Grant Agreement number: **231135**

Project acronym: **DICTA-SIGN**

Project title: Sign Language Recognition, Generation and Modelling with application in Deaf Communication

Funding Scheme: Collaborative Project

Period covered: from 01.02.2009 to 31.01.2012

Name of the scientific representative of the project's co-ordinator¹, Title and Organisation: **Dr. Eleni Efthimiou, Research Director, ILSP/ATHENA R.C.**

Tel: +302106875356
Fax: +302106854270
E-mail: eleni_e@ilsp.gr

**Project website² address:** [http://www.dictasign.eu/](http://www.dictasign.eu/)

---

¹ Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

² The home page of the website should contain the generic European flag and the FP7 logo which are available in electronic format at the Europa website (logo of the European flag: [http://europa.eu/abc/symbols/emblem/index_en.htm](http://europa.eu/abc/symbols/emblem/index_en.htm) ; logo of the 7th FP: [http://ec.europa.eu/research/fp7/index_en.cfm?pg=logos](http://ec.europa.eu/research/fp7/index_en.cfm?pg=logos)). The area of activity of the project should also be mentioned.
4.1 Final publishable summary report

Executive summary

As the web has developed it has become a place where people interact. They post opinions, modify and enhance each other’s contributions and share information. All of this is currently done via the medium of text. Dicta-Sign researched ways to enable communication between Deaf individuals through the development of sign language based human-computer interfaces (HCI). It has researched and developed sign recognition and synthesis engines that have brought these emergent technologies significantly closer to authentic signing. These advances have been demonstrated via sign language-aware Web 2.0 interfaces, combining work from the fields of:

- Sign language recognition,
- Sign language animation via avatars
- Sign language resources
- Sign language model development

All of this, with the goal of allowing Deaf users to make, edit, and review avatar-based sign language contributions online, similar to the way people nowadays make text-based contributions on the Web.

Until this research, the Deaf community has found itself excluded from interacting naturally with each other via computers. Moreover, exclusion is already experienced with regard to interpersonal communication between Deaf individuals. Sign language videos are not a viable alternative to text, for two reasons: Firstly, they are not anonymous – contributing individuals can be recognised from videos and this limits those willing to participate. Secondly, people cannot easily edit and add to a video produced by someone else for a Wikipedia-like web site. In order to make the Web 2.0 fully accessible to Deaf people, sign language contributions must be displayed by an animated avatar, which addresses both anonymisation and easy editing.

The major objective of Dicta-Sign has been to develop an integrated framework that allows contributions in four different European sign languages: Greek, British, German, and French. Users make their contributions via Kinect sensors. These are recognized by the sign language recognition component and converted into a linguistically informed internal representation, which is used to animate the contribution via an avatar. The demonstrator can also provide sign translation into the other respective three sign languages.

The project’s outcomes have justified the initially set goals with a project demonstrator that is a Sign-Wiki prototype; showcasing the potential of sign language exploitation in Web 2.0.

Other project outcomes include:

- A parallel multi-lingual corpus for four national sign languages – German, British, French and Greek (DGS, BSL, LSF and GSL respectively).
- A multilingual dictionary of 1000+ signs of the four project sign languages,
- A continuous sign language recognition system that achieves significant improvement in terms of coverage and accuracy and also has researched the novel directions of multimodal sign fusion and signer adaptation,
- A language generation and synthesis component, covering in detail the role of manual, non-manual and placement within signing space,
- Annotation tools which incorporate these technologies providing access to the corpus and whose long term utility can be judged by the up-take by other sign language researchers,
- Three bidirectional integrated prototype systems which showcase the utility of the system
components beyond the annotation tools application
All project prototypes have been exhaustively evaluated by Deaf end-users.
Summary description of project context and objectives

The development of Web 2.0 technologies has made the WWW a place where people constantly interact with each other, by posting information (e.g. blogs, discussion forums), modifying and enhancing other people's contributions (e.g. Wikipedia), and sharing information (e.g., Facebook, social news sites). Unfortunately, these technologies are not easily accessible to sign language users, because they require the use of written language.

Can't sign language videos fulfil the same role as written text in these new technologies? In a word, no. Videos have two problems: Firstly, they are not anonymous – anyone making a contribution can be recognised from the video, which holds many people back who would otherwise be eager to contribute. Secondly, people cannot easily edit and add to a video that someone else has produced, so a Wikipedia-like web site in sign language is not possible based on video.

In this context, Dicta-Sign's goal has been to develop the necessary technologies that make Web 2.0 interactions in sign language possible. Hence, in a suitable scenario, users sign to their computer using a dictation style. The computer recognises the signed phrases, converts them into an internal representation of sign language, and then has an animated avatar sign them back to the users. Content on the Web is then contributed and disseminated via the signing avatars. Moreover, the internal system representation in the envisaged configuration may also allow incorporation of a simple sign-to-sign translation service, which facilitates understanding of foreign sign language contributions.

Dicta-Sign undertook fundamental research and development in a range of technologies in order to open up new potential applications for sign language users for Deaf communication and human-computer interaction.

The technologies involved are:

- Image processing
- Advanced computer vision techniques
- Statistical methods for continuous sign recognition with multimodal fusion and adaptation
- Virtual human technology
- Sign language modelling
- Grammar and lexicon design and development
- Corpus construction

Pre-lingually deaf people (in contrast to hearing people who have gone deaf later in their lives) are estimated to make up 1% of the European population. Most Deaf people use sign language as their preferred language, and their command of the language spoken by the majority is often limited, even for reading and writing.

Today’s predominant human-computer interface, window-based graphical displays with pointing interaction, is relatively manageable for most Deaf people: The use of a language foreign to them is restricted to single words (menu items) or short phrases that they can manage with limited skills in that spoken language, or learn by heart through training. The graphical user interface puts limitations on the complexity of the human-computer communication, and therefore it is expected that it will be replaced by human language interaction for many applications. This may be in the form of assistants helping the user to navigate a graphical interface, or of assistants as the sole interface to a machine which requires tasks to be described. The first form is already reality and there are already examples for the second form today. Obviously, a far better command of the interface language is required here than in graphical environments. Most Deaf people would therefore be excluded from this future form of human-computer communication unless the computer is able to communicate in sign language. Moreover, exclusion is already experienced for interpersonal communication between deaf
individuals, given the current lack of translation tools to support not only SL-to-SL but also oral-to-SL and SL-to-oral applications.

Over recent years, considerable progress has been made on how to present Sign language content on traditional websites, be it integrated with written language or purely in sign, by digital video or signing avatars. A couple of small businesses across Europe, often driven by Deaf people, specialise in innovative approaches to fulfil anti-discrimination requirements by making websites Deaf-accessible. Current Web 2.0 technologies pose a barrier for sign language users wanting to actively participate in exchange and knowledge sharing, as these technologies are tightly built upon written language: The lack of a writing system for sign languages cannot be compensated by digital video, as was the case in Web 1.0. For example in forums, the integration of digital video is mainly a question of engineering effort, but would then leave sign language users without the possibility to anonymise their contributions – a feature often considered essential for the success of these forms of group communication. In Wiki-like applications, aiming at a shared creation of discourse, digital video would simply fail to provide a solution, adding to and editing other people’s contributions is as important as original contribution.

For the first time in Europe, this project has brought together state-of-the-art research on sign language recognition, sign language production, sign language linguistics and corpus construction to develop the technologies necessary to enable the production of sign-aware Web 2.0 applications.

The key idea is already indicated in the name of the project: Dicta-Sign. Instead of having to make use of another language, or having to learn an artificial writing system for their language, sign language users can dictate their contributions by signing to their webcams. Recognised phrases will be signed back to the user by an avatar.

Given the current state-of-the-art in sign language recognition, it is obvious that continuous dictation with an error rate comparable to today’s spoken language dictation is not achievable. Instead, users employ a dictation-style signing, and select between alternatives if the system cannot fully discriminate the input. However, as dictation is not an alternative to a well-established cultural technique, but the key to new forms of participation, users are prepared to engage with the first generation of such a new technology, as user evaluation results of Dicta-Sign outcomes prove.

The idea of a cognitive vision model making use of explicit linguistic knowledge has been experimented with. Unlike the situation with speech recognition, where linguistic modelling was outperformed by statistical modelling, massive amounts of training data are simply not available for sign languages. In fact, the project needed to create annotated corpora itself in order to have comparable data for a number of sign languages available. Notably, the structure of sign language, namely the grammatical use of space and classifier constructions consisting of non-lexical elements, requires that long-distance context is available in the discrimination process.

In addition to the dictation in the Web 2.0 context, the technological progress of the project w.r.t. cognitive vision and statistical fusion applied to sign language recognition, sign language generation and sign-to-sign translation has laid the foundations for more general bidirectional language-oriented systems to include sign language. The project will thereby contribute by reducing the barriers for Deaf people that other technological advances will bring. It will also enable producers of language-driven systems to meet the requirements defined by existing and upcoming anti-discrimination acts in many European countries.

To achieve this, Dicta-Sign has investigated both input and output of sign language; using novel recognition and avatar generation technologies. Within sign language there are several complementary channels of communication which occur concurrently; manual signing, eye gaze, facial expression and body posture. Dicta-Sign has analysed and synthesised information within these channels to attempt to encompass the full range of expressiveness. Dicta-Sign has used these
investigations to develop processing techniques and produce software and corpora for multi-lingual sign language processing. This has been undertaken in a unified framework incorporating both sign recognition and generation processes.

The **Dicta-Sign** consortium was a small collection of academic institutions with internationally established research groups who undertake leading edge research in each of their specialised areas, complemented with a commercial partner specialising in SL content production and related technological applications.

**Dicta-Sign** has achieved the following **goals**:

- Establishment of the world's only extensive, parallel, multi-lingual, corpus of annotated, video, sign language data to inform further sign language research and provide a resource to inform sign language processing technology.
- Development of the most advanced sign language annotation tools to provide access to this corpus; integrating the recognition, generation and animation technologies developed within the project.
- Use of image processing and computer vision technology, alongside the development of a statistical framework, to significantly enhance and progress beyond the state-of-the-art in continuous sign language recognition. At the same time experimenting with sign language linguistic models and exploiting multimodal fusion and adaptation.
- Development and dissemination of a morpho-phonetically based orthography, based upon HamNoSys, incorporating the manual and non-manual channels of sign language communication. This is used both for annotation purposes and as an intermediate representation for both sign language generation and recognition within system software,
- Provision of the world's largest cross-lingual sign lexicons,
- Extension of sign language generation and virtual human animation technology to the most advanced rendition of realistic, automatically generated signing,
- Development of a provisional sign-to-sign translation service.

These objectives are structured through the delivery of the following **outputs**:

- A parallel multi-lingual corpus for four national sign languages – German, British, French and Greek (DGS, BSL, LSF and GSL respectively) – of a minimum of five hours signing in each language,
- A multilingual lexicon of 1000+ signs for each represented sign language,
- A continuous sign language recognition system that achieves significant improvement in terms of coverage and accuracy of sign recognition in comparison with current technology; furthermore this system has researched the novel directions of multimodal sign fusion and signer adaptation.
- A language generation and synthesis component; covering in detail the role of manual, non-manual and placement within signing space,
- Annotation tools which incorporate these technologies providing access to the corpus and whose long term utility can be judged by their up-take by other sign language researchers,
- Three bidirectional integration prototype systems which show the utility of the system components beyond the annotation tools application,
- A Sign-Wiki showcase demonstrator which exhibits how integration of the different components can support user communication needs.

Thus **Dicta-Sign** has used leading technologies in the areas of corpus development and annotation tools, image processing and computer vision, Hidden Markov Model (HMM)-based fusion & adaptation for sign language recognition, virtual human avatar technology and sign language generation.
Main S&T results/foregrounds

Main results of the project are presented next as achieved through R&D work within WP1 to WP8.

