multilingualism: controlled vocabulary and survey

[Education](#), [Learning](#)



## **Acknowledgements**

From February 2004 10 new member states (plus Russia and Israel) have been participating in the joint European initiative of MINERVA Plus working with MINERVA to coordinate digitization efforts and activities.

Since then Minerva Plus supplementary working groups (SWG) started operation and Hungary became the coordinator of SWG Multilingual thesauri. The issue of multilingualism is becoming more and more important in making the digital cultural heritage of Europe available. Language is one of the most significant barriers to access of websites and, because of this barrier, great parts of the European digital cultural heritage cannot be found on the Internet. MINERVA Plus conducted a major survey to get an overview of the situation concerning language usage in cultural websites.

The aim of the survey was to see to what extent cultural websites and portals are available for users of different language communities and also whether websites use more languages than the language they were originally created in. Furthermore the survey intended to find out if cultural websites are using retrieval tools such as controlled vocabularies or thesauri and whether multilingual tools are available for use. The methodology used for our survey included a questionnaire completed on a voluntary basis by our target group: libraries, museums, archives and other cultural institutions operating websites.

The selection of the websites was not scientifically founded and so the sampling is not statistically representative. Nevertheless, the survey yielded a general picture of multilingualism of cultural websites and the findings will

be a good starting point for more systematic and statistically valid research in the future. I would like to thank our Israeli colleagues for letting us use their questionnaire (Registry of Controlled Vocabularies related to Jewish Cultural Heritage and Israel) as basis for our survey. I am also very grateful to our respondents for collecting and mailing the requested information.

Last but not least I would like to express my gratitude to the editorial board of this document. Ivan Ronai NRG member for Hungary " We dedicate this report to the memory of the late Stephen Conrad. " Editorial Committee Stephan Conrad (Germany) David Dawson (The United Kingdom), Christophe Dessaux (France), Kate Fernie (The United Kingdom), Antonella Fresa (Italy), Dr. Allison Kupietzky (Israel), Ivan Ronai (Hungary), Martina Rozman Salobir (Slovenia), Gabriella Szaloki (Hungary) 4 Contributors Jitka Zamrzlova (Czech Republik) Marju Reismaa (Estonia) Veronique Prouvost (France) Dimitrios A. Koutsomitropoulos (Greece) Stephan Conrad (Germany) Szaloki Gabriella (Hungary) Marzia Piccininno (Italy) Guiliana di Francesco (Italy) Dr. Allison Kupietzky (Israel) Domitilla Fagan (Ireland) Laila Valdovska (Latvia) Pierre Sammut (Malta) Jos Taekema (The Netherlands) Lars Egeland (Norway) Maria Sliwinska (Poland) Piotr Ryszewski (Poland) Ana Alvarez Lacambra (Spain) Martina Roznan Salobir (Slovenia) Elena Kuzmina (The Russian Federation) Martin Katuscak (Slovak Republik) Kate Fernie (The United Kingdom) Guy Frank (Luxembourg) Minna Kaukonen (Finland)

## **Introduction**

FOREWORD " Immer werden jene vonnoten sein, die auf das Bindende zwischen den Volkern jenseits des Trennenden hindeuten und im Herzen der Menschheit den Gedanken eines kommenden Zeitalters hoherer Humanitat

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

glaubig erneuern“ Stefan Zweig: Triumph und Tragik des Erasmus von Rotterdam There will always be necessary those who look on the binding parts between peoples beyond the separating ones, reinvigorating, in the heart of mankind, the thought of a forthcoming century of superior humanity. What is multilingualism? The European context " Multilingualism refers to both a person's ability to use several languages and the co-existence of different language communities in one geographical area. " 1 In fact, the more languages you know, the more of a person you are (Kolko jazykov vies, tolkokrat si clovekom), says the Slovak proverb that opens the Commission's communication on multilingualism. The European Commission adopted in November 2005 the communication to the Council " New Framework Strategy for Multilingualism" document<sup>2</sup>, which underlines the importance of multilingualism and introduces the European Commission's multilingualism policy. The Commission's multilingualism policy has three aims:

- to encourage language learning and promoting linguistic diversity in society;
- to promote a healthy multilingual economy, and
- •to give citizens access to European Union legislation, procedures and information in their own languages.

Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions - A New Framework Strategy for Multilingualism COM(2005) 596 final Brussels, 22. 11. 2005 <http://europa.eu.int/languages/servlets/Doc?d=913> European Commission press release

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

<http://europa.eu.int/rapid/pressReleasesAction.do?reference=IP/05/1451&format=HTML&aged=0&language=EN&guiLanguage=en#fn1>  
<http://europa.eu.int/languages/servlets/Doc?id=9136> The Tower of Babel is an ancient symbol of the multilingualism in the Bible<sup>4</sup> Ever since the European Year of Languages in 2001<sup>5</sup> was organised by the European Council, the European Day of Languages has been held every September 26 to help the public appreciate the importance of language learning, to raise awareness of all the languages spoken in Europe and to encourage lifelong language learning.

It is a celebration of Europe's linguistic diversity. The European Commission has also launched recently a new portal for European languages<sup>6</sup>, which is available in all the 20 official languages of the European Union. It is a useful information source of multilingualism and can be a starting point for every project. The resource given has been prepared for the general public and covers a range from the Union's policies to encourage language learning and linguistic diversity.

