



R A P P O R T A N N U E L 2 0 1 2



Impressum

Réalisation: Céline Aymon Fournier, Relations publiques, Idiap

Rédaction: Le fin mot... Communication, Martigny

Traduction: Dave Brooks, ELCS

Conception graphique: Atelier Grand, Sierre

Crédits photographiques: Céline Ribordy, Sion; Sedrik Nemeth, Sion; Idiap, Martigny

Impression: calligraphy.ch, Sierre

Tirage: 1800 exemplaires

SOMMAIRE

Messages

«L'Idiap contribue au renforcement de la place économique valaisanne.» Olivier Dumas, président du Conseil de fondation de l'Idiap	2
«Nous sommes plus que jamais en lien avec les préoccupations actuelles de notre société.» Prof. Hervé Bourlard, directeur de l'Idiap	3

Recherche

Idiap, carte d'identité	5
Activités de recherche - Sélection	
Mobile Data Challenge: Le smartphone au service de la qualité de la vie	9
Mediaparl: Le Grand Conseil comme si vous y étiez!	10
Projet armasuisse: Pour communiquer en toute sécurité	12
Projet Ninapro: Nouvelle génération de main artificielle	13

Réseau

Partenaires industriels: La confiance des plus grands	15
International Create Challenge: De l'idée à la start-up en trois semaines	16
L'actualité des start-up: Croissance et nouveaux engagements	18

Visages

Lakshmi Saheer, première docteur indienne de l'Idiap: «Mes collègues ici sont épatants.»	21
Equipe gestion de projets: Les managers de la recherche	22
Antoine Dorsaz, aide-comptable: Les fruits de la passion	24
Allées et venues	25
Distinctions	26
Thèses achevées	27

Finances

Compte d'exploitation	29
Sources de financement / Charges / Commentaires	30
Bilan	31

Organisation

Organigramme opérationnel	33
Collaborateurs	34
Conseil de fondation	36
Comité d'accompagnement	38
Principaux partenaires	39

Encarté scientifique

Idiap Research Areas: Human and Media Computing	II
Scientific Progress Report	IV
Main projects in progress	XVIII
Major publications / Conferences	XXVI

BILLET DU PRÉSIDENT

«L'IDIAP CONTRIBUE AU RENFORCEMENT DE LA PLACE ÉCONOMIQUE VALAISANNE.»



Olivier Dumas, président du Conseil de fondation de l'Idiap

Vingt ans, et après... On aurait pu croire, après les festivités du 20^e anniversaire de l'Idiap, en 2011, et la satisfaction d'avoir vu le petit institut devenir grand en conservant sa créativité et sa convivialité, qu'un léger essoufflement se manifesterait. Il n'en a rien été.

Le projet Mediaparl, mené en partenariat avec la télévision régionale Canal 9 et le Service parlementaire valaisan, est l'un des aboutissements majeurs de cette année 2012. Ce projet a permis de mettre au service du canton les compétences de niveau mondial de l'Idiap en matière de reconnaissance automatique de la parole. Grâce à un projet de recherche de deux ans, soutenu par la Loterie Romande et le Grand Conseil, l'institut a pu décliner ce savoir-faire en français et en allemand, et offrir ainsi au Parlement valaisan un outil d'archivage et d'indexation des vidéos de ses sessions. Une plateforme que plusieurs autres parlements cantonaux et même la Berne fédérale souhaitent aujourd'hui adopter. (Voir pages 10-11).

En développant toujours davantage sa mission de transfert de technologies, notamment via sa collaboration avec le site IdeArk de la Fondation The Ark, l'Idiap contribue également au renforcement de la place économique valaisanne. Preuve en est la progression de ses principales start-up qui ont toutes, en 2012, engagé un nouveau collaborateur, confirmant ainsi leur dynamisme que plusieurs prix et récompenses sont venus couronner. (Voir pages 18-19).

De plus, relevons le renouvellement du Conseil de fondation de l'Idiap. Je saisi ici l'occasion de remercier pour leur précieuse collaboration les membres sortants et de souhaiter la bienvenue aux trois arrivants, MM. Berclaz, Furrer et Montserrat. Respectivement représentants de la HES-SO Valais Wallis et du futur pôle EPFL Valais Wallis, d'Euresearch et de venturelab/venture kick. Bienvenue également aux deux nouveaux représentants de l'EPFL, Mme Adrienne Corboud Fumagalli et le professeur Philippe Gillet, respectivement vice-présidente pour l'innovation et la valorisation, et vice-président pour les affaires académiques. Gageons qu'ils offriront à l'institut de nouvelles perspectives et que leur vision nous permettra, demain, d'intensifier encore notre collaboration avec les autres acteurs de la scène scientifique régionale, nationale et internationale. (Voir pages 36-37).