WP1: Visual Tracking and Feature Extraction

Workpackage 1 is concerned with Visual Tracking and Feature Extraction and work during the Dicta-Sign project has covered several tasks advancing the state of the art in this field. In order to recognise signs, a low level feature description of the signers actions need to be extracted. These extracted features can either be used for sign language recognition or as the basis of annotation tools.

Hand detection and tracking in 2D and 3D

A signers hand motion makes up a large part of the information contain in a sign. For this reason several approaches to hand tracking have been used. In the first instance, this was done in 2D. A single frame is used to remove the background and build the basis of an adaptive colour model for the skin. The hands and face are segmented on each frame, noting key frames where the hands and face are easily separated. Using trackers between these key frames, forward and backward prediction and smoothing are used to produce a final set of trajectories as shown in Figure 1a. While 2D information is sufficient to identify many signs, others require depth information. The Dicta-Sign corpus was captured from 3 views (see Figure 1b) and this allows us to track in 3D.

As before, a skin segmentation algorithm is applied in all 3 views. Then a novel SceneFlow algorithm is used to calculate the 3D velocity of all the skin pixels (particles) in the scene as shown in Figure 1.[Hadfield2011] These particles are then clustered into 3 objects based on their position and motion. Then, using a similar smoothing as for the 2D the trajectories can be created.[Hadfield2012] This method has been shown to work on both the Dicta-Sign corpus data (Figure 1b) and a internal stereo dataset with a cluttered background Figure 1c.)

Visual Hand Detection/Tracking and Motion Feature Extraction

Detection of face and hands can also be based on skin colour segmentation via modelling the skin colour distribution with Gaussian Mixture Models. Skin regions are segmented and processed by morphological filtering on the skin probability domain. For the tracking of the face and hands a probabilistic framework is employed which exploits the statistics of hand and face regions. To resolve cases with occlusions a forward/backward linear prediction and template matching is employed [Roussos2010a, Roussos2010b]. Figure 6(a-c) shows examples from this tracking.

Body Part Detection/Estimation from Depth

With the release of the Kinect mid-way through the project, we have been investigating methods for localising body parts from depth data. Two approaches have been developed, the first uses poselets (subsets of connected body parts in a given pose) and uses these in a scanning window over the full pose. Each poselet which is discovered gives its predictions of surrounding joint positions. These
estimations are combined and used to infer final joint positions. See Figure 2a for a graphical representation. The second approach uses direct regression to estimate the offsets from the centre of the scanning window to the joints and accumulates these predictions in likelihood maps before inferring final joint positions. Examples of poses inferred from depth are shown in Figure 2b.

Facial feature localisation and tracking

Non Manual Features make up a significant proportion of Sign Language. To this end a facial feature tracker has been developed that can accurately track points on a signer’s face [Ong2011]. Parts of the Dicta-Sign corpus have been tracked using this method as shown in Figure 3. This tracker has been further modified to include failure-self-righting and remove the need for manual intervention during tracking. It is due to these changes that it is now suitable for use as annotation tool within the A³ interface.

Face Modelling for Tracking and Feature Extraction

An alternative to tracking the facial points independently is to build a face model. In this field features have been explored based on Active Appearance Modelling (AAM) of facial images [Rodomagoulakis2011]. A Global AAM consists of characteristic points of interest around facial regions such as the mouth, the eyes and the eyebrows (see Figure 4, left). Inverse compositional methods for AAM fitting and their adaptive and constrained variants are appropriate to deal with variation in head pose and lighting conditions. Once the global face graph is fitted, we extend it with supplementary Local AAMs which offer a descriptive representation for specific regions of linguistic interest [Rodomagoulakis2011, Antonakos2012] (see Figure 4right).
**Feature Extraction**

While the hand tracking trajectories give information about the motion of the hands and face there is also more information about the sign available from the handshape. This can be done in either the appearance or the depth domain. In the appearance domain, HOGs (Histograms of Oriented Gradients) are employed on skin segmented, normalised hand patches [Cooper2011]. We have also use GIST descriptors and these have been applied to hand shape classification in both the visual and depth domains. This filter is a multi-scale approach allowing both the overall shape of the hand to be described as well as the finer detail, as shown in Figure 5[Pugeault2011].

![RGB Depth Gabor filter convolution response](image1)

**Figure 5 Examples of the GIST features extracted from a handshape**

**Handshape Modelling and Feature Extraction**

Handshapes can also be described via affine-invariant Shape-Appearance modelling (SAM) [Rousos2010a, Rousos2012]. This approach consists of: 1) Shape-Appearance representation of the hand images that models the handshapes without any landmark points. These hand images are modelled with a linear combination of affine-free eigen-images followed by an affine transformation, which effectively accounts for modest 3D hand pose variations. 2) A regularized fitting of the SAM that exploits prior information about the handshape and its dynamics in order to achieve robustness against occlusions. This process yields an accurate tracking of the hand as well as handshape features. Figure 6(d-f) shows results of the regularized fitting of the Shape-Appearance model.

![Figure 6 (a-c) Tracking sample for a portion of the DictaSign Continuous GSL Corpus. (d-e) Regularized Shape-Appearance model fitting for handshape feature extraction. In each original frame, we superimpose the fitted model-based reconstruction of the hand at the original domain. In each down-right corner, we display the reconstruction at the model domain.](image2)

**Mapping Body Motion to HamNoSys Features**

HamNoSys offers a linguistically based method for describing the component parts of a sign. By mapping the features we extract to HamNoSys we can produce robust, signer independent inputs to sign level classification. Two approaches are taken towards performing this mapping; either the features are learnt using machine learning (specifically boosting and multi-class random forests) or they are derived deterministically using heuristics.

**Location:** Since 2D and 3D tracking give the position of the hands and head we can learn classifiers to describe the relationship between co-ordinates. When the full skeleton information is available we can describe sign location in relation to the other skeletal joints. Alternatively, we can use the HamNoSys description of signing space and derive the relevant area from the given co-ordinates. For example Figure 7b shows the height information which can be derived from either the 2D, 3D or Kinect tracking. This can be used in combination with the horizontal-plane polar co-ordinate description based around the signer.
Motion: Using either 2D or 3D tracking, motions can be described as changes in co-ordinates. When learnt, these changes are mapped using binary pattern classifiers and boosting. Deterministically, the tracking can be mapped to linear motions using heuristics about direction of motion. Examples of the mapped motions are shown in Figure 7.

Handshape: Using the HOG features extracted from either appearance or Gabor features from depth, classifiers are learnt using multi-class random forests. This has proven to work well on fingerspelling handshapes [Pugeault2011] and by training for the canonical HamNoSys handshapes (Figure 8) it offers advantages to sign level classification [Cooper2011]. Alternatively, a data driven approach can be taken by clustering the feature space to reduce the dimensionality [Ong2012].

WP2: Continuous Sign Recognition

Concerning Sign Language Recognition, several advances has been made in the field of Sub-Unit (SU) based Continuous SL Modelling and Recognition. These include: 1) Dynamic-Static SUs, 2) Phonetic SUs, 3) Raw/Canonical SUs, 4) Multiple Signers and Adaptation, 5) Facial Features, 6) Fusion of multiple cues/modalities in sub-unit level and continuous data for sign recognition. The following sections present selected highlights of the afore-mentioned contributions.

Dynamic-Static Sub-Units

These sub-units provide a novel unsupervised method for data-driven sub-unit construction which accounts for sequential structure inspired by movement-hold concepts. The data-driven modelling is performed separately for dynamic (movement) and static (non-movement) parts of sign. This is automatically performed and employs for each type of segment the appropriate feature(s) with the appropriate model architecture and parameters [Pitsikalis2010, Theodorakis2011a].
Phonetic-based Sub-Units

These sub-units explore novel directions for incorporating phonetic transcriptions into sub-unit based statistical models for sign language recognition [Pitsikalis2011]. Firstly, a new symbolic processing approach for converting sign language annotations, based on HamNoSys, into sequences of labels according to the Posture-Detention-Transition-Steady Shift model [Vogler2011] is employed. Then, these labels and their correspondence with visual features is exploited to construct phonetics-based sub-unit models. By employing the statistical sub-unit construction/decoding, the sequences are aligned with the visual data, producing the time boundary information which they would otherwise lack. The resulting phonetic sub-units offer new perspectives for sign language analysis, phonetic modelling and recognition. An example result of a GSL sign is shown in Figure 9 after incorporating the proposed phonetics-based sub-units.

![Figure 9 GSL Sign for PILE: Segments after incorporation of PDTS phonetic labels into Phonetic Sub-unit Construction, Training and Alignment. Superimposed start/end frames of each segment, accompanied with an arrow for transitions/epenthesis. Each segment corresponds to a phonetic label. PDTS segments labels are of type Epenthesis (E), Posture (P), Transition (T).](image)

Multiple Signers and Adaptation

In order to deal with multiple signers and signer adaptation the following concepts are introduced: 1) Raw-canonical (SU-CanRaw) signer-independent visual phonetic models [Theodorakis et al. 2012]. These are built by uniformly sampling the feature space and constructing statistical HMM models that carry, by construction, phonetic information. Then by encapsulating data-driven phonetic information, related to the dynamic-static parts [Theodorakis et al. 2011a] to handle sequentiality of movements-postures (Figure 9). 2) Canonical (SU-Can) signer-adapted visual phonetic models. By employing the SU-CanRaw models and a small set of development data from the new signer, the SU-Can models can be trained and generated. These models are then adapted to the new signer, while still retaining their phonetic information (see Deliverable D2.3, Theodorakis et al. 2012).

Facial Features and Fusion

Classification experiments have been carried on continuous sign language corpora by employing global and local AAM facial features for sign-related facial cues. The coarse global features have proved appropriate for head pose estimation in contrast to the local ones that can measure finer facial cues including the eyebrows position and the aperture of the eyes.

Unsupervised Extreme States Classification for Facial Cues: By exploiting the global & local AAM framework, and selecting the appropriate features, a new method for the unsupervised detection and classification of facial events was introduced. These events include the head pose from the side to side turning of the head, and eye blinking; see Figure 10 [Antonakos2012].

Facial Unsupervised Detections and Linguistic Phenomena: Facial cues are studied concerning their relation to indicative linguistic phenomena such as sign and sentence boundaries or higher level phenomena such as alternative constructions and enumerations.
Fusion of Movement-Position and Handshape cue

The proposed method combines the data-driven subunits (SUs) exploiting the dynamic-static information and the affine shape-appearance handshape SUs [Pitsikalis2010, Theodorakis2011a, Roussos2010]. Specifically, we exploit the dynamic-static nature of the SU models by considering handshapes only during postures. Moreover, by employing the scheme of [Theodorakis2011b] we end-up with the fused SU models of Static Positions and Handshape SU Models. An example of these fused SU models for two signs is depicted in Figure 11.

Sequential Patterns for Sign Recognition

By approaching the sign recognition challenge using a discriminative classifier approach such as sequential pattern trees, the resultant classifiers are able to ignore noise and better generalise with fewer training examples. Specifically, sequential patterns encode what information must be present in a sign and in what order. In the example shown in Figure 12, a plausible pattern is shown for the GSL sign ‘bridge’. In the first instance, both hands move up, then they move apart, then they both move down whilst still moving apart before coming to rest. The training of these classifiers uses a tree based approach to effectively search all possible patterns finding those which best describe the given signs. It is this classification architecture which underlies the sign look up tool and the dictation style recognition of the wiki. [Elliott2011, Ong2012, Cooper2012]
WP3: Synthesis and Animation

Dicta-Sign has improved the state-of-the-art in synthesis and animation of sign language through 3D virtual characters. The most mature aspect of this work is used within the project bidirectional prototypes of WP7 and the Sign-Wiki project demonstrator of WP8. Synthesis is based on the use of SiGML, Signing Gesture Markup Language, which is an XML language based heavily on HamNoSys. HamNoSys transcriptions can be mapped directly to SiGML, although SiGML has some additional capabilities.