The main areas covered are:

- linguistic diversity
- language learning
- language teaching
- translation
- interpretation
- languagetechnology

A wide range of information is given for each of them from EU and national rules to a round up of employment opportunities for professional linguists with the Union's institutions. In fact, the Communication also stresses the importance of language skills to worker mobility and the

4 5 6 Pieter Breugel: Tower of Babel [http://europa.eu.int/comm/education/policies/lang/awareness/year2001\\_en.html](http://europa.eu.int/comm/education/policies/lang/awareness/year2001_en.html)

7 competitiveness of the EU economy. The Commission will publish a study next year on the impact on the European economy of shortages of languages skills. It is worth mentioning the Eurobarometer<sup>7</sup> survey published on the web site that was carried out between May and June 2005 among European citizens including those of the accession countries (Bulgaria and Romania), of candidate countries (Croatia and Turkey) and the Turkish Cypriot Community. One of the most interesting results is that half of the people interviewed say that they can hold a conversation in a second language apart from their mother tongue.

Tower of Babel in the Maciejowski Bible<sup>8</sup> Why Multilingualism is important? In Europe we want to live in a socially inclusive society in which diverse cultures live in mutual understanding, building at the same time a common European identity. Language, together with the shared knowledge and traditions, which passed from one generation to another, is an important part of an individual's cultural identity. We strongly believe that the diversity of languages, traditions and historical experiences enriches us all and fosters our common potential for creativity.

Let us make languages connect people and cultures not divide them. This is an important role for cultural institutions. Take the case of museums;

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

multilingualism is of significant importance. Museums define their sphere of tasks as collecting, making available, preserving, researching and exhibiting objects. A multilingual exchange of information on objects supports museums in their tasks on the one hand and on the other hand the users of the products of museum work (visitors). Museums collect objects whose meaning renders them unique and one-of-a-kind.

However, the physical objects can only be available in one place at one particular time, making them accessible only to a few people. In order to make information about museum pieces available to as wide a target group as possible, a special importance lies in the accessibility of the relevant information on the Internet and in overcoming language barriers. Web sites are extremely powerful mean to do that. Nonetheless, multilingual exchange of information about museum pieces is also of interest for cultural tourism and therefore for economic reasons.

A museum visitor wants to know how to access such objects, in other words, which museum is displaying the objects at what point in time. Museums need to be able to make this information available in different languages in order to reach visitors from neighbouring countries. 7 8 Europeans and languages. A survey in 25 EU Member States, in the accession countries (Bulgaria and Romania), the candidate countries (Croatia and Turkey) and among the Turkish Cypriot Community <http://europa.eu.int/languages/en/document/80/20> [http://en.wikipedia.org/wiki/Image:Maciejowski\\_Tower\\_of\\_Babel.jpg](http://en.wikipedia.org/wiki/Image:Maciejowski_Tower_of_Babel.jpg) 8

Multilingualism is of special interest to smaller and local museums in Europe, to preserve local and national differences and to make available their <https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

peculiarities and unique characteristics to others. Objects that originally belonged together have been spread around the world by means of exchange, purchase, division of goods and also by theft or violent conflict. To recreate relationships between the parts of collections that have been dispersed to multiple institutions and countries, it is essential to exchange relevant information and for this to happen multilingual accessibility is a prerequisite.

Further, it can be assumed that many objects can be qualified through a provenancereconstructionthat crosses borders. The single objects mutually contextualise one another. And cross-border communication implies the use of multiple languages. Another point of view is the quality and effectiveness of communication on the Internet. Information technologies dramatically changed users' behaviour at the end of twentieth century and a constant increase in demands and expectations from new services can be observed. Some countries report that the number of virtual visits to cultural institutions is becoming higher than real visits.

Therefore each institution should take care about communication on the Internet and the best medium for this is an institutional website. Cultural institutions have become aware of the power of websites and have been creating their own websites since the 1990s. Beyond the problem of guaranteeing a regular maintenance of the information provided, multilingualism plays again a strategic role, The majority of websites are addressed to their own small communities, such as university members, public library readers or the citizens of a town in which a museum is located.



However, the more useful information that can be found on a website, the more Internet users visit them regardless of borders. Language is the major barrier to foreigners in making use of these websites. Whilst policies and initiatives aimed at preserving languages are the primary responsibility of Member States, community action can play a catalytic role at European level adding value to the Member States' efforts. The development of multilingualism on the Internet has been stimulated in recent years by the European Commission by supporting trans-national projects, fostering partnership between digital content owners and language industries.

However, support for high quality multilingual resources still needs to be enhanced. A pan-European inventory and library of mature linguistic tools, resources and applications as well as qualified centres of competence and excellence would provide helpful support. Online access to this inventory, oriented towards problem-solving, providing cultural institutions with appropriate solutions for specific problems related to linguistic and cultural customization would be beneficial for the improvement of multilingualism in the web cultural applications.

This Handbook is intended as a contribution to this pan-European inventory. Europe's experiences in multiculturalism and multilingualism represent an enormous strength that European cultural institutions should be able to exploit by positioning themselves in the new digital sphere of information and knowledge society.

## **EXECUTIVE SUMMARY**

This document was created for cultural institutions to emphasize the importance of multilingualism, and to provide them information and tools for establishing multilingual access to their collections.

In the Introduction we summarize the whole survey process carried out by the WP3 working group in the scope of the MINERVA Plus Project. The aim of the survey was mapping the multilingualism of the cultural sites and collecting information on multilingual thesauri in use. The survey lasted for a year from June 2004 to June 2005 in two runs, the results are presented in the following chapters. During the survey process we realized that we need to learn about official and minority languages and legislation within different countries and so we started to collect Country reports.