Finalement, l'année 2012 a également été marquée par l'annonce de la signature en janvier de l'accord entre le Conseil d'Etat et l'EPFL pour l'implantation en Valais d'une antenne active dans les domaines de l'énergie, de la chimie et de la santé, par la création de onze chaires de recherche ainsi que de programmes d'enseignement de niveau master. Un projet qui devrait se réaliser en 2015 et dans lequel l'Idiap devra jouer sa partition, forte de sa collaboration historique avec l'EPFL. A n'en pas douter, «EPFL Valais Wallis» permettra, en établissant des liens étroits avec les institutions de recherche et la Fondation The Ark, d'élargir l'offre de formation en Valais, de renforcer les activités de recherche mais aussi de proposer aux Valaisans et Valaisannes de nouveaux débouchés dans leur canton.

MESSAGE DU DIRECTEUR

«NOUS SOMMES PLUS QUE JAMAIS EN LIEN AVEC LES PRÉOCCUPATIONS ACTUELLES DE NOTRE SOCIÉTÉ.»



Prof. Hervé Boulard, directeur de l'Idiap

Depuis ses débuts, en 1991, l'Idiap a toujours été guidé par l'esprit de son fondateur, l'industriel italien Dalle Molle, celui de mettre la machine au service de l'humain. L'idée peut paraître désuète. Et pourtant. A une époque où l'informatique se glisse dans chaque parcelle de notre quotidien, l'ambition prend tout son sens. Et l'Idiap, malgré une croissance extraordinaire, un déménagement dans les spacieux locaux du Centre du Parc et un renforcement de sa place sur la scène mondiale, n'a pas oublié. Bien au contraire.

En 2012, notre institut a choisi d'ancrer davantage encore ses recherches dans la réalité actuelle en se diversifiant et en valorisant ses compétences clés dans des domaines d'application en lien avec les préoccupations de notre société.

Ainsi, pas moins de dix domaines d'application bénéficient désormais du savoir-faire de haut niveau de nos scientifiques (voir page 7). Energie, santé, sécurité, interaction homme/homme et homme/machine sont quelques-uns de ces domaines dans lesquels l'Idiap mène des projets de recherche nationaux et européens avec dynamisme, affirmant ainsi avec davantage de clarté et de visibilité encore le rôle moteur qu'il entend jouer dans le progrès de notre société. Un vœu qui semble d'ailleurs se réaliser, le magazine *Bilan* ayant désigné en juin 2012 deux de nos scientifiques parmi les 300 personnes les plus influentes de notre pays. Sébastien Marcel, l'un de nos chercheurs seniors, y est reconnu pour ses activités internationales en biométrie, imposant l'Idiap en Europe comme un acteur clé dans ce domaine, tandis que j'y suis pour ma part mentionné pour avoir fait de l'Idiap «l'un des principaux acteurs mondiaux des interfaces homme-machine avec des start-up prometteuses comme Koemei et Klewel» (voir pages 18-19).

L'Idiap cherche et trouve, c'est une réalité, et celle-ci me réjouit. Ce bilan positif, je le partage avec les trois maîtres d'enseignement et de recherche (MER) qui assurent avec moi la Direction scientifique de l'Idiap et avec notre nouveau directeur adjoint, François Foglia. Entrepreneurs et autonomes, ceux-ci réalisent un travail formidable. Les chiffres d'ailleurs parlent d'eux-mêmes: avec la gestion de 42 projets en 2012, notre institut a établi un nouveau record en la matière!