The SiGML animation system is primarily implemented through the JASigning software which is supported on both Windows and Macintosh platforms. This is achieved by the use of Java with OpenGL for rendering and compiled C++ native code for the Animgen component that converts SiGML to conventional low-level data for 3D character animation. JASigning includes both Java applications and web applets for enabling virtual signing on web pages. Both are deployed from an Internet server using JNLP technology that installs components automatically, but securely, on client systems.

Françoise: The Dicta-Sign Avatar

To support inter-operability the project has developed an avatar named Françoise that can be used both within the JASigning software and by the Octopus and GeneALS software of partner CNRS. This is based on a virtual signer skin (Figure 13 and Figure 14) which has been adapted to the ARP standard [Jennings 2010] for compatibility with JASigning.

The initial designs were evaluated by WebSourd and a revised design now forms the basis of the Françoise avatar which is included in the JASigning distribution and will be installed automatically for users of the project demonstrator, the Sign-Wiki.
ARP avatars can be exported to a range of industry standard tools. The main export is as a 3ds Max project containing the avatar and some animations. This can be read for use in the standard animation packages, 3ds Max and Maya. An exporter has also been written to generate BVH data files for animations, for import by CNRS.

**Enhanced Flexibility and Precision in SiGML**

During the project, enhancements have been made to SiGML that allow precise control of the timing of animations. This was exploited to show that an animation synchronised with a sign language video can be produced by annotating the signs used, along with their timings, and playing back HamNoSys transcriptions of the signs.

Fine control of the location of sign postures, and the details of transitions between postures, has been provided. For example, in addition to the discrete values inherited from HamNoSys, more continuous attribute values can now be used. Hence locations in signing space can be provided numerically. In addition, the HamNoSys concept of the attribute value “between” two other values (locations, orientations, or shapes) was extended to allow specification of any point from one extreme to the other.

When a transition in a sign involves a change in hand posture, HamNoSys allows modifications to several attributes to be described. Firstly there may be a movement involving a change in hand location. Secondly there may be a change in hand shape, and finally there may be a change in hand orientation (palm and extended finger direction). The model used by the SiGML animation system assumes that all these changes are synchronised so that the final handshape is only achieved when the movement has been completed.

However, colleagues at UHH suggested that orientation and shape were often established early in a movement. [Hanke 2011] reports enhancements made to SiGML and the Animgen software, as well as early experiments on whether the resulting animation is seen as more realistic.
Signing Avatar Components

The JASigning software is based on Java components that have been used to build a number of free-standing Java applications and Java applets for use on web pages. The key components are Animgen, which generates animation data from SiGML, an OpenGL-based renderer, and coordinating code.

It is straightforward to incorporate the Java components within new Java applications and this approach was used to provide signing client software that connects to sign recognition servers in the WP7 prototype sign search and sign look-up tools.

The components can be scripted on web pages using JavaScript interfaces and controlled through a range of user interface components. In the Sign-Wiki, an enhanced applet provides sign synthesis and animation, it also coordinates sign recognition.

The OpenGL browser interface from Java has limited support, making cross-platform implementation difficult. With the development of HTML5 and the WebGL standard that is supported on most leading browsers (Firefox, Chrome, Safari, but not Internet Explorer) it becomes attractive to move to more integrated browser support. With this in mind, a WebGL implementation of the renderer has been developed. The WebGL implementation should become more reliable as browser support for HTML5 increases. It will also make it practical to provide the software on operating systems such as Linux and potentially on mobile devices without producing bespoke implementations. There is also a compatible renderer written in C# enabling .NET applications to include SiGML animation components.

WP4: Linguistic Modelling

Workpackage 4 was concerned with the high-level linguistic framework relevant to sign languages. The aim was to develop lexicon and grammar models to assist in generation and recognition.

Lexicon models

The task consisted of representing the lexical units of sign languages. Models have been developed to describe these lexical units in a formal way, so that the descriptions may be used either as an output from a sign recognition system, or as an input to an animation system. This required including both phonetic and abstract levels in the representation.

Phonetic representation: This was based both on HamNoSys and on the Johnson & Liddell phonetic notation, which provides more information about the temporal structure of a sign. The project developed an automated method to translate HamNoSys into a segmented structure. This is a phonetic representation of a sign that is both directly applicable to recognition and which can also help with modelling linguistic processes for both recognition and animation.
Abstract representation: This was based on Zebedee, a sign description model which uses a geometric tool kit and a time structure. The representation allows signs to be specified, or any lexical unit and its associated variability (whether or not carrying meaning) in the signing space. It is abstract enough for each sign description to capture many variations in its final production and make it reusable in different signing contexts.

The project developed AZedit (Figure 16) and AZfilter, two tools dedicated to editing and manipulating descriptions in Zebedee, as well as searching through a description database for specific queries.

Grammar modelling

This task focuses on representing the sentence structure. The main issues for sentence representation are:

- Flexible sign order
- Signing space representation
- Non-manual gestures
- Synchronisation of multiple body articulators
- Role shift
- Prosody

![DictaSign](dictasign.png)
It clearly constitutes more than could possibly be handled in the project lifetime. Understanding the need for priorities, the consortium decided to focus on a list of linguistic constructions.

**Grammar for recognition:** Models developed at the lexical level enable inferences to be made about features that vary during sign language performance. Examples include the locations at which signs are performed, and the movements linked to directional verbs that identify pronouns. Work has been done to build on partial recognition of signs that identifies a subset of linguistic features (expressed in HamNoSys). Using this approach, fully inflected sign sequences can be inferred for a class of sign language phrases.

**Grammar for generation:** This allows a more comprehensive inclusion of non-manual signing-components (chest, shoulders, head, gaze, and facial components: eye blinks, eyebrows, cheeks, mouth, tongue) in generation. Non-manual components occur during interaction and follow synchronisation rules that need to be described. This is the same for all body components, whether manual or non-manual. Thus, we have designed a description language, *Azalee*, to enable synchronisation of multi-tiered sign specifications, sign languages conveying simultaneous information (Figure 17). It was built from the Dicta-Sign corpus in an effort to generalise synchronisation patterns observed in the data, so should account for the linguistic nature and specificities of SLs.

![Figure 17 Three pieces of simultaneous signed information arranged on a time-line](image)

Corpora have been used intensively to build the linguistic models from scratch, but they have also been used to verify existing models. Most of the models were originally derived from the annotation of the LSF corpus. In the second stage, we verified some of our hypotheses on the other SL corpora. This verification resulted in refinements to existing hypotheses and implementation. This has also exhibited the common and differing properties for a given set of linguistic structures between the studied SLs. Three languages were included in the study, which constitutes the first cross-linguistic formalisation of structures pertaining to all levels of language. The use of grammar-based and corpora-based language models for SL generation is one of the originalities of the Dicta-Sign project. Building corpora-based language models is an entirely novel approach.

We have studied six structures, covering various levels from the lexical unit to discourse structure, and useful structures for use in the SignWiki which is the final Dicta-Sign demonstrator. The final items studied were:

- **enumerations**, sets of unordered elements signed in sequence
- **alternatives**, where options of a choice are listed sequentially
- **qualification/naming**, where one sign ($S_0$) in a sequence denotes an entity specified by the others (e.g. named, finger-spelled, adj-qualified)
- **neutral questions**, those where the speaker is genuinely asking for an answer which he is not able to predict
- **quantifiers** (small and big)
- **announcing titles**, announcing a topic of a discourse section

The first three cover parts of utterances, the fourth one entire clauses, the fifth is for lexical units and the last is in reference to a discourse structure.
For example, within alternative structures the observation relates to chin motion. The base of the chin reaches a new position towards the end of each option, it then moves down into a short rest, before being released for the next item (see Figure 18 for an example).

![Figure 18 Example of an alternation structure in LSF expressing two seating possibilities in a plane, near the window (1st image) or (2nd image) near the aisle (3rd image).](image)

It was interesting to find that for some of the structures (enumerations, alternatives, questions) observations made so far confirm not only the existence of invariants for the rules in a language, but also some generalisation across languages. For others (quantifiers, naming/qualifying), frequent properties were found, but no common and systematic features.

**WP5: Annotation tools**

The Dicta-Sign project proposed a complete solution that integrates various automatic tools into the annotation process, driven by the needs of the linguist annotators. The assumption is that the scientific community produces software that is specialised to specific needs and could be useful during the annotation process, but they are not designed to work together. These could be annotation tools for sign language studies or for more general-purpose animation and synthesis tools, image-processing or recognition tools. One of the challenges of the project is to take the software developed within the recognition workpackages and to integrate it in tools to aid manual annotation. Indeed, manual annotation is time-consuming, error prone and irreproducible. Moreover, the quality of the annotation depends on the experience and the knowledge of the annotator. In addition, for some measures (e.g. gaze direction), humans are not reliable. All this shows why automation of annotation, even partial, can be beneficial. Since such technologies are at the forefront of research, the knowledge (algorithms, code and even binaries) is sensitive or confidential. The solution proposed by Dicta-Sign is specified as a distributed-system architecture called Automatic Annotator Assistants System. It uses software developed in different programming languages, operating systems and software platforms. It also allows outsourcing of annotation tasks to other machines such as computing clusters. This solution introduces four agents:

- **Annotation Tool (AT):** annotation software on which the user is working. The user can select the appropriate functionality to apply to a given corpus and any existing annotation.
- **Automatic Annotator Assistant (A3):** software providing the automatic annotation service. Each A3 has a specific functionality and requires parameters relating to the data.
- **A3 Supervisor (A3S):** this acts like a service directory, inventorying available A3s.
- **Video File Server (VFS):** hosts videos in the corpus and manages access. It is used to share the video corpus between the AT and the A3.

The interaction between the different agents of the distributed system architecture is illustrated in Figure 19.
Initially, the AT sends a request to the A3S for the list of the available A3 functionality. It notifies the annotator who is then able to directly call any available A3. The exchange between the AT and the A3 can be implemented in two modes: synchronously or asynchronously. In the synchronous mode, the AT calls the function and waits for the result. The user has to wait for the result to continue his work (Figure 20). In the asynchronous mode, the AT calls the A3 and the answer is the remaining time estimated until the end of the processing. The AT can send a request for the results after this time has elapsed. The time can be seconds, hours or days, in the meanwhile the user can carry on with their work.

As one can see, the security of sensitive code used in A3s is provided by locating programs on remote servers. However, the security of the annotation data and the video corpus must be taken into account too. Indeed, annotations and videos are sensitive data in the same manner as A3 code. For example, videos are commonly under restrictive image rights. The security of data is designed over three axes: authenticity, authorisation and confidentiality.

- Authenticity enables agents to identify one another and to ensure that there is no impersonation. Transport Layer Security (TLS) is used, where each agent owns an identity certificate.
- Authorisation covers rights. Allowing each agent to choose whether to serve or not another agent. Since it hosts the videos, it is the VFS which centralises the video authorisation; it manages ‘trust’ in AT users and A3s.
- Confidentiality secures communications against listening agents. This is resolved with TLS which ciphers all communications.
From a software engineering point of view, a simple and open-source middleware, SOAP, is used to implement all the communications. We are using Annotation Graph (AG) [Bird2001] format for the data structure and an open data model inspired by Schmidt et al. [Schmidt2009] that facilitates interoperability and makes it possible to use various common, annotation-tool data-models. Thus, all data are represented using an XML description. The consistency of the data exchanged is guaranteed thanks to the specification of a Common Type Library (CTL).

There are implementations of an AS3 and a VFS, four prototypes of A3:

- Body Part Tracking
- Facial Feature Tracking
- Signing Detection
- Sign Segmentation

Access to these A3s has been added to the existing annotation tools AnColin and Ilex. The facial feature tracker from the work done in WP1 has been extended to work within the A3 framework. The results shown in Figure 3 were obtained using this tool. Also created was a signing detector to allow high level annotation of the corpus. This tool uses optical flow to detect when a subject is signing and when they are not. Examples of the optical flow extracted are shown in Figure 21, by analysing the quantity of optical flow in a scene, the tool can determine whether there is sufficient motion for it to be signing. This reduces the work of human annotators by allowing them to concentrate on the sections where the signing occurs.