This information should be the starting point in each European Union project because it helps to understand the differences between countries. Each report has the same structure: multilingual diversity of the country, an evaluation of the participation in the survey and use of multilingual thesauri or controlled vocabularies. One of the practical aims of the MINERVA Project is to share the Best practice examples. Country representatives were asked to nominate the best practice examples for multilingual websites and thesauri.

We have summarized the results of the nominations for Best practice examples for multilingual thesauri and introduced some of them in detail, which are already in use in many different countries. In the survey we collected 657 multilingual websites<sup>9</sup> from all over Europe. We present the Best practice examples of multilingual cultural websites, which are available <https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

in two or more languages, and meet the requirements of the 7th chapter of the Quality Principles for cultural Web Sites: a handbook<sup>10</sup> published by the MINERVA Plus WP5 working group.

Some of them implements thesaurus for information retrieval. From the results, and findings we set up the Conclusions about the importance of multilingualism, and the use of multilingual thesauri. We also made same proposal for the future in the Future perspectives about supporting the translation of the well-tested thesauri, the quality test beds for thesauri, and the further collection of multilingual thesauri.

## **INTRODUCTION – ABOUT THE SURVEY**

After accession to the European Union the new member states became a part of a multicultural and multilingual community.

At present there are 20 official, and about 150 estimated minority and immigrant languages are spoken in the enlarged European Union<sup>11</sup>. Thus information retrieval whether on the Web or in a common database can be a serious problem. That is why, at the kick off meeting of the MINERVA Plus Project in Budapest February 2004, it was decided to establish a working group specializing in multilingual issues, especially on multilingual thesauri. The working group was a follow up of the

9 10 11 MINERVA Institutions <http://www.minervaeurope.org/institutions.htm>

Calimera Guidelines: Cultural Applications: Local Institutions Mediating Electronic Resources, Multilingualism, 2004. <http://www.calimera.org/Lists/Guidelines/Multilingualism.htm>

work carried out by the working group by the MINERVA Project Work Package 3 (WP3), led by France. Instead of creating a brand new multilingual thesaurus for the project's purposes, we decided to make a survey of multilingual websites and thesauri. This also gave us a good opportunity to discover the usage of multilingual thesauri all over Europe.

The survey was completely voluntary, and we declare that our results cannot be considered to be statistically relevant. They can be best referred to as a random sampling. The reason for this is explained by the different customs of the member states, different methods of circulating and gathering information implemented by the national representatives and the different social attitudes of each country towards the issue of multilingualism and consequently the different levels of maturity of the digital products in terms of multilingual features.

The coordinators' attitudes, working fields and positions made a major impact on their countries' results. Some countries, including Israel, The Netherlands and Slovakia, had just finished a survey and were able to contribute these results offline. Other countries, including Poland, Greece and Russia, decided to send offline results because of a shortage of time or resources; these were added to the online results in the same format. The aim of the survey was mapping multilingual access to the European digital cultural content. To implement the survey we compiled a website <http://www.ek.oszk.hu/minerva/survey>, which was used for data collection and displaying the actual results. The online questionnaire could be reached from the front page. The questionnaire had two major parts. The first section was for auditing the multilingualism of the cultural websites. The second part <https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

could be filled out only by institutions that declared the use of controlled vocabularies for information retrieval in their database. This part was based on an Israeli questionnaire that was developed for a different survey. The results could be continuously followed online.

There were separate links from the front page to the " Statistics", to the registered " Institutions", and to the " Controlled vocabularies" grouped by the countries. The statistics were calculated by individual countries and also for the results of the whole survey. The institution's types, the number of the languages available on the site, the site availability in English and the type of searching tools were analyzed. " Institutions" showed the names of the registered institutions linked to their websites, so that the site could be easily reviewed. Controlled vocabularies" showed the names of the registered thesauri and their registration form. The first run of the data collection started in June 2004 and ended in August. In the first analysis there were 236 answers from 21 states. This high score indicated also the diversity of participation. From 1 to 40 institutions answered per state and registered their websites in our database. There were 67 libraries, 63 museums, 35 archives, 21 cultural sites, and 45 other institutions. The results of the first run demonstrated that the 30% of the websites were still monolingual, 3% were bilingual, and about 26% were multilingual. There were 31 thesauri registered: 13 from Italy, 10 from the United Kingdom, 6 from Hungary, 1 from the Netherlands, and 1 from Austria. The working group had its first meeting on 12th of November 2004 in Budapest. The members of the working group presented a short country report. The slides are available on the official website of the survey by clicking on the "

Download the slide shows". It was clear, that there are different legislation and customs in each member states and so we planned to collect country reports of multilingual aspects.

The group agreed on new rules for the survey and restrictions for the results. We started a second run of the survey for those countries that were underrepresented in the first run. We also decided to create a mailing list (WP3 list) for circulating general information and discussion. We set up the criteria for the best practices examples and agreed on definitions. The second meeting took place in Berlin on the 8th of April 2005, during the two day WP5 meeting on quality of the websites. We gained useful experiences. We realised that it would be useful to get to know about the multilingual issues from each country in a sophisticated way and so we decided to collect country reports. This will also help us to find the best practices examples to share. We agreed on the form of the country reports and the deadline for preparing them. The second run of the survey closed at the end of May 2005. The combined results of the two runs of the survey doubled those of the first. There were 676 websites registered from 24 countries. Some countries, like Germany, Italy, Greece, Israel and Malta sent additional information, but no information came from Cyprus, Latvia, Lithuania or Luxembourg.