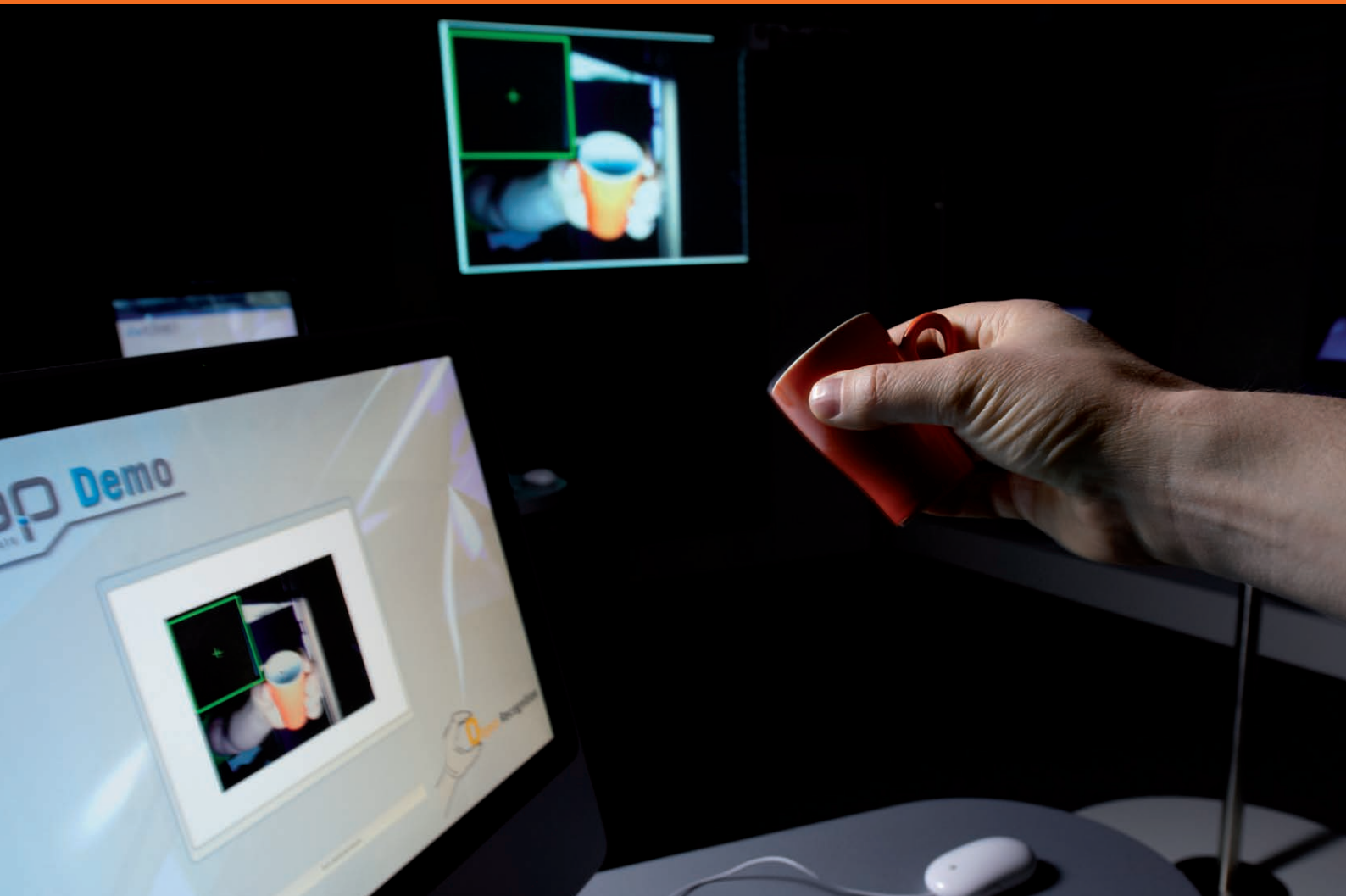
Cette mission de recherche méritait d'être encore mieux soutenue, et c'est désormais chose faite grâce à l'engagement d'une nouvelle équipe de gestion de projets (voir pages 22-23). Trois nouveaux collaborateurs accompagnent les chercheurs dans des tâches administratives qui se révèlent de plus en plus lourdes, mais qui revêtent toute leur importance.

Je profite aussi de ce message pour souhaiter la bienvenue, et plein succès, à l'initiative «EPFL Valais Wallis». Ceci est une nouvelle opportunité pour tous les partenaires, y compris l'Idiap, associé à l'EPFL depuis sa création, de se développer encore davantage. Notre institut aura certainement un rôle clé à jouer, ce qui sera sans doute à notre agenda 2013.

Enfin, je tiens ici à féliciter tous les collaborateurs de l'Idiap – ils étaient plus de 100 en 2012 – pour leur engagement. Chercheurs, doctorants, développeurs, administratifs, tous ont œuvré à notre succès. Je leur souhaite, je nous souhaite, de vivre une année 2013 sous le signe de l'audace et de la réussite, et cela dans l'esprit de collaboration et de haut standard de qualité qui a toujours animé l'Idiap.

A handwritten signature in blue ink, which appears to read 'H. Boulard'. The signature is stylized and fluid.

RECHERCHE



UN INSTITUT À TAILLE HUMAINE ET AU RAYONNEMENT INTERNATIONAL

Fort d'une centaine de collaborateurs et de domaines de recherche en lien avec les défis actuels, l'Idiap s'engage pour un progrès scientifique au service de l'homme.