In conclusion, we propose a distributed system architecture, which adds automatic processing support within existing annotation tools. This architecture is multi-platform and multi-language (including RealBasic, C/C++ and Java) and the model used for data exchange is adaptable to many annotation formats. We specified a distributed system architecture that includes security mechanisms to ensure annotation data, video files and program code integrity. All the specifications of API for A3 and the source code of A3 template and A3S are publicly available. The program code of the VFS and the AnColin annotation tool are open-source and available for the community. A VFS, an A3S and all the A3 that have been developed during the project Dicta-Sign, will be available on our server in the near future.
WP6: Sign Language Corpora

Within the Dicta-Sign project, sign language resources were compiled for four European sign languages: British, French, German, and Greek. These resources were used to inform progress in other research areas within the project, especially sign recognition, linguistic modelling, and sign generation.

As a first step, a multilingual lexical database providing a core lexicon of approximately 1000 entries in the four project sign languages was built. The shared list of concepts chosen for the lexicon is made up of signs of everyday use or those specifically related to the field of Dicta-Sign's main topic, European travel. Signs were recorded for each language and annotated by assigning gloss labels, form description (HamNoSys) and a rough meaning (see Figure 22).

The biggest achievement in the area of language resources within the project is Dicta-Sign’s multilingual corpus on the domain “Travel across Europe”. Prior to the project, parallel corpus collection for sign languages had only been undertaken in minimal sizes, or for spoken language simultaneously interpreted into several sign languages. Until now there hasn’t been a parallel corpus of semi-spontaneous signing by native signers. Due to the “oral” nature of sign language, and the risk of influences from written majority languages, the collection of parallel sign language data is a difficult task. The aim for the data collection within Dicta-Sign was to elicit data as authentic as possible under lab conditions. At the same time the data were supposed to be semantically close enough to be comparable across the different languages. In order to achieve this, elicitation tasks were specifically designed for this purpose. One key point in the planning of the data elicitation was to film Deaf informants in pairs, interacting with each other. Therefore, most tasks require the active involvement of both conversational partners, asking them to discuss and negotiate on certain topics, or to describe and explain things to the partner. The elicitation material presented to the informants included tasks such as advertising holiday destinations, explaining the procedures at the airport, and telling vacation stories [Matthes2010].
Data collection took place in all four countries involved in the project, using seven different cameras to film the informants from different perspectives (front, side, and top view) as well as additional stereo cameras that provided video footage for automatic processing. In each country, 14 to 16 informants were filmed in sessions lasting about two hours each. This resulted in applicable language data of at least 8 hours per language. After various steps of video post-processing the data were annotated using iLex, an annotation environment that is linked to a database [Hanke2008]. The detailed annotation, conducted for parts of the corpus data, includes segmentation of the continuous signing into individual signs, lemmatisation (i.e. assigning glosses), a form description of the signs using HamNoSys as well as English translation. In iLex lemmatisation is done by linking tags to entries in the database, which results in filling a video transcript and a growth of the sign language database at the same time. Additionally, content tags that reflect the topics, were assigned to most of the corpus data. These tags allow video sections with comparable content across individual informants and languages to be found.

The Dicta-Sign language resources are made available via a web portal that provides data on different access levels and with different approaches to access the data. The access levels range from publicly available data to restricted access for researchers or future project partners. Depending on
the level, the portal provides elicitation task material, video clips, transcripts, and informants' metadata in different levels of detail. Different approaches are offered to access the data, e.g. by selecting a certain language and informant, by choosing a certain task, or by selecting a specific content tag to see data of various languages and informants signing about the same topic.

**WP7: Bidirectional Integration**

Dicta-Sign has developed technologies to support sign language users, both by providing tools that can automatically recognise signs performed and also by providing synthesis technology that can be used to present animated sign language using realistic 3D virtual characters. The aim of WP7 is to bring together the results of different parts of the project to produce interactive prototypes which use sign language both for input and output in communication with deaf users. These prototypes build on all parts of the project since they rely on linguistic models of sign language during processing and exploit the multilingual lexicon and corpora.

Three prototypes have been developed:

- A search tool trained to identify the closest matches to a sign it is given
- A look-up tool which is able to display corresponding signs in several sign languages
- A Sign Wiki for collaborative development of sign language documents

It was decided early on that sign recognition would be improved by using 3D information about signs. The wisdom of this decision became clear when, during the lifetime of the project, the arrival of the Xbox Kinect made low-cost 3D input a reality for the general user. All the WP7 prototypes are based on the use of the Kinect for sign recognition, linked to modules that can synthesis signing using 3D virtual characters.

The search tool is trained to recognise signs in GSL. When a sign is presented, the tool seeks the closest matches in its training set and returns an ordered list of the identities of candidate signs. Each of the training signs has also been described in HamNoSys, which provides a precise definition of the hand postures and movements that make up a sign. The Dicta-Sign synthesis and animation system can take a HamNoSys description and generate a real-time animation of the sign. The user of the tool is therefore presented with a list of potential matches for their sign.

The look-up tool is an extension of the search tool. Feedback from expert deaf evaluators was used to make the user interface more intuitive for deaf users. Also, more importantly, the system was extended to recognise signs from DGS as well as GSL. As the signs used in training are in the Dicta-Sign multilingual lexical database, which contains corresponding sign in all four sign languages studied by the project, a selected sign can be presented in BSL, DGS, GSL, and LSF by using the

![Figure 25 Schematic of Search and Look-up Tools](image-url)
HamNoSys descriptions stored in the database. This provides a rudimentary translation tool which can be used in two directions: a user can perform a sign in their own language to see what they should sign when visiting another country; or the user can perform an unknown sign seen while away from home and could look up the corresponding sign in their own language.

The overall operation of the tool is given by the diagram in Figure 25.

The leftmost section refers to an offline training process. In operation, the tools involve a client and server module. The server processes video and depth data from the Kinect and uses parameters from the training process to match signs against the training set in a chosen language. The client receives a ranked list of glosses, or identifiers for signs. These are looked up in the multilingual lexicon and the relevant HamNoSys (HNS) descriptions are used to generate an animation.

**Sign Look-Up Tool**

Both an offline and online version of the sign look up tool were created. The offline version uses a query video to search a dictionary database of signs and returns the top 10 matches, like a search engine. This was based on the information extracted in the previous sections and used a bank of Markov Chains to classify examples. This was then extended to use the Kinect sensor as an input, it was based on the 3D features extracted and the more robust Sequential pattern learning to make it user independent and capable of working in a variety of environments.

**Kinect Server**

The Kinect server was built on the work of the Sign Look-Up Tool and communicates with both the sign wiki and the continuous sign recognition module. It extracts features relating to the signer as well as some HamNoSys information, it can either use this internally to recognise dictation style signs or pass the information on to the continuous sign module for recognition of signs with linguistic modifications. The server has two Kinect-buttons which the user can K-click to start and stop data input as shown in Figure 26.

![Figure 26 Kinect Buttons being K-clicked by a user.](image)

**Presentation Client**

On the client side of the look-up tool there are two main presentation modes, between which the user may switch interactively as desired:

- Sign Matches Display mode, in which the ranked list of possible lexicon matches for signs in chosen input sign language is presented to the user by the virtual human signer;
- Sign Translation Display mode, in which the user can view translations, into the four supported sign languages, of any individual sign suggested by the matching process.

We show the enhanced user interface developed for the look-up tool in Figure 27. The Kinect server may deliver a single candidate or many candidates, if a number of signs have similar features. Up to four possible matches can be displayed at a time and will be animated simultaneously:
In Figure 28 we illustrate the translation feature, in this case signs for “gun” which are (unsurprisingly) very similar in all four languages.

**Sign-Wiki**

The Sign-Wiki is the most advanced of the prototypes. Since this was felt to be of most interest to the Deaf community, it forms the basis of the Project Demonstrator developed by WP8. The aim is that a Kinect device would be the main source of sign language input, but where that is not possible the user can take advantage of the multilingual lexicon using the lexicon search facility shown in Figure 29.

Having selected the sign language they wish to use, the user can enter text in a search box. The text can be a full word or part of a word. If the search text matches all, or part, of the name or gloss, of a sign, that sign will be shown as an option. The sign can either be animated by clicking on its icon, or dragged into the sign language document.
In addition, the lexicon has a list of synonyms, in English, for the concepts represented by each sign. The synonym list is matched as well as the sign gloss. The example given shows that a user may be presented with a range of options which can be helpful if they are not sure what word would best characterise a sign.

![Figure 29 Sign-Wiki Search Mode](image)

Early Sign-Wiki prototypes were evaluated by deaf experts and as a result of feedback some changes were made to the user interface of the Sign-Wiki to form the project demonstrator. A comprehensive description of the Sign-Wiki is given in the next section.

**Evaluation**

The underlying philosophy of the design and testing process of each prototype was to involve Deaf signers at each step of the project. Thus, the evaluation was designed to be as close as possible to real use cases. The evaluation focuses on using the prototypes in context. This means that all tests are based on real sentences created by experienced signers. The evaluation was also based on several tasks that are close to the intended final use of the prototype being tested (content understanding, translation of a message, creation of wiki content).

The following evaluation tests were performed with help from the Deaf community:

**Avatar**: Among both Deaf, and Hearing signers learning sign language; in various situations (context-less / interaction); isolated signs and sentences; and varying the conditions of visualisation (speed point of view). These evaluations revealed that comprehension is sensitive to the speed; it had to be slowed down for isolated signs. The use of the Dicta-Sign avatar currently results in a lower comprehension level than a real signer. Facial expression is required but it should be noted that it only augments comprehension if it is accurate enough, and well synchronised with the hand gestures.

**Look-up tool**: The user had to use the tool in order to translate sentences from GSL to LSF. These tests revealed that the tool was usable as a dictionary.

**Wiki**: Evaluation performed online where the users watch a short sentence signed by a virtual signer, answer some questions about this sentence, create and modify the wiki content. The main interface issues are related to the sign language video display. The outline display was well understood by 75% to 85% of the evaluators. However, the display of low-level items could be enhanced. Regarding new sign creation, hand orientation user friendliness should be improved.

The results of the evaluation process prove that the Dicta-Sign achievements lay the foundations for various applications such as:
**SL learning field:** It would be possible to create large-scale databases that could help individuals learn sign language. The sign databases could be embedded in Dictionaries or lexicons. For this to be the case the evaluation underlines the need for sign validation (both for the face expression and the body movement).

**Computer Games:** The Dicta-Sign virtual signer could be used to convey simple messages within a gaming context. For the moment, comprehension is limited for full sentences. This limitation may be caused by a lack or an incoherence of the non-manual expression with the body movement.

**Anonymisation:** This is often presented as one of the ultimate goals of sign language recognition. Current applications would need much more accuracy to be able to reproduce the signer non-manual components before this could be realised.

**Automatic translation:** This could be done either from SL to a written language or to another SL. Even if it is currently out of reach for sentences, Dicta-Sign has made the first steps forward in enabling a sign-to-sign translation for a large set of signs. With even a small set of signs this would be possible with a constrained domain; an application such as a translator that increases understanding of the Internet videos from the European Union of the Deaf for example.

**WP8: Project Demonstrator**

In workpackage 8 the task was to develop the project demonstrator prototype, taking on board the results of the evaluations and showcasing all technological advancements and research work executed in order to obtain the project goal; enabling sign language in Human-Computer Interaction in the Web 2.0.

The Sign-Wiki demonstrator shares the technical characteristics and GUI design of the Sign-Wiki prototype application. It incorporates all the properties to serve the Dicta-Sign proof of concept purpose. Deaf internet users have an environment that allows them for the first time to create, edit, modify and upload sign language texts.