There were 265 museums, 138 libraries, 98 archives, 65 cultural sites, and 129 other websites registered. 179 of them were monolingual, the majority (310) were bilingual, 123 were available in 4 languages, 14 in 5 languages, 10 in 6 languages, 4 in 7 languages, 3 in 9 languages, and 1 in 34 languages. 491 out of the 676 websites were available in English. There were 106 registered controlled vocabularies in our database: 1 from Austria, 3 from

France, 22 from Germany, 6 from Hungary, 30 from Israel, 13 from Italy, 19 from Russia, 1 from Sweden, 1 from The Netherlands and 10 from the United Kingdom.

The third meeting took place in Budapest on the 8th of September 2005. The participants of the meeting established an editorial board of this document. We agreed on the timeline, set up the structure of the deliverable and shared the tasks among the group. 1. 4 DEFINITIONS Definition of terms used in the survey: Cultural Site: is a website of a cultural institution (libraries, museums, archives) or a website providing cultural information having a digital collection (virtual galleries, cultural databases, historical sites).

Multilingual website: is a website providing information in two or more languages We understand that thesaurus is a special type of controlled vocabulary, in which the relations between the terms are specified. We are looking for multilingual thesauri focusing on cultural coverage, which can be used for online information retrieval on a cultural website. A controlled vocabulary<sup>12</sup> is a list of terms that have been explicitly enumerated. This list is controlled by and is available from a controlled vocabulary registration authority.

All terms in a controlled vocabulary should have an unambiguous, non-redundant definition. This is a design goal that may not be true in practice. It depends on how strict the controlled vocabulary registration authority is regarding registration of terms into a controlled vocabulary. As a minimum the following two rules should be enforced: If the same term is commonly used to mean different concepts in different contexts, then its name is explicitly qualified to resolve this ambiguity. If multiple terms are used to

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

mean the same thing, one of the terms is identified as the preferred term in the controlled vocabulary and the other terms are listed as synonyms, aliases or non-preferred. A thesaurus is a networked collection of controlled vocabulary terms. This means that a thesaurus uses associative relationships in addition to parent-child relationships. The expressiveness of the associative relationships in a thesaurus vary and can be as simple as “related to term” as in term A is related to term B. 12 What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model? <http://www.metamodel.com/article.php?story=20030115211223271> 12 A thesaurus has two kinds of links: broader/narrower term, which is much like the generalization/specialization link, but may include a variety of others (just like a taxonomy). In fact, the broader/narrower links of a thesaurus is not really different from a taxonomy, as described above. A thesaurus has another kind of link, which typically will not be a hierarchical relation, although it could be. This link may not have any explicit meaning at all, other than that there is some relationship between the two terms.

Additional information about thesauri: What controlled vocabularies, taxonomies, thesauri, ontologies, and meta-models all have in common are:

- They are approaches to help structure, classify, model, and or represent the concepts and relationships pertaining to some subject matter of interest to some community.
- They are intended to enable a community to come to agreement and to commit to use the same terms in the same way.



- There is a set of terms that some community agrees to use to refer to these concepts and relationships. The meaning of the terms is specified in some way and to some degree.
- They are fuzzy, ill-defined notions used in many different ways by different individuals and communities.

Controlled Vocabulary vs Free Text<sup>13</sup> When you search an electronic database for information on a specific topic, you must find a balance between achieving high precision and achieving high recall. A search which results in high precision will be narrow, including only records which are very focused on your topic. However, this type of search may be so focused that you miss out on some information which may be relevant.

A search which results in high recall will be broader and more inclusive, but may retrieve irrelevant information which you then have to sort through.

Controlled Vocabulary Most electronic databases allow you to search a subject by controlled vocabulary. This is often the best way to strike that balance between precision and recall. Controlled vocabulary is a set of predetermined terms which are used consistently to describe certain concepts. Experts in a discipline analyze an article and choose the appropriate terms from the controlled vocabulary which best characterize what the article is about.

All articles which address the same concept will be indexed using the same term or combination of terms. . Thesaurus Of course, to use controlled vocabulary, you must know what the terms are. The list of these terms is called a thesaurus. Many electronic databases allow you to search the thesaurus online to find the appropriate term for your search. Some

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

databases, including OVID databases, will automatically map, or translate the term you type to the closest matching controlled vocabulary term and perform the search on that controlled vocabulary term. 3 Information adapted by Shauna Rutherford, University of Calgary Library, from: Barclay, Donald (ed). 1995. Teaching Electronic Information Literacy: A How-To-Do-It Manual. New York: Neil Schuman. (p. 63-64). 13 Controlled vocabulary terms can usually be found in the subject headings or descriptor fields of a database record. When you search by controlled vocabulary, the system is looking for those terms only in the subject heading or descriptor fields, not in the other fields of the database.

Advantages: Controlled vocabulary ensures that you retrieve all records which address the same topic, regardless of which words the authors use to describe that topic. Synonyms are all indexed under the same controlled vocabulary term, so the searcher is spared the job of thinking of and searching every term that describes a certain topic. Controlled vocabulary also avoids problems with spelling variations. Disadvantages: There will be times when using controlled vocabulary does not result in the exact search that you need.

New topics are not well represented by controlled vocabulary. As well a very specific and defined topic may not be represented in the controlled vocabulary which provides a subject heading which is much too broad. Free Text Almost all electronic databases allow free-text or keyword searching. In this type of search, the system usually looks for your search terms in every field of the record (not just in the subject heading or descriptor fields) and it

looks for those terms to occur exactly as you type them, without mapping or translating them to controlled vocabulary terms.