En 1991 déjà, date de sa fondation, l'institut de recherche Idiap se donnait pour mission de s'engager pour un progrès scientifique au service du bien-être des hommes. Aujourd'hui, vingt ans plus tard, l'institut place toujours les intérêts de la société au cœur de sa démarche.

Déferlement de produits technologiques

En ce début de XXI^e siècle, on assiste à un déferlement permanent de nouveaux outils technologiques. S'ils permettent des gains considérables en termes d'efficacité et de confort, ils bouleversent aussi les habitudes des gens, laissant une part des utilisateurs démunis et une autre lassée par la modification récurrente des systèmes. Dans ce contexte, l'Idiap travaille essentiellement à l'amélioration des relations homme-machine, et à l'optimisation de la communication humaine.

Réseau national et international

Au niveau suisse, l'Idiap travaille avec les écoles polytechniques, les hautes écoles et les universités où se pratiquent également des activités de recherche. Impliqué dans plusieurs projets européens, l'institut entretient des liens étroits avec de nombreux partenaires, essentiellement en France, en Angleterre et en Allemagne. Outre-Atlantique, c'est la Californie avec l'International Computer Science Institute (ICSI) de Berkeley qui se positionne comme le partenaire privilégié.

Portrait express

Structure / Fondation à but non lucratif, l'Institut de recherche Idiap a été fondé en 1991 par la Ville de Martigny, l'Etat du Valais, l'Ecole polytechnique fédérale de Lausanne (EPFL), l'Université de Genève et Swisscom. Autonome, l'institut est cependant lié à l'EPFL par un plan de développement commun.

Financement / Le budget de l'Idiap – 10 millions de francs suisses – est assuré à 60% par des projets de recherche décrochés au terme de processus compétitifs, et à 40% par des fonds publics (voir répartition des sources de financement, page 30).

Equipe / En 2012 l'Idiap emploie plus d'une centaine de collaborateurs, dont quelque 80 chercheurs (professeurs, chercheurs seniors, chercheurs, postdoctorants et doctorants).

Locaux / L'Idiap est installé depuis 2007 dans l'aile ouest du Centre du Parc de Martigny. Il y occupe 2600 m² de locaux, répartis sur quatre étages.

Missions

■ Recherche

Mener des projets de recherche fondamentale au plus haut niveau dans ses domaines de prédilection, s'assurant ainsi une place parmi les meilleurs à l'échelle nationale, européenne et mondiale. Sur la scène internationale, l'Idiap bénéficie d'un large réseau de partenaires et collabore activement avec de grandes universités, des centres de recherche publics ou privés, etc.
Les projets de recherche compétitifs garantissent 60% du financement de l'institut.

■ Formation

Former la relève en faisant découvrir le monde de la recherche à des stagiaires, en accueillant de jeunes chercheurs talentueux préparant leur doctorat, et en dispensant de nombreux cours à l'EPFL et en interne.
Un collaborateur scientifique sur deux est un doctorant.

■ Transfert de technologies

Assurer le transfert des technologies à travers la dissémination la plus large possible de ses résultats de recherche dans la communauté scientifique, mais aussi et surtout en tissant des liens étroits avec le monde industriel.
Grâce à l'incubateur The Ark sur le site d'IdeArk, l'Idiap permet l'éclosion de nombreuses start-up.



5 THÈMES DE RECHERCHE & 10 DOMAINES D'APPLICATION

Les travaux menés à l'Idiap sont regroupés dans cinq thèmes de recherche et valorisés dans dix domaines d'application.

5 thèmes de recherche

Depuis sa fondation, en 1991, l'Idiap a développé l'essentiel de ses recherches autour de la communication homme-machine et homme-homme facilitée par la machine. Une thématique qui se divise en cinq thèmes de recherche:

1 Les systèmes perceptifs et cognitifs

(reconnaissance automatique de la parole, vision par ordinateur, reconnaissance de l'écriture, traitement de documents multimédias, robotique, traduction automatique)

2 La modélisation du comportement humain et social

(médias de communication web et mobiles, interactions sociales, traitement des signaux sociaux)

3 Les interfaces d'information et de présentation

(systèmes d'information multimédias, contextualisation et personnalisation, etc.)

4 L'authentification biométrique

(vérification du locuteur, vérification du visage, etc.)