An example of the wiki in edit mode is shown in Figure 30, here the user can drag and drop signs around or delete content. The wiki also incorporates a sign builder mode (Figure 31) which allows a user to modify an existing sign at the lexical sub-unit level. The example shown is changing the orientation of the palm in the sign for ‘me’.
Figure 31 Sign Builder – Selecting the Palm Orientation

The environment of the Sign-Wiki allows for a number of input modes which exploit visual means, while it provides representation of sign utterances by a virtual signer (avatar) exploiting sign synthesis and animation technologies. The input modes include the search facility described in the previous section as well as the Kinect input mode shown in Figure 32. The signer can then enter their content either using a dictation style sign or continuous sign for a more constrained lexicon.

Figure 32 Kinect input mode

Once the signer has recorded their signs they are processed by the recognition engines. The server then sends the recognised signs to the wiki which displays them to the user (Figure 33). At this point the user can choose to accept the sentence as it has been recognised, pick and choose from the recognised signs or re-do the Kinect input from the beginning.
The whole interaction process integrates sign recognition and sign synthesis technologies, while it exploits properly coded sign language resources of significant volume and makes provisions for enhancements based on implementation of linguistic models.

Finally, since the Sign-Wiki incorporates multilingual data, it allows for the application of simple sign translation services, which provide rough translations of input data into the project’s sign languages. This is demonstrated in Figure 34, where the top row shows content in GSL and the bottom the equivalent content in DGS.

The Dicta-Sign Sign-Wiki provides Deaf Internet users the opportunity to communicate using their natural language; it replaces videos with a module that allows dynamic creation and editing of sign language messages, while preserving anonymity of content creators.
**Bibliography**


Potential impact and main dissemination activities and exploitation of results

Potential impacts

In the framework of the declared European policies regarding Equal Opportunities for All Citizens, e-inclusion has become one of the main requirements in view of the Information Society. It is of upmost importance that new forms of communication should find their role in society at large without new barriers evolving with them. Dicta-Sign has researched the key technologies required to promote sign language communication. These include providing Web 2.0 services and other HCI technologies to Deaf sign language users, an important linguistic minority in Europe; and up till now excluded from these valuable developments.

As the field of sign language technology is still young, it has been beyond the scope of a 3-year STREP to catch up completely with mainstream language technology and to deliver end-user products. Dicta-Sign concentrated on significantly advancing enabling-technologies by a multidisciplinary approach; allowing them to come close enough so that designers of future natural language systems can take sign languages into account. It has therefore been vital to disseminate the results of the project, not only within its home disciplines, but to the wider community.

This has not only been done by the respective activities of those partners involved in general language technology (esp. ILSP and CNRS-LIMSI), but also by means of active dissemination of the research results of all scientific partners, as well as via targeted dissemination events towards industry, the wider public and of course the Deaf Community on both a national and European level.

Furthermore, the project’s impact will be presented at the 5th Workshop on Sign Language Representation and Processing (LREC 2012), which will take place after the end of the project (May 2012), while an edited volume will gather all major project scientific and technical achievements.

The tools integrated in WP7 are demonstrated in WP8 and also as part of the WP9 activities, fitting into the scenarios expected by the attendees from natural language technology and the ECA (Embodied Conversational Agents) communities.

Apart from the major academic institutions which have expressed their interest in scientific collaboration from the beginning of the project, several organisations active in virtual agents as well as telecom operators and broadcasters have already expressed their interest in Dicta-Sign technology, and this interest can be expected to attract further members of the industry after the end of the project and the dissemination of the achieved results. The interest of service providers to the Deaf community to get early access to the technology to be developed in Dicta-Sign, in order to produce end-user products they consider essential for their customers, is strong evidence for the success of this approach.

The areas of image processing, computer vision, language technology and virtual human technology are all research areas where advances have significant impact on contemporary society, with respect to the kind of data stored, and the manner in which humans interact with computers. The integration of these technologies is essential to support bi-directional communication in sign language. The combination of leading researchers in these areas has uniquely equipped the Dicta-Sign consortium to make significant progress to assist deaf people in the use of new and developing computer systems.

Robustness of the technologies involved is a key success factor and therefore has taken an important role in the development especially of sign language recognition within Dicta-Sign. The novel approach to combine cognitive vision with knowledge about language has explored the question of what is achievable by artificial cognitive models.
At the end of the Dicta-Sign project, its impact can be demonstrated by the prototype products and services that have been built using Dicta-Sign technologies.

Due to the nature of sign languages as small minority languages, the development of systems such as those which have been envisaged by Dicta-Sign could not feasibly be funded nationally for each individual sign language, but could only be achieved on a European level. While European sign languages are generally considered much closer to each other than many European oral languages, they differ from each other both in grammar and lexicon. Dicta-Sign has therefore treated four sign languages in parallel, to make sure that the methods developed are general enough to cover a multitude of sign languages – hopefully most European sign languages and many beyond. The experiences gained and methodologies created in this project will then also make the construction of comparable language resources for other sign languages easier and more efficient. The establishment of parallel linguistic resources in multiple sign languages will give scope for the use of corpus-based sign language processing techniques which have previously lacked the necessary support.

With three different avatar technologies from Dicta-Sign partners alone, interoperability between different avatars has been a key issue for this project. Dicta-Sign aimed to go beyond that and cluster with the developers of other signing avatar technologies existing worldwide. Towards this end consortium members have been involved in organising and participation in two focused events, namely: the First International Workshop on Sign Language Translation and Avatar Technology. (Berlin, 10-11 January 2011) and the Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT). (23 October 2011, University of Dundee, UK)

Several parties have already suggested that the Hamburg Notation System and its XML counterpart SiGML be proposed as candidates for international standards. Thus far, partners have been hesitant about this step due to the comparatively small number of sign languages for which SiGML has been used. With two more languages extensively used in Dicta-Sign (LSF and GSL) and the interoperability problems addressed, Dicta-Sign partners are now in a strong position to propose its sign animation description as a W3C standard and the HamNoSys glyph set for integration to Unicode.

The next step beyond interoperability of existing signing avatar technology is the capability to sign for any avatar technology available. In order to come closer to this goal, Dicta-Sign will publish specifications for avatars to be fulfilled on top of established standards such as H-Anim or MPEG-4 in order to be driven by Dicta-Sign or compatible generation modules. These specifications will then hopefully be integrated in future H-Anim revisions.

Conversely, Dicta-Sign has used proposed standards from related fields such as gesture research and Embodied Conversational Agents (ECAs), such as APML, and having reported its findings has thereby contributed to the maturing of these standards.

ILSP/ATHENA RC and UHH members of the Dicta-Sign consortium are stakeholders in the recently approved European Research Infrastructure CLARIN and will continue cooperation in this context, especially with respect to ISOCat for sign languages. The Dicta-Sign corpus was selected (board approval pending) as a best-practice example on multimodal resources for CLARIN-D (the German part of CLARIN funded by the federal government until at least 2014). This will allow us, also in cooperation with the University of Tübingen, Center for Sustainability of Linguistic Data (NaLiDa), to keep the Dicta-Sign corpus up-to-date in the light of upcoming developments especially in the area of metadata and processing workflows.

Dicta-Sign partners are in close contact with leading academic institutions in the USA, and both sides are eager to coordinate research on a long term basis.
Main dissemination activities and exploitation of results

Dicta-Sign has had a strong dissemination strategy, which has always been viewed as a crucial forerunner for further exploitation of the project outcomes. This is reflected in the fact that a whole work-package devoted to dissemination and exploitation activities has been included in the project’s work plan (cf. section 1.3), and a purpose-formed committee that has managed, coordinated and effectuated these efforts has been designed as part of the management strategy of the project. In particular, the dissemination strategy of Dicta-Sign has been deployed along the axes: **dissemination channels** (e.g. internet, mass-media, scientific publications), **dissemination events** (e.g. conferences, workshops, symposia, lectures, debates) and **open-access strategies** (e.g. public releases of data).

The Dicta-Sign project covers a broad scientific spectrum, inter-relating sign languages technologies and linguistic study that have promoted Deaf accessibility in the environment of HCI applications. Repercussions have been awaited in knowledge in the various disciplines but also in methodology for the transverse exploitation of this knowledge in the project’s WPs.

Dissemination of this knowledge at the scientific community has been done with presentation of research outcomes in conferences interested by this topic. The events where Dicta-Sign research and development results were presented during the projects lifecycle are:

- The 8th International Gesture Workshop (GW-2009), held in Bielefeld- Germany on February 25-27, 2009
- The IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2009), held in Taipei-Taiwan on April 19-24, 2009
- The SLCN Workshop 2: *Metadata*, Nijmegen, the Netherlands, November 2009
- The 5th International Conference on Multimedia, Information and Communication Technologies in Education (m-ICTE 2009), April 09, Lisbon, Portugal.
- The Universal Access in Human-Computer Interaction (UAHCI) Conference, hosted by the HCII-09, July 09, San Diego, USA
- The CerLiCO, Poitiers, June 09
- The ICTA Conference, Tunisia, May 09
- The IEEE workshop on Video-Oriented Object and Event Classification, Kyoto, Japan.
- The IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, CVPR09
- International Conference on Acoustics, Speech and Signal Processing (ICASSP-2010), Dallas, Texas, USA, March 2010
- LREC-2010
- 4th Workshop on “Representation and Processing of Sign Languages: Corpora and Sign Language Technologies”, satellite workshop to LREC-2010
- International Conference ECCV 2010, Crete, September 2010
- “Workshop on Sign, Gesture and Activity” Satellite workshop of the ECCV 2010
- The SLCN Workshop 3: “*Annotation*”, Stockholm, Sweden, 14-16 June 2010.009
- The SLCN Workshop 4: “*Exploitation and publication of signed corpora*”, Berlin,
Germany, 3-4 December, 2010
- International Conference on Image Processing, Hong-Kong, September 2010.
- Traitement Automatique des Langues Naturelles (TALN), Montréal, Canada, 20-23 July.
- Traitement Automatique des Langues des Signes (TALS), Montréal, Canada, 20-23 July.
- Signed Language Phonology Workshop, 6-8 July 2011, Université de Bretagne-Sud, Vannes, France.
- 2nd Sign Language Translation and Avatar Technologies Workshop (satellite to ACM ASSETS), 23 October 2011, Dundee, UK.
- Reconnaissance des formes et Intelligence Artificielle (RFIA 2012), 24-27 January 2012, Lyon, France.
- Sign Language Translation and Avatar Technologies Workshops, 2011.
- European Signal Processing Conference (EUSIPCO2011).
- Int. Conference Computer Vision, ICCV2011
- IEEE Workshop on Consumer Depth Cameras for Computer Vision 2011.
- 7th International Conference on Computer Vision Theory and Applications (VISAPP’12).
- IEEE International Workshop on Social Behavior Analysis 2011
- IEEE Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams 2011
- British Machine Vision Conference BMVC’11
- Pattern Recognition and Image Analysis. 5th Iberian Conference, IbPRIA 2011. Best Paper Award.
- IEEE Workshop on Consumer Depth Cameras for Computer Vision 2011
- IEEE Workshop on Human Interaction in Computer Vision 2011
- Computer Vision and Pattern Recognition Conference CVPR 2012
- The 10th International Conference on Greek Linguistics, 1-4 September 2011, Komotini, Greece.
- COST Action: IS1006 - Unraveling the grammars of European sign languages: pathways to full citizenship of deaf signers and to the protection of their linguistic heritage. Meeting 15-16/11/2011, Georg-August-Universität, Göttingen, Germany.

The co-ordinator and the partners have further made known the project’s scientific impact and outcome to the international scientific community via organisation of related workshops in
outstanding international conferences and scientific events and, at the same time, submitting their work for publication in renowned international journals.

Training material resulting from the project has been gathered and used by the academic partners as part of their course materials at postgraduate level. The research activities within the project are also leading to PhD or post-doctoral degrees for some of the participants.