**Advantages** Free-text searching can often provide more results in a shorter time because you are not reviewing the thesaurus for the controlled subject heading. It is appropriate for very specific searches or when the topic you are looking for is relatively new. **Disadvantage** Free-text searching often results in missed records that are very relevant to your search topic. You must spend more time planning your search strategy to ensure that you are searching all appropriate synonyms of your search term.

Success, therefore, often depends on your familiarity with the search topic and your ability to identify appropriate keywords and their synonyms.

2. 1 Country reports CZECH REPUBLIC 2. 1. 1 Population and Languages spoken The number of inhabitants in the Czech Republic is about 10 million. 90. 4% of the population is Czech by nationality although many other nationalities are represented (see the table below). 94, 1% citizens speak Czech, which is the official language of the Czech Republic.

Nationalities in the Czech Republic in 2001 Population in total 10 230 060  
 Czech Moravian Slovakian Polish German Ukrainian Vietnamese Hungarian  
 Russian Romany/gypsy Silesian Bulgarian Grecian Serbian Croatian  
 Romanian Albanian Others U/I 9 249 777 380 474 193 190 51 968 39 106 22  
 112 17 462 14 672 12 369 11 746 10 878 4 363 3 219 1 801 1 585 1 238  
 690 39 477 172 827 100% 90, 4 3, 7 1, 9 0, 5 0, 4 0, 2 0, 2 0, 1 0, 1 0, 1 0, 1  
 0 0 0 0 0 0 0, 4 1, 7 2. 1. 2 The survey in the Czech Republic In the first  
 round of the survey, 15 cultural institutions were chosen; the survey was  
 completed by studying their web sites via the Internet.

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

This seemed to be the most suitable method of the obtaining valid results. The cultural institutions were grouped into 4 categories: museums, memorials, galleries and libraries. Museums Museum of Decorative Arts in Prague National Technical Museum in Prague Technical Museum in Brno National Museum of Agriculture in Prague The Moravian Museum Comenius Museum in Uhersky Brod Museum of Puppets in Chrudim 15 Languages used CZ, EN CZ, EN CZ CZ, EN, DE CZ, EN CZ CZ, EN, DE, FR, NL, IT

Hussite Museum in Tabor Museum of Glass and Jewellery in Jablonec nad Nisou The Wallachian Open Air Museum in Roznov Memorials Memorial Lidice Memorial Terezin Galleries National Gallery in Prague Moravian Gallery in Brno Libraries National Library of the Czech Republic CZ, EN, DE CZ CZ, DE, EN CZ, EN, DE CZ, EN, DE CZ, EN CZ, EN CZ, EN In the second round of the survey, a random sample of the websites of members of the Association of the Museums and Galleries of the Czech Republic (AMG) were checked.

The AMG has 856 official members. In Prague there are 51 institutions; 26 museums and 25 other cultural institutions (galleries, memorials etc). The survey found that, among the Prague museums websites 19. 2% were monolingual, 69. 2 % were bilingual and 11. 6% were multilingual; 80. 8 % were available in English. The survey found that, among the non-Prague museums websites: 33 % were monolingual, 40% were bilingual websites and 27% were multilingual websites; 67% were available in English.

The results from non-Prague museums were as follows: Museums The City of Tesin Museum The Pharmaceutical Museum Kuks Regional Museum Ceska Lipa City Museum Usti nad Labem Regional Museum of Kromeriz South  
<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

Moravian Museum in Znojmo The Museum of Moravian Slovakia East Bohemian Museum in Pardubice The Museum of Mlada Boleslav Region The Museum of Romani (Gypsy)CultureSports Cars Museum Lany Museum Podkrkonosi in Trutnov Museum of Historical Motorcycles and Bohemian Toys Museum Kasperske Hory Regional Historical and Geographical Museum in Sumperk The Town Museum Nova Paka Technical Museum in Koprivnice

Languages used CZ CZ CZ CZ CZ CZ, EN CZ, EN CZ, EN CZ, EN CZ, EN CZ, EN CZ, EN, DE CZ, EN, DE CZ, EN, DE, PL CZ, EN, DE CZ, EN, DE, FR, PL, RU

16 Portals The Server muzei a galerii CR (Association of the Museums and Galleries of the Czech Republic: <http://www.cz-museums.cz/>) is the most comprehensive portal of the Czech museums. An English version is currently under construction. Information on museums and cultural heritage can be found on the portal Startpage (<http://muzea.startpage.cz/>) as well, but only in the Czech language, the same situation is on the Seznam web catalogue (Muzea).

Prague museums are described and listed on the portal , Prague - Heart of Europe' ([http://www.heartofeurope.cz/museum\\_national1.html](http://www.heartofeurope.cz/museum_national1.html)), which is used by foreign visitors to Prague and has an English version. There are several library portals. A list of the most comprehensive can be found on the home page of the National Library of the Czech Republic (<http://www.nkp.cz/en/>). Uniform Information Gateway (<http://www.jib.cz/>), Conspectus (<http://conspectus.nkp.cz/>), Memoria Project – Manuscriptorium (<http://www.memoria.cz/>) and Kramerius (<http://kramerius.nkp.cz/>) all have an English version. Czech libraries can be found also on Kknihovny.cz [http://www.knihovny.cz/en\\_index.hp3](http://www.knihovny.cz/en_index.hp3)), which provides information on Czech libraries,

their collections, information resources, services and how to access and use them. Comparison of findings Websites included in the MINERVA Survey: 86.7% available in English. Websites included in the survey of Prague cultural institutions: 80.8% available in English. Websites of other Czech museums and institutions: 67% available in English. 2. 1. 3 Thesauri and controlled vocabularies used No multilingual thesauri with cultural coverage were found to be available online among the institutions included in the survey. Relations between terms were mostly done using links or some other hypertext methods. Some of the institutions used free text indexing, but most did not use any sophisticated retrieval tools.