5 L'apprentissage automatique

(modélisation statistique, réseaux de neurones, modèles mathématiques)



L'Idiap en chiffres (année 2012)

Ressources humaines

- 1 professeur
- 3 MER
- 9 chercheurs permanents et seniors
- 18 postdoctorants
- 34 doctorants
- 16 ingénieurs système et ingénieurs de développement
- 11 stagiaires et visiteurs (moyenne/année)
- 10 collaborateurs administratifs
- 7 titres de docteur décernés
- 42 postes dans les start-up du site IdeArk
- 31 nationalités représentées

Activités scientifiques

- Pôle national de recherche IM2 (Gestion interactive et multimodale de systèmes d'information) depuis 2001
- Participation à 42 programmes de recherche
- Direction de projet dans 6 consortiums
- Participation à la stratégie de développement économique du Canton du Valais à travers le programme The Ark et en particulier la société IdeArk
- 240 publications scientifiques
- Participation à de nombreuses conférences internationales

10 domaines d'application

Que ce soit au travers des projets européens ou nationaux dans lesquels il est engagé ou des partenariats industriels qu'il entretient, l'Idiap valorise ses compétences clés dans des domaines d'application qui, aujourd'hui, sont en lien étroit avec les préoccupations de notre société, à l'image de l'énergie, de la sécurité, des systèmes mobiles ou encore de l'exploitation des archives multimédias.

1

Commande d'appareils par la voix, systèmes de traduction voix/voix, systèmes de navigation, etc.

**INTERACTION HOMME-HOMME
ET HOMME-MACHINE/ROBOT**

Indexation sémantique, reconnaissance d'objets, résumé de contenu audiovisuel, analyse de documents YouTube, analyse de l'héritage culturel, etc.

2

**EXPLOITATION D'ARCHIVES
MULTIMÉDIAS RICHES**

3

Salle de réunion intelligente, vidéo-conférence, accès et indexation multimédias, collaboration inter-langues, etc.

**SYSTÈMES COLLABORATIFS
ET CRÉATIFS**

Traitement du signal pour plateformes mobiles, réseaux sociaux mobiles, recommandation et publicité sur mobile, etc.

4

**GESTION DES
SYSTÈMES MOBILES**

5

Réseaux énergétiques hétérogènes, réseaux de senseurs, modélisation de l'activité humaine pour anticiper les besoins, etc.

GESTION DE L'ÉNERGIE

Gestion intelligente des patients, prothèses (auditives, corporelles, etc.), modélisation de systèmes biologiques, interfaces pour les handicapés, acquisition et indexation de documents médicaux

6

SANTÉ ET BIO-ENGINEERING

7

Contrôles d'accès (espaces physiques ou virtuels), vérification du locuteur, vérification du visage, vidéosurveillance, modélisation des risques naturels, sécurité des espaces publics et privés

**SÉCURITÉ ET GESTION
DES RISQUES**

Ecologie, management environnemental, réduction de la pollution, réduction du trafic, meilleur usage du réseau routier, réduction du bruit

10

**VILLES DU FUTUR
(«SMART CITIES»)**

9

Jeux en réseau multilingues, jeux familiaux à distance («togetherness»)

DIVERTISSEMENT ET JEUX

Analyse de scènes multi-senseurs, analyse du comportement humain, adaptation au comportement humain (chauffage, etc.), sécurité ménagère

10

DOMOTIQUE



GROUPES DE RECHERCHE

A l'IIdiap, les recherches sont menées au sein de huit groupes de recherche.

Speech & Audio Processing

Prof. Hervé Bourlard, Dr John Dines, Dr Phil Garner, Dr Mathew Magimai Doss, Dr Petr Motlicek

Le traitement de la parole est depuis de nombreuses années l'un des thèmes de recherche majeurs de l'IIdiap. L'institut occupe d'ailleurs une place de choix sur la scène internationale dans ce domaine. Au cœur des travaux du groupe de recherche le plus important de l'IIdiap: la reconnaissance automatique de la parole de manière statistique, la synthèse vocale et le traitement générique de l'information audio (localisation de la source, des réseaux de microphones, segmentation par locuteur, indexation de l'information, codage du signal vocal à de très bas débits, analyse du bruit de fond).

Computer Vision & Learning

Dr François Fleuret

Les chercheurs de ce groupe travaillent sur le développement de nouvelles techniques d'apprentissage statistique pour la vision par ordinateur, avec un intérêt particulier pour leurs propriétés computationnelles. Leurs recherches ont pour principaux domaines d'application la reconnaissance automatique d'images, en particulier l'analyse de scènes, la détection d'objets, le suivi de personnes et de structures biologiques.