Dicta-Sign has led to knowledge with direct impact for the development of several tools, including tools for sign recognition, tools for image processing, tools for sign synthesis, tools of annotation and editors of linguistic models (lexicon, grammar, signing space) as well as the project’s demonstration showcase and laboratory prototypes, allowing to test the integration of the models and the methods developed in the project.

This knowledge is shared by the consortium members but the set of showcase and laboratory prototypes will be presented to industry in scientific events, complemented with exhibitions and will also be put at the disposal of the project supporting organisations and the Deaf community. Dissemination of the project outcomes will be facilitated through the adoption of “open-access” processes, which will allow the research community to build on findings and achievements within Dicta-Sign. We believe that generation, distribution and access to open research content are crucial for promoting continuous state-of-the-art innovation in the European research landscape.

The prototypes have been used by deaf people to test their usability but also to showcase new practices (access to information in sign language, mode of writing of the sign language, generation in sign language) and thus to encourage new requirements in term of applications (tools usable by general public) and in scientific questions.

Dissemination to scientific community and related industry has established the means and procedures to communicate the scientific advancements of the project to the international scientific community, to support organisations, other interested organisations and standardisation bodies.

The list of contacts includes major international Deaf educational institutions heavily involved in research on sign language and related technologies (i.e. Boston University and Gallaudet University, USA, the Kyoto Institute of Technology, Japan and The Beijing Academy of Science, China). All potentially interested parties have been invited to actions of promotion of project scientific results, organised in the framework of national and international conferences and exhibitions. They have also been invited, and will continue to be invited, to participate in workshops organised by the consortium, in the framework of scientific events most attractive to industry, due to presentation of innovative technologies. Partners have also used their contacts with media to ensure that the public in general, and the Deaf communities at large, are kept informed about the project and its aims (e.g. by means of broadcast features), rather than only those Deaf involved in scientific contexts or those contacted as language consultants.

Moreover, the consortium will further keep a steady contact after the project’s end with those organisations having expressed interest and support for Dicta-Sign from the beginning of the project. This group has been open also to other organisations that were sensitised via dissemination activities in the course of the project’s lifecycle.

The partners have also used their existing national contacts as well as contacts to magazines and other journals of wide interest, to promote the project’s outcomes through popularisation.

Aiming at acceptance by Deaf communities, Dicta-Sign has organised familiarisation events with the project’s national deaf communities as a complementary action to communication with related industry and scientific community, since approval by final users is essential to strengthening the commercialisation potential of the project’s outcomes, the latter being a major perspective within the consortium.
Project news, views and findings/results have been reflected in two newsletter releases which have reached a wide audience and the Deaf communities of the project countries.

In short, Dicta-Sign dissemination objectives have been targeted towards:

- Acceptance by partner countries’ Deaf communities of the tools developed within the project
- Acceptance by international scientific community, industry and standardisation bodies of the project’s scientific innovations
- Communication of innovations and feedback from organisations interested in signed content production
- Popularisation of project’s scientific innovations

Complementary to scientific events, communication with industry and acceptance by the users related events, the Dicta-Sign web site (http://www.dictasign.eu/Main/HomePage) setup from the early stages of the project, has functioned as a communication forum for the consortium, and as a means for reaching the wider public and the research and academic community, providing multilingual material (written language and sign language videos) and reflecting the project’s aims, research progress, scientific impact, etc.

**Exploitation**

Because of the nature of the research in Dicta-Sign, i.e. basic research on scientific and technological fronts rather than applied research with industrial participation, dissemination activities are in focus within the project’s structure and workplan; still, exploitation opportunities are explored, investigating the employment of the project research outcomes within commercial applications (e.g. signing avatar applications, Deaf communication and education tools). Details on the project exploitation plan are available in the related deliverable, which provides a platform for the exploitation of the project’s outcomes far beyond the project lifecycle.

Regarding management of knowledge and of IPRs, the PCC has already worked on the consortium agreement addressing issues such as:

- Intellectual Property Rights, confidentiality issues, voting rules, monetary obligations, etc.
- Specific confidentiality issues should they arise.
- Software rights issues, existing and future.

This has been proven to be sufficient to meet the legal obligations taken on by the consortium in undertaking the project. Subject to further agreement will be conditions of possible future commercialisation of the prototypes and resources deriving from the project. Issues like monetary obligations, distribution of any surpluses, rights to further developments of the deliverables, etc. will be then faced.
## 4.2 Use and dissemination of foreground

### Section A (public)

Template A1: list of scientific (peer reviewed) publications. Sorting is made on the basis of year of publications, since a significant number of Dicta-Sign publications can be characterised as important and also they may belong to different scientific domains.

<table>
<thead>
<tr>
<th>No</th>
<th>Title</th>
<th>Main author</th>
<th>Title of the periodical or the series</th>
<th>Number, date or frequency</th>
<th>Publisher</th>
<th>Place of publication</th>
<th>Year of publication</th>
<th>Relevant pages</th>
<th>Permanent identifiers (if available)</th>
<th>Is/Will open access provided to this publication??</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sign Language Recognition using Sub-Units</td>
<td>Cooper H, Bowden R</td>
<td>Special Topic on Gesture Recognition &amp; Pattern Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accepted to appear</td>
<td>Yes</td>
</tr>
</tbody>
</table>

---

3 A permanent identifier should be a persistent link to the published version full text if open access or abstract if article is pay per view) or to the final manuscript accepted for publication (link to article in repository).

4 Open Access is defined as free of charge access for anyone via Internet. Please answer "yes" if the open access to the publication is already established and also if the embargo period for open access is not yet over but you intend to establish open access afterwards.
<table>
<thead>
<tr>
<th>No</th>
<th>Title</th>
<th>Authors</th>
<th>Conference/Proceedings</th>
<th>Date</th>
<th>Venue</th>
<th>Pages</th>
<th>DOI</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Title</td>
<td>Authors</td>
<td>Venue/Conference</td>
<td>Date</td>
<td>Location</td>
<td>Year</td>
<td>Pages/Volume</td>
<td>Conference/Website</td>
</tr>
<tr>
<td>----</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------</td>
<td>--------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>15</td>
<td>MIMiC: Multimodal Interactive Motion Controller</td>
<td>Okwechime D, Ong E J, Bowden R,</td>
<td>IEEE Trans. on Multimedia</td>
<td>Vol 13(2)</td>
<td>IEEE</td>
<td>2010</td>
<td>255-265</td>
<td>10.1109/TMM.2010.2096410</td>
</tr>
<tr>
<td>16</td>
<td>Social Interactive Human Video Synthesis</td>
<td>Okwechime D, Ong E J, Gilbert A, Bowden R,</td>
<td>Procs. of Asian Conf. on Comp. Vision</td>
<td>Vol 6492</td>
<td>Springer LNCS</td>
<td>2010</td>
<td>253-266</td>
<td>10.1007/978-3-642-19315-6_20</td>
</tr>
<tr>
<td>17</td>
<td>Facial Expression Recognition using Spatiotemporal Boosted Discriminatory Classifiers</td>
<td>Moore S, Ong E J, Bowden R,</td>
<td>Procs. of International Conf. on Image Analysis &amp; Recognition</td>
<td>Vol 6111</td>
<td>Springer LNCS</td>
<td>2010</td>
<td>405-414</td>
<td>10.1007/978-3-642-13772-3_41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>in Deaf Communication</td>
<td>Bowden, A., Braffort, C. Collet, J. Maragos and F. Goudenove</td>
<td>Languages : Corpora &amp; Sign Languages Technologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Title</td>
<td>Authors</td>
<td>Conference/Editor</td>
<td>Date</td>
<td>Location</td>
<td>Web Link</td>
<td>Status</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------------</td>
<td>-----------------------</td>
<td>------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Head Tracking and Hand Segmentation during Hand over Face Occlusion in Sign Language</td>
<td>Gonzalez, M., C. Collet and R. Dubot</td>
<td>Workshop on Sign, Gesture and Activity (SGA), 11th European Conference on Computer Vision (ECCV)</td>
<td>Septem ber 2010</td>
<td>Heraklion, Crete, Greece</td>
<td>2010</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Corpus de langue des signes : premières réflexions sur leur conception et leur représentativité</td>
<td>Segouat, J., A. Braffort and A. Choisier</td>
<td>Travaux linguistiques du Cerlco, L’exemple et le corpus, quel statut</td>
<td>23</td>
<td>Presses Universitaires de Rennes</td>
<td>Rennes 2010</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Diva, une architecture pour le support des agents gestuels interactifs sur Internet</td>
<td>Braffort, A., J.-P. Sansonnet, J.-C. Martin and C. Verrecchia</td>
<td>Technique et Science Informatiques, (Agents Conversationnels Animés)</td>
<td>29/7</td>
<td>2010</td>
<td>777-806</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Sign Language Synthesis : Skeleton Modelling for More Realistic Gestures</td>
<td>Delorme, M</td>
<td>ACM SigAccess Newsletter</td>
<td>96</td>
<td>2010</td>
<td><a href="http://www.springerlink.com/content/978-3-642-12552-2/?MUD=MP">http://www.springerlink.com/content/978-3-642-12552-2/?MUD=MP</a></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Toward Modelling Sign Language Coarticulation</td>
<td>Segouat, J. and A. Braffort</td>
<td>Gesture in Embodied Communication and Human Computer Interaction, Lecture Notes in Computer</td>
<td>Vol. 5934</td>
<td>Springer-Verlag</td>
<td>Bielefeld, Germany 2010 325-336</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>Authors</td>
<td>Conference/Event</td>
<td>Volume</td>
<td>Publisher</td>
<td>Year</td>
<td>Pages</td>
<td>URL</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>----------------------------------------------</td>
<td>--------</td>
<td>--------------------</td>
<td>------</td>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Title</td>
<td>Authors</td>
<td>Journal/Proceedings</td>
<td>Volume/Issue</td>
<td>Year</td>
<td>Start/End</td>
<td>Digital Object Identifier</td>
<td>Status</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>--------------</td>
<td>------</td>
<td>-----------</td>
<td>----------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>47</td>
<td>Sign Language Recognition</td>
<td>Cooper H, Bowden R</td>
<td>Looking at People: Automatic visual analysis of humans</td>
<td>Part 4</td>
<td>2011</td>
<td>539-562</td>
<td>10.1007/978-0-85729-997-0_27</td>
<td>yes</td>
</tr>
<tr>
<td>48</td>
<td>Kinecting the Dots: Particle Based Scene Flow From Depth Sensors</td>
<td>Hadfield S, Bowden R</td>
<td>Procs. of International Conf. on Comp. Vision</td>
<td>6-11 Nov</td>
<td>2011</td>
<td>2290-2295</td>
<td>10.1109/ICCV.2011.6126509</td>
<td>yes</td>
</tr>
<tr>
<td>49</td>
<td>Capturing the Relative Distribution of Features for Action Recognition</td>
<td>Oshin O, Gilbert A, Bowden R</td>
<td>Procs. of International Conf. on Automatic Face &amp; Gesture Recognition &amp; Workshops</td>
<td>21-25 March</td>
<td>2011</td>
<td>111-116</td>
<td>10.1109/FG.2011.5771382</td>
<td>Yes</td>
</tr>
<tr>
<td>50</td>
<td>Visualisation &amp; Prediction of Conversation Interest through Mined Social Signals</td>
<td>Okwchime D, Ong E J, Gilbert A, Bowden R</td>
<td>Procs. of International Conf. on Automatic Face &amp; Gesture Recognition &amp; Workshops</td>
<td>21-25 March</td>
<td>2011</td>
<td>951-956</td>
<td>10.1109/FG.2011.5771380</td>
<td>Yes</td>
</tr>
<tr>
<td>51</td>
<td>Learning Sequential Patterns for Lipreading</td>
<td>Ong E J, Bowden R</td>
<td>Procs. of British Machine Vision Conf.</td>
<td>29 Aug - 2 Sep</td>
<td>2011</td>
<td>55.1-55.10</td>
<td>10.5244/C.25.55</td>
<td>Yes</td>
</tr>
<tr>
<td>52</td>
<td>Learning Temporal Signatures for Lip Reading</td>
<td>Ong E J, Bowden R</td>
<td>Procs. of International Conf. on Comp. Vision Workshop: Analysis &amp; Retrieval of Tracked Events &amp; Motion in Imagery Streams</td>
<td>6-11 Nov</td>
<td>2011</td>
<td>958-965</td>
<td>10.1109/ICCVW.2011.6130355</td>
<td>Yes</td>
</tr>
<tr>
<td>53</td>
<td>Reading the Signs: A Video Based Sign Dictionary</td>
<td>Cooper H, Pugeault N, Bowden R</td>
<td>Procs. of International Conf. on Comp. Vision Workshop: Analysis &amp; Retrieval of Tracked Events &amp; Motion in Imagery Streams</td>
<td>6-11 Nov</td>
<td>2011</td>
<td>914-919</td>
<td>10.1109/ICCVW.2011.6130349</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Retrieval of Tracked Events &amp; Motion in Imagery Streams</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>There is more than one way to get out of a car: Automatic Mode Finding for Action Recognition in the Wild</td>
<td>Oshin O, Gilbert A, Bowden R,</td>
<td>Procs. of Pattern Recognition &amp; Image Analysis. 5th Iberian Conf, IbPRIA 2011</td>
<td>Vol 6669</td>
<td>Springer LNCS Las Palmas de Gran Canaria, Spain 2011 41-48</td>
<td>10.1007/978-3-642-21257-4_6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Notes:**
- **Yes** indicates that the publication is included in the project report.
- **No** indicates that the publication is not included in the project report.
<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Authors</th>
<th>Conference/Workshop</th>
<th>Date</th>
<th>Page(s)</th>
<th>URL</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>Robust body parts tracking using particle filter and dynamic</td>
<td>Gonzalez, M. and C. Collet</td>
<td>IEEE International Conference on Image Processing</td>
<td>11-14 September 2011</td>
<td>537-540</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Template</td>
<td>Title</td>
<td>Authors</td>
<td>Workshop</td>
<td>Location</td>
<td>Year</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>---------</td>
<td>----------</td>
<td>----------</td>
<td>------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>Formalism for multi-linear rules of Sign Language grammars</td>
<td>Filhol, M.</td>
<td>Formal and experimental advances in Sign Language theory (FEAST)</td>
<td>Venice, Italy</td>
<td>2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Phonetic Encoding in Sign Language Corpus Annotation</td>
<td>Hanke, T. and S. Matthes</td>
<td>Signed Language Phonology Workshop</td>
<td>Vannes, France</td>
<td>2011</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>Deixis in interaction: a multimodal approach</td>
<td>Koutsombera, M.</td>
<td>10th International Conference on Greek Linguistics</td>
<td>Komotini, Greece</td>
<td>2011</td>
<td><a href="http://www.icgl.gr/">http://www.icgl.gr/</a> (abstracts)  Yes</td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Authors</td>
<td>Title</td>
<td>Conference/Event</td>
<td>Date</td>
<td>Location</td>
<td>Publication/URL</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>------------------</td>
<td>------</td>
<td>----------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>Ong EJ, Cooper H, Pugeault N, Bowden R</td>
<td>Sign Language Recognition using Sequential Pattern Trees</td>
<td>Proc. of Comp. Vision &amp; Pattern Recognition Conf.</td>
<td>June 18-20</td>
<td>Providence, Rhode Island, USA</td>
<td>To appear Yes</td>
<td></td>
</tr>
</tbody>
</table>