The same situation is true of online controlled vocabularies or eglossaries. Library of Congress Subject Headings (LCSH) Library of Congress Subject Headings (LCSH) are currently used in the Czech Republic as a source of English equivalents of subject terms, a Czech translation does not exist. UNESCO Thesaurus A Czech translation of the UNESCO thesaurus does not exist. 17 2. 2 ESTONIA 2. 2. 1 Population and Languages spoken Estonia has about 1.351 million inhabitants (as of January 2005). The largest ethnic groups are Estonians (68%), Russians (26%), Ukrainians (2%), Belarussians (1%) and Finns (1%). Estonian is the only official language in Estonia in local government and state institutions.

The Estonian language belongs to the Finno-Ugric language family and is closely related to Finnish. Finnish, English, Russian and German are also widely spoken and understood in Estonia. 2. 2. 2 The survey in Estonia In 2004, 8 Estonian institutions took part in the MINERVA survey of multilingualism in cultural websites. These included 2 archives, 1 library, 4

<https://assignbuster.com/multilingualism-controlled-vocabulary-and-survey/>

museums and 1 other cultural organisation: Estonian State Archives (<http://riigi.arhiiv.ee/>) Estonian Historical Archives (<http://www.eha.ee/>) National Library of Estonia (<http://www.nlib.ee/>) Estonian Theatre and Music Museum (<http://www.tmm.ee/>) Estonian National Museum (<http://www.rm.ee/>) The Art Museum of Tartu (<http://www.tartmus.ee/>) Estonian Museum of Applied Art and Design (<http://www.etdm.ee/>) Conservation Centre Kanut (<http://www.kanut.ee/>) As this was not a representative sample, 34 additional websites were surveyed via the Internet. These included 30 museums (museums within the government of the Ministry of Culture, county museums and municipal museums financed by the Ministry of Culture), 20 libraries (research and special libraries and central libraries) and 4 archives (governmental and national archival institutions). Est Museums within the government of the Ministry of Culture:

Central museums: Art

Museum of Estonia Estonian Health Care Museum Estonian History Museum Estonian Maritime Museum Estonian National Museum Estonian Open Air Museum Estonian Sports Museum Estonian Theatre and Music Museum <http://www.ekm.ee/> <http://www.tervishoiumuuseum.ee/> <http://www.eam.ee/> <http://www.meremuuseum.ee/> <http://www.erm.ee/> <http://www.evm.ee/> <http://www.spordimuuseum.ee/> <http://www.tmm.ee/> + + + + + + + + + + + + + + free text free text Eng Rus Ger Fin Search tools Museum of Estonian Architecture <http://www.arhitektuurimuuseum.ee/> State museums: Estonian Museum of Applied Art and Design Tartu Art Museum County museums: <http://www.etdm.ee/> <http://www.tartmus.ee/>

18 Harjumaa Museum Hiiumaa Museum Iisaku Museum Jarvamaa Museum  
Laanemaa Museum Mahtra Peasantry Museum Memorial Museum of Dr. Fr.

R. Kreutzwald Parish School Museum of Oskar Luts at Palamuse Polva  
Peasantry Museum Parnu Museum Saaremaa Museum Tartumaa Museum  
Valga Museum Viljandi Museum Foundations: Anton Hansen Tammsaare  
Museum at Vargamae Museums of Virumaa Municipal museums: Juhan Liiv  
Museum Muhu Museum Setu Farm Museum Libraries: Research and special  
libraries: AcademicLibrary of Tallinn University Estonian National Library  
Estonian Repository Library Tartu University Library Central libraries: Harju  
County Library Jogeva County Central Library Johvi Central Library Jarva  
County Central Library Kohtla-Jarve Central Library <http://www.muuseum.harju.ee/> <http://www.muuseum.hiiumaa.ee/> <http://www.hot.ee/iisakumuuseum/> <http://www.jarva.ee/?CatID=201> <http://www.muuseum.haapsalu.ee/> <http://www.hot.ee/mahtram/> <http://www.hot.ee/muuseumvoru/> <http://www.palmuseum.ee/> [http://www.talurahvamuseum.polvamaa.ee/](http://talurahvamuseum.polvamaa.ee/) <http://www.pernau.ee/> <http://www.saaremaamuuseum.ee/> <http://www.tartumaamuuseum.ee/> <http://www.valgamuseum.ee/> <http://www.muuseum.viljandimaa.ee/> <http://ice.estnet.ee/albuuv/?id=257> <http://www.svm.ee/> <http://www.muusa.ee/> <http://www.muuseum.ee/> <http://www.hot.ee/setomuuseum/> + + + + + + + + + + + + + + + + + + + + free text free text - + + + + free text + + - + ree text - + - + + <http://www.tulib.ee/> <http://www.nlib.ee/> <http://www.hoiuraamatukogu.ee/> <http://www.utlib.ee/> <http://www.poltsamaa.ee/index.php?lang=est=72,661,664> <http://raamat.paide.ee/> <http://k-jarve.lib.ee/> + + + + + + + free text free text - + - + + + + free



text Korvekula Library / Tartu County <http://tmk.tartuvv.ee/tmk/> Central Library Kardla Central Library Laane County Central Library Narva Central Library Polva Central Library Parnu Central Library <http://www.lib.haapsalu.ee/> <http://www.raamat.polva.ee/> <http://www.pkr.ee/> + + + + + - Laane-Viru County Central Library <http://www.lvkrk.ee/> + + - 19