Social Computing

Dr Daniel Gatica-Perez

Le «social computing» est un domaine interdisciplinaire qui intègre les théories et les modèles de l'informatique mobile et omniprésente, de l'apprentissage automatique, du multimédia et des sciences sociales. Tous ces domaines sont réunis pour détecter, analyser et interpréter les comportements humains et sociaux au quotidien dans le but de créer des instruments qui soutiennent les interactions et la communication. Les lignes de recherche actuelles comprennent l'étude des interactions face-à-face, l'analyse comportementale de vidéos sociales et le traitement de données urbaines à l'aide de téléphones mobiles et de réseaux sociaux géo-localisés.

Perception & Activity Understanding

Dr Jean-Marc Odobez

Ce groupe de recherche s'intéresse à l'analyse des activités humaines à partir de données multimodales. Cette analyse repose sur le développement d'algorithmes fondés sur des méthodes de vision par ordinateur, d'apprentissage, et de fusion de données pour résoudre des tâches fondamentales comme la détection et le suivi de personnes, leur représentation et la caractérisation de leur état, ainsi que la modélisation de données séquentielles et leur interprétation sous forme de gestes, comportements ou relations sociales. La surveillance, l'analyse de comportements, les interfaces homme-robot et l'analyse de contenus multimédias constituent les principaux domaines d'application.

Artificial Cognitive Systems

Dr Barbara Caputo

Ce groupe travaille sur le développement d'algorithmes d'apprentissage multimodaux qui permettent aux vecteurs artificiels d'agir de manière autonome dans un cadre réaliste. Un accent particulier est mis sur la capacité de détecter de manière indépendante les lacunes de connaissance et de les combler de manière autonome avec des stratégies d'apprentissage ouvertes et flexibles. L'objectif visé est la conception d'algorithmes de principe qui sont informatiquement efficaces et fournissent une performance robuste dans un cadre très réaliste tout en donnant des garanties théoriques sur le comportement attendu.

Applied Machine Learning

Dr Ronan Collobert

Ici, les chercheurs développent des algorithmes informatiques capables d'«apprendre» un comportement pour résoudre une tâche, au contraire des algorithmes plus classiques dont le comportement est régi par un ensemble de règles établies heuristiquement. La recherche est motivée par des applications du monde réel, tel que le langage naturel, la vision et l'audio, impliquant le traitement de grandes bases de données. Un accent particulier est mis sur les algorithmes d'apprentissages génériques qui ne requièrent a priori que peu de connaissance des données.

Biometric Person Recognition

Dr Sébastien Marcel

En informatique, la biométrie se rapporte à la reconnaissance automatique d'individus par l'utilisation de caractéristiques comportementales et biologiques. Les chercheurs étudient et développent de nouveaux algorithmes de traitement d'image et de reconnaissance des formes pour la reconnaissance de visage (2D, 3D et proche infrarouge), la reconnaissance du locuteur, la détection d'attaques (anti-spoofing) ainsi que des modalités biométriques émergentes. Ce groupe encourage la reproduction des résultats de recherche et le transfert de technologies en utilisant sa propre librairie de traitement du signal et d'apprentissage automatique par ordinateur.

Natural Language Processing

Dr Andrei Popescu-Belis

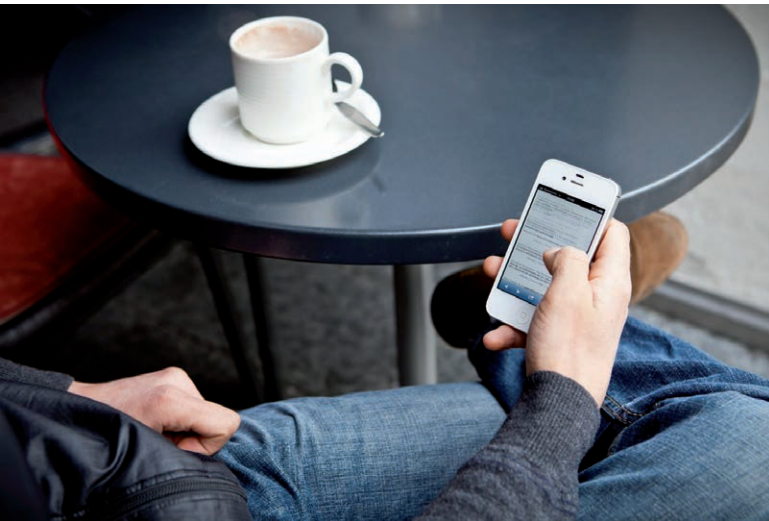
Les chercheurs de ce groupe étudient comment l'analyse des textes au niveau sémantique et pragmatique permet d'améliorer les performances dans deux tâches principales: la traduction automatique et la recherche d'information. Ils s'intéressent également à la façon dont la recherche d'information sur des données en réseau, y compris multimédias, peut être améliorée en utilisant l'information linguistique et celle provenant du réseau.