http://hal.archives-ouvertes.fr/view_by_stamp.php?&halsid=j6ggabhq


<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
<th>Conference Details</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>Title</td>
<td>Authors</td>
<td>Conference/Journal</td>
<td>Date</td>
</tr>
<tr>
<td>-----</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>90</td>
<td>A combination of two synchronisation methods to formalise Sign</td>
<td>Filhol, M.</td>
<td>Gesture in Embodied Communication and Human Computer Interaction, Lecture Notes in Computer Science/Lecture Notes in Artificial Intelligence</td>
<td>Vol. 7206</td>
</tr>
<tr>
<td>Paper ID</td>
<td>Title</td>
<td>Authors</td>
<td>Journal/Media</td>
<td>Volume/Publication Date</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>92</td>
<td>Computer modelling</td>
<td>Safar, E. and J. Glauert</td>
<td>Handbook of Sign Language Linguistics</td>
<td></td>
</tr>
<tr>
<td>NO.</td>
<td>Type of activities</td>
<td>Main leader</td>
<td>Title</td>
<td>Date</td>
</tr>
<tr>
<td>-----</td>
<td>-------------------</td>
<td>-------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Workshop</td>
<td>Richard Bowden, Philippe Dreuw, Petros Maragos, Justus Piater,</td>
<td>International Workshop on Sign, Gesture and Activity</td>
<td>11 Sep 2010</td>
</tr>
<tr>
<td>3</td>
<td>Workshop</td>
<td>Philippe Dreuw, Eleni Efthimiou, Thomas Hanke, Trevor Johnston, Gregorio Martinez Ruiz, Adam Schembri</td>
<td>4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies</td>
<td>22-23 May 2010</td>
</tr>
<tr>
<td>4</td>
<td>Workshop</td>
<td></td>
<td>1st International Workshop on Sign Language Translation and Avatar Technology</td>
<td>10 – 11 January 2011</td>
</tr>
<tr>
<td>5</td>
<td>Workshop</td>
<td></td>
<td>2nd International Workshop on Sign Language Translation and Avatar Technology</td>
<td>23 October 2011</td>
</tr>
<tr>
<td>6</td>
<td>Newsletters</td>
<td>François Lefebvre-Albaret, Jérémie Segouat</td>
<td>May 11</td>
<td>May 2011</td>
</tr>
</tbody>
</table>

5 A drop down list allows choosing the dissemination activity: publications, conferences, workshops, web, press releases, flyers, articles published in the popular press, videos, media briefings, presentations, exhibitions, thesis, interviews, films, TV clips, posters, Other.

6 A drop down list allows choosing the type of public: Scientific Community (higher education, Research), Industry, Civil Society, Policy makers, Medias ('multiple choices' is possible.)
<table>
<thead>
<tr>
<th>#</th>
<th>Newsletters</th>
<th>Authors</th>
<th>Date</th>
<th>Location</th>
<th>Target Audience</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Newsletters</td>
<td>François Lefebvre-Albaret, Jérémie Segouat</td>
<td>December 11</td>
<td>December 2011</td>
<td>Civil society, Deaf community, Medias</td>
<td>International</td>
</tr>
<tr>
<td>8</td>
<td>Flyers</td>
<td>Eleni Efthimiou</td>
<td></td>
<td></td>
<td>All</td>
<td>International</td>
</tr>
<tr>
<td>9</td>
<td>Video</td>
<td>UniS</td>
<td></td>
<td>February 2011</td>
<td>UK</td>
<td>All</td>
</tr>
<tr>
<td>10</td>
<td>Articles in newspaper Der Spiegel</td>
<td>UHH</td>
<td></td>
<td>2011</td>
<td>Germany</td>
<td>All</td>
</tr>
<tr>
<td>11</td>
<td>Presentation</td>
<td>François Lefebvre-Albaret</td>
<td>February 2011</td>
<td>France</td>
<td>Industry</td>
<td>France</td>
</tr>
<tr>
<td>12</td>
<td>Presentation</td>
<td>Jérémie Segouat</td>
<td>March 2011</td>
<td>France</td>
<td>Industry (innovation)</td>
<td>France</td>
</tr>
<tr>
<td>13</td>
<td>Presentation</td>
<td>François Goudenove</td>
<td>May 2011</td>
<td>Belgium</td>
<td>Policy makers</td>
<td>Europe</td>
</tr>
<tr>
<td>14</td>
<td>Presentation</td>
<td>Elise Leroy</td>
<td>June 2011</td>
<td>France</td>
<td>Industry (social entrepreneurship)</td>
<td>France</td>
</tr>
<tr>
<td>15</td>
<td>Presentation</td>
<td>François Lefebvre-Albaret</td>
<td>January 2011</td>
<td>France</td>
<td>All</td>
<td>France</td>
</tr>
<tr>
<td>24</td>
<td>Presentation</td>
<td>Jérémie Segouat</td>
<td>May 2011</td>
<td>Spain</td>
<td>Industry, scientific community, Deaf community</td>
<td>Europe</td>
</tr>
<tr>
<td>25</td>
<td>Presentation</td>
<td>François Lefebvre-Albaret</td>
<td>June 2011</td>
<td>France</td>
<td>Deaf community</td>
<td>France</td>
</tr>
<tr>
<td>26</td>
<td>Presentation</td>
<td>François Lefebvre-Albaret</td>
<td>July 2011</td>
<td>France</td>
<td>Deaf community</td>
<td>Europe</td>
</tr>
<tr>
<td>27</td>
<td>Presentation</td>
<td>Julia Pelhate</td>
<td>September 2011</td>
<td>France</td>
<td>Deaf community</td>
<td>France</td>
</tr>
<tr>
<td>28</td>
<td>Presentation</td>
<td>Julia Pelhate</td>
<td>September 2011</td>
<td>Italy</td>
<td>Translators and interpreters (deaf and hearing) community</td>
<td>International</td>
</tr>
<tr>
<td>29</td>
<td>Exhibition (booth)</td>
<td>Jérémie Segouat</td>
<td>May 2010</td>
<td>Malta</td>
<td>Scientific community</td>
<td>International</td>
</tr>
<tr>
<td>30</td>
<td>Article</td>
<td>John Glauert</td>
<td>March 2012</td>
<td>England</td>
<td>Educators of the deaf</td>
<td>UK</td>
</tr>
</tbody>
</table>

* TEMPLATE A2 provides information on dissemination activities organised by the Dicta-Sign consortium, both as regards dissemination to the scientific community, the Deaf community and the wide public.

Information on the overall dissemination activity of the consortium members towards the scientific community is provided in TEMPLATE A1.
4.3 Report on societal implications

A General Information (completed automatically when Grant Agreement number is entered.)

<table>
<thead>
<tr>
<th>Grant Agreement Number:</th>
<th>231135</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of Project:</td>
<td>Sign Language Recognition, Generation and Modelling application in Deaf Communication</td>
</tr>
<tr>
<td>Name and Title of Coordinator:</td>
<td>Dr. Eleni Efthimiou</td>
</tr>
</tbody>
</table>

B Ethics

1. Did your project undergo an Ethics Review (and/or Screening)? If Yes: have you described the progress of compliance with the relevant Ethics Review/Screening Requirements in the frame of the periodic/final project reports? Special Reminder: the progress of compliance with the Ethics Review/Screening Requirements should be described in the Period/Final Project Reports under the Section 3.2.2 'Work Progress and Achievements'

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>X</td>
</tr>
</tbody>
</table>

2. Please indicate whether your project involved any of the following issues (tick box):

- Research on Humans
  - Did the project involve children? [ ]
  - Did the project involve patients? [ ]
  - Did the project involve persons not able to give consent? [ ]
  - Did the project involve adult healthy volunteers? [x]
  - Did the project involve Human genetic material? [ ]
  - Did the project involve Human biological samples? [ ]
  - Did the project involve Human data collection? [ ]

Research on Human embryo/foetus
- Did the project involve Human Embryos? [ ]
- Did the project involve Human Foetal Tissue / Cells? [ ]
- Did the project involve Human Embryonic Stem Cells (hESCs)? [ ]
- Did the project on human Embryonic Stem Cells involve cells in culture? [ ]
- Did the project on human Embryonic Stem Cells involve the derivation of cells from Embryos? [ ]

Privacy
- Did the project involve processing of genetic information or personal data (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)? [ ]
- Did the project involve tracking the location or observation of people? [ ]

Research on Animals
- Did the project involve research on animals? [ ]
Were those animals transgenic small laboratory animals?
Were those animals transgenic farm animals?
Were those animals cloned farm animals?
Were those animals non-human primates?
Research Involving Developing Countries
Did the project involve the use of local resources (genetic, animal, plant etc)?
Was the project of benefit to local community (capacity building, access to healthcare, education etc)?