Rapla Central Library Saare County Central Library Sillamae City Central Library Tallinn Central Library Tartu City Central Library Valga Central Library Viljandi City Library Voru County Central Library Archives: <http://www.raplakrk.ee/> <http://www.kuessaare.ee/skr/> <http://www.slib.ee/> <http://www.keskraamatukogu.ee/> <http://www.luts.ee/> <http://www.valgark.ee/> <http://www.raamatukogu.viljandi.ee/> <http://lib.werro.ee/> + + + + + + +

+ + + + free text free text free text free text Governmental and national archival institutions: Estonian Filmarchive Estonian Historical Archives Estonian State Archives National Archives of Estonia <http://www.filmi.arhiiv.ee/> <http://www.eha.ee/> <http://www.riigi.arhiiv.ee/> <http://www.ra.ee/> + +

+ + + + + + + free text free text - 4 of these websites were monolingual while 30 were multilingual as follows: • 20 sites were available in 2 languages • 7 sites were available in 3 languages • 2 site was available in 4 languages • 1 site was available in 5 languages 4 foreign languages were represented including English (28), Russian (9), German (4) and Finnish (3).

The extent to which the contents are available in these languages varies. On the web pages there are many signs of work-in-progress: pages in other languages being announced or in an early stage of development. 2. 2. 3

Thesauri and controlled vocabularies used At present there are no multilingual thesauri in use on the Web by any Estonian cultural institution.

15 sites provide free text search. 20 2. 3 FRANCE 2. 3. Population and Languages spoken The linguistic situation in France The political context Since the launch in 1998 of the Government Action Plan for the Information Society (PAGSI) that made culture and language two of its priorities the French government has been actively supporting and promoting research efforts and applications in the field of language. Here are some examples of this multifaceted involvement:

- Taking part in multilingual European projects with an educational content (eg Linguanet : <http://www.linguanet-europa.org>)
- Taking part in multilingual European projects with a cultural content (eg Herein: <http://www.european-heritage.et/sdv/herein/> , Narcisse, Minerva or Michael)
- Taking part in multilingual European projects with a scientific content (eg Cismef: <http://www.chu-rouen.fr/cismef/>)
- Taking part in the equipment of language (terminology committees ) or in the industrialisation of language through supporting the creation of basic tools and linguistic resources in the oral and written areas (support of the linguistic research laboratories by the ministries in charge of industry, culture and education)
- Promoting knowledge in linguistic engineering by making it available to the communities of research workers and industrialists (eg Technolanguage project: <http://www.technolanguage.net>)

With its rich past and its large dissemination the French language can face the future with confidence.

In order to guarantee its national and international role in an ever changing world the general policy in favour of the French language has taken into account all areas: the role of the French language in the social cohesion, its teaching (in France and abroad), its enrichment (creation of new words), its

display through the new technologies and on the web, and its dissemination, but also its relationship with the other languages. General overview and actions taken France possesses a very rich linguistic heritage. The languages of France are our common good and contribute to the creativity of our country and to its cultural influence at the side of the French language. By the expression “ languages of France” we mean the regional or minority languages that are traditionally spoken by French citizens on the territory of the Republic and that are not the official language of any state. For this reason neither Portuguese nor Chinese are languages of France though spoken by many French citizens.

Apart from the fact that they are not endangered languages, they are regularly taught within the education system as foreign languages. Western Armenian is the language of the diaspora and thus a language of France whereas Eastern Armenian is the official language of the Republic of Armenia. Colloquial Arabic is the language that is actually spoken by many French people. It differs from literary or classical Arabic that is the official language of Arabic countries and is used in the media but not by the general population. These definition criteria are adapted from the European Charter for Regional or Minority Languages. France’s linguistic policy is indeed developed within the European framework. 21

Languages that transcend political frontiers, such as the Basque, Catalan, Flemish and Frankish languages, illustrate the internal plurality and the unity of our common cultural space. They open doors on neighbouring countries. From this angle the languages of France can be viewed as means of cultural invention and as the components of a polyphonic ensemble where the

imaginary, intellectual and affective worlds of the men and women of our country can express themselves freely. On the basis of these criteria more than seventy-five languages of France can be counted in Metropolitan France and overseas areas. They are characterized by a great diversity.

In Metropolitan France: Romance, Germanic, Celtic languages as well as Basque, a non-Indo-European language. Overseas: Creoles, Amerindian, Polynesian, Bantu (Mayotte) and Austronesian (New Caledonia) languages, among others. There is also a great demographic diversity between these languages. Three or four million people are speaking Arabic in France whereas Neku or Arha are spoken only by a few dozen people. In between, the various Creoles or the Berber languages are spoken by about two million people in France. The 1999 national census revealed that 26 % of adults living in France had regularly practiced in their youth a language other than French – Alsatian (660 000 speakers), Occitan (610 000), Oil languages (580 000), Breton (290 000).

For each of these languages one can add an equal – at least – number of occasional speakers. However language transmission in France is almost not effective any more in the family circle and relies today mostly on the teaching of these languages and their creativity in the artistic domain. Metropolitan France Regional languages : Alsatian, Basque, Breton, Catalan, Corsican, Western Flemish, Moselle Franconian, Franco-provencal, Oil languages (Franc-Comtois, Walloon, Champenois, Picard, Norman, Gallo, Poitevin-Saintongeais, Lorrain, Bourguignon-Morvandiau), Oc languages or Occitan (Gascon, Languedocien, Provencal, Auvergnat, Limousin, Vivarese-Alpine).