LE SMARTPHONE AU SERVICE DE LA QUALITÉ DE LA VIE

L'Idiap et le centre de recherche Nokia ont organisé en 2012 le Mobile Data Challenge, un défi lancé à la communauté mondiale des chercheurs. L'enjeu: une meilleure compréhension du comportement humain et social via les téléphones mobiles.

Entre 2009 et 2011, l'Idiap avait mené un projet d'envergure avec Nokia. Intitulé LS-CONTEXT, celui-ci visait en priorité la mise en place d'une base de données d'informations issues des téléphones mobiles, une matière d'autant plus difficile à récolter qu'elle touche à la sphère privée (voir encadré). En 2012, cette aventure a trouvé son épilogue à travers le Mobile Data Challenge (MDC), une initiative de recherche proposée à l'échelle mondiale.

Daniel Gatica-Perez, chercheur senior et responsable du Social Computing Group à l'Idiap, est le co-initiateur de ce vaste projet. «Le téléphone mobile constitue un excellent outil d'étude des comportements sociaux: on l'a en permanence avec soi, on l'utilise pour appeler ou envoyer des SMS mais aussi pour rester connecté à ses réseaux virtuels, s'informer ou encore organiser ses loisirs, son travail.»



108 équipes de chercheurs du monde entier

Ainsi, afin d'offrir à ces informations une analyse à la hauteur de leur richesse, l'Idiap et Nokia Research Lausanne eurent l'idée d'organiser une recherche à large échelle. Début janvier 2012, les données étaient transmises à plus de 500 participants individuels, sous forme de copies marquées et numérotées (watermark). Au printemps 2012, 108 équipes de scientifiques du monde entier, certaines issues d'universités et institutions de haut vol (Carnegie Mellon University, University of Illinois at Urbana-Champaign, CERN, EPFL, ETHZ, TU Berlin), soumettaient leurs résultats. Par un accord formel, tous s'étaient auparavant engagés à traiter les données de manière totalement éthique, uniquement à des fins de recherche. Deux propositions leur ont ensuite été adressées: le Dedicated Track ou l'Open Track. Dans le premier cas, il s'agissait de prédire la mobilité de l'utilisateur, d'estimer son profil démographique ou encore de déterminer le type d'endroit où se trouve l'utilisateur (école, restaurant, parc, cinéma, etc.). Par exemple: des dispositifs bluetooth équidistants et immobiles, un samedi à 20h... on pourrait être dans un cinéma!

«Les données recueillies nous permettront un jour de faciliter la vie des utilisateurs», explique Daniel Gatica-Perez. Prenons l'exemple de Monsieur X qui, après son travail, se rend systématiquement à la gare. Son téléphone pourrait comprendre que l'heure du train approche, et signifier à M. X qu'il est temps de quitter le bureau ou qu'il ne faut pas oublier ses clés.

Le second défi faisait appel à la créativité des scientifiques, libres de proposer leurs propres tâches, d'étudier par exemple les relations entre la mobilité et les conditions météorologiques. Un volet qui a produit plusieurs nouvelles découvertes. Les lauréats ont été invités au Mobile Data Challenge Workshop en juin 2012 à Newcastle au Royaume-Uni.

Une base de données exceptionnelle

Durant 12 à 18 mois, près de 200 volontaires romands ont utilisé un smartphone qui a enregistré leurs déplacements, leur vitesse, la proximité d'autres téléphones, mais aussi des informations relatives à l'utilisation de l'appareil, à la fréquence des appels, à l'utilisation de l'appareil photo, du lecteur MP3, etc.

«Autant de gens concernés sur une aussi longue durée, et autant de types de données récoltées, c'est unique pour la science, explique Daniel Gatica-Perez. Cela nous a permis d'ouvrir des champs de recherche extrêmement vastes.» La dimension éthique n'en a pas pour autant été oubliée. «Nous ne connaissons pas le contenu des appels ou des SMS, mais uniquement dans quelles circonstances l'utilisateur a utilisé ces fonctions.»



LE GRAND CONSEIL COMME SI VOUS Y ÉTIEZ!