Dual Use
Research having direct military use 0 Yes x No
Research having the potential for terrorist abuse 0 Yes x No

C Workforce Statistics

3. Workforce statistics for the project: Please indicate in the table below the number of people who worked on the project (on a headcount basis).

<table>
<thead>
<tr>
<th>Type of Position</th>
<th>Number of Women</th>
<th>Number of Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientific Coordinator</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Work package leaders</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Experienced researchers (i.e. PhD holders)</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>PhD Students</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>25</td>
<td>22</td>
</tr>
</tbody>
</table>

4. How many additional researchers (in companies and universities) were recruited specifically for this project?

Of which, indicate the number of men:
For ILSP among the people reported in 3, 3 male and 1 female researchers were recruited for the project.
For UHH among the people reported in 3, 3 female and 1 male researchers were recruited for the project.
For CNRS among the people reported in 3, 2 female post-graduate engineers and 1 male researcher were recruited for the project.
For UniS among the people reported in 3, 1 female and 1.5 male researchers were recruited for the project.
For NTUA among the people reported in 3, 3 female and 8 male researchers were recruited for the project.
For UEA among the people reported in 3, 4 female and 6 male researchers were recruited for the project.
## Gender Aspects

5. Did you carry out specific Gender Equality Actions under the project?  
   - Yes  
   - No

6. Which of the following actions did you carry out and how effective were they?  
   - Design and implement an equal opportunity policy  
     - Not at all effective  
     - Very effective  
   - Set targets to achieve a gender balance in the workforce  
   - Organise conferences and workshops on gender  
   - Actions to improve work-life balance  
   - Other:

7. Was there a gender dimension associated with the research content – i.e. wherever people were the focus of the research as, for example, consumers, users, patients or in trials, was the issue of gender considered and addressed?  
   - Yes - please specify  
     - Regarding SL native signers it was designed to treat male and female informants equally for the corpus collection process  
   - No

## Synergies with Science Education

8. Did your project involve working with students and/or school pupils (e.g. open days, participation in science festivals and events, prizes/competitions or joint projects)?  
   - Yes - please specify  
   - No

9. Did the project generate any science education material (e.g. kits, websites, explanatory booklets, DVDs)?  
   - Yes - please specify  
     - Multilingual SL lexicon, SL corpora  
   - No

## Interdisciplinarity

Which disciplines (see list below) are involved in your project?  

- Main discipline\(^7\): 1.1, 6.2  
- Associated discipline\(^7\):  

---

\(^7\) Insert number from list below (Frascati Manual).
### G Engaging with Civil society and policy makers

**11a**  Did your project engage with societal actors beyond the research community? *(if ‘No’, go to Question 14)*

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

**11b**  If yes, did you engage with citizens (citizens' panels / juries) or organised civil society (NGOs, patients' groups etc.)?

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes- in determining what research should be performed</th>
<th>Yes - in implementing the research</th>
<th>Yes, in communicating /disseminating / using the results of the project</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**11c**  In doing so, did your project involve actors whose role is mainly to organise the dialogue with citizens and organised civil society (e.g. professional mediator; communication company, science museums)?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**12.**  Did you engage with government / public bodies or policy makers (including international organisations)

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes- in framing the research agenda</th>
<th>Yes - in implementing the research agenda</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**13a**  Will the project generate outputs (expertise or scientific advice) which could be used by policy makers?

<table>
<thead>
<tr>
<th></th>
<th>Yes – as a primary objective (please indicate areas below- multiple answers possible): Research and Innovation, Information Society</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes – as a secondary objective (please indicate areas below - multiple answer possible): Education, Training, Youth</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

**13b**  If Yes, in which fields?

|-------------|----------------------|--------|-------------|-----------|---------|--------|------------------------------------------|--------------------------|-------------------------------|--------|--------------|-------------|-------------|------------------------|---------------|--------------------------|-------------|------------------------|---------|------------------|----------------|----------------|--------------------------|------|---------|---------|

### 13c If Yes, at which level?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>Local / regional levels</td>
</tr>
<tr>
<td>x</td>
<td>National level</td>
</tr>
<tr>
<td>x</td>
<td>European level</td>
</tr>
<tr>
<td></td>
<td>International level</td>
</tr>
</tbody>
</table>

### H Use and dissemination

#### 14. How many Articles were published/accepted for publication in peer-reviewed journals?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>21:</td>
<td>10 in peer-reviewed journals and 11 in peer-reviewed multi-author book chapters and contributions to edited volumes</td>
</tr>
<tr>
<td>Moreover the following scientific output is also achieved</td>
<td></td>
</tr>
<tr>
<td>6:</td>
<td>Workshop proceedings, special session and peer-reviewed multi-author book editings</td>
</tr>
<tr>
<td>78:</td>
<td>Articles in refereed Conferences and Workshops</td>
</tr>
<tr>
<td>4:</td>
<td>Popular Science</td>
</tr>
</tbody>
</table>

#### To how many of these is open access\(^8\) provided?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12 out of the 20 journals and multi-author book chapters and contributions to edited volumes</td>
<td></td>
</tr>
<tr>
<td>81 out of the 84 workshop proceedings editings, refereed conferences and workshops &amp; popular science</td>
<td></td>
</tr>
</tbody>
</table>

---

\(^8\) Open Access is defined as free of charge access for anyone via Internet.
### How many of these are published in open access journals?

<table>
<thead>
<tr>
<th>2 peer-reviewed journals</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 Peer-reviewed multi-author book chapters and contributions to edited volumes</td>
</tr>
<tr>
<td>Sprinkellink Sense provides free academic access to all Springer publications</td>
</tr>
</tbody>
</table>

### How many of these are published in open repositories?

| 77 workshop/conference papers can be accessed freely from academic domains and relevant event sites. |

### To how many of these is open access not provided?

| 3 |

#### Please check all applicable reasons for not providing open access:

- x publisher's licensing agreement would not permit publishing in a repository
- x no suitable repository available
- q no suitable open access journal available
- x no funds available to publish in an open access journal
- x lack of time and resources
- q lack of information on open access
- q other*: ..............

### 15. How many new patent applications (‘priority filings’) have been made? (*Technologically unique*: multiple applications for the same invention in different jurisdictions should be counted as just one application of grant).

| 0 |

### 16. Indicate how many of the following Intellectual Property Rights were applied for (give number in each box).

| Trademark | 0 |
| Registered design | 0 |
| Other | 0 |

### 17. How many spin-off companies were created / are planned as a direct result of the project?

Indicate the approximate number of additional jobs in these companies: 0

---

* For instance: classification for security project.
18. Please indicate whether your project has a potential impact on employment, in comparison with the situation before your project:

- □ Increase in employment, or □ In small & medium-sized enterprises
- □ Safeguard employment, or □ In large companies
- □ Decrease in employment, x None of the above / not relevant to the project
- □ Difficult to estimate / not possible to quantify

19. For your project partnership please estimate the employment effect resulting directly from your participation in Full Time Equivalent (FTE = one person working fulltime for a year) jobs:

Indicate figure:

- Difficult to estimate / not possible to quantify

I Media and Communication to the general public

20. As part of the project, were any of the beneficiaries professionals in communication or media relations?

Please fill in accordingly

- O Yes x No

21. As part of the project, have any beneficiaries received professional media / communication training / advice to improve communication with the general public?

- O Yes x No

22. Which of the following have been used to communicate information about your project to the general public, or have resulted from your project?

- □ Press Release
- □ Media briefing
- □ TV coverage / report x TV coverage / report
- □ Radio coverage / report
- □ Brochures / posters / flyers x Brochures / posters / flyers
- □ DVD / Film / Multimedia
- □ Coverage in specialist press
- □ Coverage in general (non-specialist) press x Coverage in general (non-specialist) press
- □ Coverage in national press x Coverage in national press
- □ Coverage in international press x Coverage in international press
- □ Website for the general public / internet x Website for the general public / internet
- □ Event targeting general public (festival, conference, exhibition, science café) x Event targeting general public (festival, conference, exhibition, science café)

23. In which languages are the information products for the general public produced?

- □ Language of the coordinator x English
- □ Other language(s), French, Greek, German for project results and the local SLs, namely BSL, LSF, GSL, DGS
**Question F-10:** Classification of Scientific Disciplines according to the Frascati Manual 2002 (Proposed Standard Practice for Surveys on Research and Experimental Development, OECD 2002):

**FIELDS OF SCIENCE AND TECHNOLOGY**

1. **NATURAL SCIENCES**
   1.1 Mathematics and computer sciences [mathematics and other allied fields: computer sciences and other allied subjects (software development only; hardware development should be classified in the engineering fields)]
   1.2 Physical sciences (astronomy and space sciences, physics and other allied subjects)
   1.3 Chemical sciences (chemistry, other allied subjects)
   1.4 Earth and related environmental sciences (geology, geophysics, mineralogy, physical geography and other geosciences, meteorology and other atmospheric sciences including climatic research, oceanography, vulcanology, palaeocology, other allied sciences)
   1.5 Biological sciences (biology, botany, bacteriology, microbiology, zoology, entomology, genetics, biochemistry, biophysics, other allied sciences, excluding clinical and veterinary sciences)

2. **ENGINEERING AND TECHNOLOGY**
   2.1 Civil engineering (architecture engineering, building science and engineering, construction engineering, municipal and structural engineering and other allied subjects)
   2.2 Electrical engineering, electronics [electrical engineering, electronics, communication engineering and systems, computer engineering (hardware only) and other allied subjects]
   2.3. Other engineering sciences (such as chemical, aeronautical and space, mechanical, metallurgical and materials engineering, and their specialised subdivisions; forest products; applied sciences such as geodesy, industrial chemistry, etc.; the science and technology of food production; specialised technologies of interdisciplinary fields, e.g. systems analysis, metallurgy, mining, textile technology and other applied subjects)

3. **MEDICAL SCIENCES**
   3.1 Basic medicine (anatomy, cytology, physiology, genetics, pharmacy, pharmacology, toxicology, immunology and immunohaematology, clinical chemistry, clinical microbiology, pathology)
   3.2 Clinical medicine (anaesthesiology, paediatrics, obstetrics and gynaecology, internal medicine, surgery, dentistry, neurology, psychiatry, radiology, therapeutics, otorhinolaryngology, ophthalmology)
   3.3 Health sciences (public health services, social medicine, hygiene, nursing, epidemiology)

4. **AGRICULTURAL SCIENCES**
   4.1 Agriculture, forestry, fisheries and allied sciences (agronomy, animal husbandry, fisheries, forestry, horticulture, other allied subjects)
   4.2 Veterinary medicine

5. **SOCIAL SCIENCES**
   5.1 Psychology
   5.2 Economics
   5.3 Educational sciences (education and training and other allied subjects)
   5.4 Other social sciences [anthropology (social and cultural) and ethnology, demography, geography (human, economic and social), town and country planning, management, law, linguistics, political sciences, sociology, organisation and methods, miscellaneous social sciences and interdisciplinary , methodological and historical S1T activities relating to subjects in this group. Physical anthropology, physical geography and psychophysiology should normally be classified with the natural sciences].

6. **HUMANITIES**
   6.1 History (history, prehistory and history, together with auxiliary historical disciplines such as archaeology, numismatics, palaeography, genealogy, etc.)
   6.2 Languages and literature (ancient and modern)
   6.3 Other humanities [philosophy (including the history of science and technology) arts, history of art, art criticism, painting, sculpture, musicology, dramatic art excluding artistic "research" of any kind, religion, theology, other fields and subjects pertaining to the humanities, methodological, historical and other S1T activities relating to the subjects in this group]