Non-territorial languages: Colloquial Arabic, Western Armenian, Berber, Judeo-Spanish, Romany, Yiddish) Overseas Caribbean area: French-based Creoles: Guadeloupean, Guyanese, Martiniquan, Reunionese; Bushinenge Creoles of Guyana (Anglo-Portuguese-based): Saramaca, Aluku, Njuka, Paramaca; Amerindian languages of Guyana: Galibi (or Kalina), Wayana, Palikur, Arawak (or Lokono), Wayampi, Emerillon; Hmong New Caledonia: twenty-eight Kanak languages. Grande terre: Nyelayu, Kumak, Caac, Yuaga, Jawe, Nemi, Fwai, Pije, Pwaamei, Pwapwa, VohKone, Cemuhi, Paici, Ajie, Arha, Arho, Orowe, Neku, Siche, Tiri, Xaracuu, Xaragure, Drubea, Nume; LoyaltyIslands: Nengone, Drehu, Iaai, Fagauvea. French Polynesia: Tahitian, Marquesan, Tuamotuan and Mangareva languages, languages spoken in the Austral Islands: Raivavae, Rapa and Rurutu languages. Wallis and Futuna Islands: Wallisian, Futunian. Mayotte : Maore, Malagasy dialect of Mayotte.

French Sign Language (Langue des signes française or LSF) It is traditionally used by French citizens and is also a language of France. Several legislative provisions and regulations define the place of the languages of France in the areas of culture, education and the media. The law of 4 August 1994 relative to the use of the 22 French language specifies: “ The provisions of the present law apply without prejudice to the legislation and regulations relative to regional languages in France and is not against their use” (article 21). The recognition by the French State of the specific position held by the languages of France in the nation’s cultural life was materialized by the creation in October 2001 of the Delegation generale a la langue française et aux langues de France.

The Ministry of Culture and Communication support and promote the languages of France through its multiple fields of intervention: music, literature, theatre, ethnological heritage, archives, museums, cinema, ... Moreover, specific credits have been allotted to the Delegation generale a la langue francaise et aux langues de France for the following priority objectives :

- Helping publications in or about the languages of France;
- Supporting the fields - such as the performing arts, singing, television and radio - where language act as a vector for creation
- Ensuring the presence of the languages of France through the new information and communication technologies;
- More generally, putting the emphasis on the interaction between language and culture and their importance in a living society.

<http://www.languagedefrance.om/> At a time when economical or cultural transactions between countries are growing rapidly and Europe<sup>14</sup> gets stronger a policy that guarantees the presence of the French language for French citizens and French national interests is asserted: ([http://www.culture.gouv.fr/culture/dglf/Actualites/communication\\_ministre\\_17\\_mars.pdf](http://www.culture.gouv.fr/culture/dglf/Actualites/communication_ministre_17_mars.pdf)). The French linguistic situation has come under close scrutiny and has been studied conjointly by the two national analysis institutes - INSEE<sup>15</sup> and INED<sup>16</sup> - on the occasion of the last national census: [http://www.ined.fr/publications/pop\\_et\\_soc/pes376.pdf](http://www.ined.fr/publications/pop_et_soc/pes376.pdf) However this national policy goes together with a decided openness towards the other languages:

- Through the promotion of minority and/or regional languages that are part of France's heritage. France's linguistic heritage consists of 75 languages.

However this list puts together idioms that have a variety of socio-linguistic status. It goes from the mostly spoken Creoles, which are the mother

tongues of more than one million speakers and probably the regional languages that are the most alive, to the mostly written Bourguignon-Morvandiau which is only spoken by a few people nowadays and with no mother-to-baby transmission. See: [http://www.culture.gouv.fr/culture/dglf/lang-reg/rapport\\_cerquiglini/langues-france.html](http://www.culture.gouv.fr/culture/dglf/lang-reg/rapport_cerquiglini/langues-france.html) [http://www.culture.gouv.fr/dglf/politique-langue/assises/actes\\_assises.pdf](http://www.culture.gouv.fr/dglf/politique-langue/assises/actes_assises.pdf) [http://www.insee.fr/fr/insee\\_regions/idf/rfc/docs/alapage226.pdf](http://www.insee.fr/fr/insee_regions/idf/rfc/docs/alapage226.pdf) • Through the creation of an Observatory of Linguistic Practice (Observatoire des pratiques linguistiques).

The observatory's task is to study current linguistic practices in France as well as the modalities and the effects of the contact between languages. The observatory was established in 1999 within the Delegation generale a la langue francaise et aux langues de France (DGLFLF), an interdepartmental service for culture and communication. It aims at inventorying, developing, making available, the knowledge pertaining to France's linguistic situation, in order to provide information useful for developing cultural, educational or social policies. One of its aims is also to make more widely known the common linguistic heritage [http://culture.gouv.fr/culture/dglf/lois/politiques\\_ling\\_eng\\_Europe.df](http://culture.gouv.fr/culture/dglf/lois/politiques_ling_eng_Europe.df) INSEE : Institut National de la Statistique et des Etudes Economiques (French National Institute of Statistics and Economic Studies) INED : Institut National d'Etudes Demographiques (French National Institute for Demographic Studies) 14 15 16 23 that consists of all the languages and linguistic varieties spoken in France and that contribute to its diversity. The activity of the observatory is organized around four axes: • Research and study work: the observatory

does not carry out research work as such but it is supporting and coordinating projects or research programmes on themes that are of interest to the French Ministry of Culture and Communication and - more generally - to the public authorities, the representatives, the decision