A travers le projet Mediaparl, l'Idiap proposera bientôt aux députés valaisans – et au grand public – de retrouver n'importe quelle intervention parlementaire en vidéo. Un outil avant-gardiste qui intéresse la Berne fédérale.

«Impossible de parler du projet Mediaparl sans évoquer celui du parlement sans papier. Les deux sont interconnectés», explique d'emblée Alexandre Nanchen, ingénieur de développement depuis 2008 à l'Idiap et responsable du projet. En 2010, partant du constat, notamment, que chaque député utilisait environ 33 kilos de papier par année, Jean-Albert Ferrez, alors premier vice-président du Grand Conseil et directeur adjoint de l'Idiap, lança le projet d'un parlement équipé de tablettes informatiques. En parallèle se réalisait la numérisation de 213 000 pages produites depuis 1851 par le Grand Conseil valaisan, afin de permettre un accès facilité aux archives.

Un système capable de reconnaître le français et l'allemand

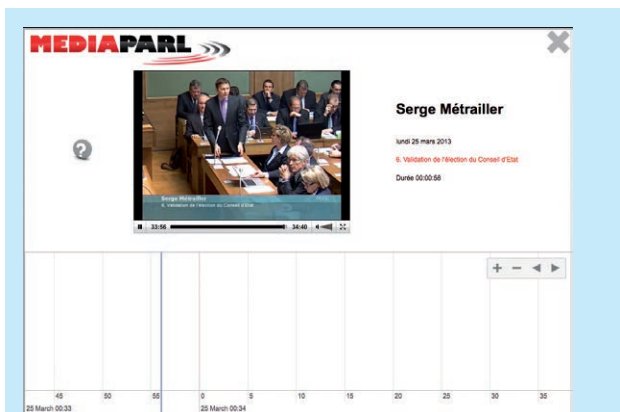
Une fois cette première étape réalisée, s'est naturellement posé la question des archives télévisuelles et du meilleur moyen de les utiliser, Canal 9 enregistrant depuis 2007 toutes les sessions parlementaires. «L'Idiap fait partie des meilleurs instituts de recherche au monde en matière de reconnaissance automatique de la parole, explique Alexandre Nanchen. Nous avons donc eu envie de monter un projet pour indexer toutes ces archives audiovisuelles.» Petit hic: tous les moteurs de reconnaissance automatique développés à ce jour à l'Idiap l'ont été en anglais, et il s'agissait ici de concevoir un outil qui fonctionne avec le français et l'allemand!

Un projet de recherche est alors conçu par Jean-Albert Ferrez et Yann Rodriguez, responsable du transfert de technologie à l'Idiap. Soutenu par la Loterie Romande et le Grand Conseil, Mediaparl a démarré en janvier 2012 et devrait trouver son aboutissement début 2014. Il prévoit le développement parallèle de deux outils; d'une part un outil d'indexation, qui extrait automatiquement tous les mots d'un flux de parole, les indexe, et permet ensuite de retrouver n'importe quelle séquence vidéo par une simple recherche par mot clé, comme on le ferait dans un texte. Le second volet concerne un système de retranscription automatique: «Aujourd'hui le Service parlementaire a besoin de six semaines pour coucher sur papier le contenu d'une session. Avec un tel système, il devrait gagner un temps précieux, même si une relecture et une mise en forme «humaine» seront quand même nécessaires.»

Un échantillon de parole des députés

A la base de chacun des deux volets, un outil informatique capable d'identifier un flux de paroles. «Pour mettre au point un système qui est performant dès le départ, il faut lui fournir des exemples représentatifs de la réalité qu'il devra traiter par la suite. Dans un premier temps, nous avons donc fait correspondre le texte prononcé par les députés avec leur parole grâce aux enregistrements et transcriptions des sessions précédentes du Grand Conseil. Au total vingt heures pour chaque langue.» Un exercice qui a permis de «calibrer» le système, et une étape d'autant plus nécessaire que le Grand Conseil valaisan a ses particularités! «D'une région à l'autre, les accents sont très différents, ce qui ne facilite pas la reconnaissance par l'ordinateur.» Sans compter qu'au Grand Conseil valaisan, 10% des interventions sont faites en allemand, et que rien ne prévient le système du changement de langue! «Ensuite, des notions de construction de phrase et de phonétique ont été intégrées afin que le système puisse apprendre le français et l'allemand.» Un aspect particulièrement important, la syntaxe et le vocabulaire politique étant très particuliers.

A quelques mois de la mise en route du prototype, Alexandre Nanchen tient à rappeler que ce projet a véritablement été un travail d'équipe. «Phil Garner, chercheur senior à l'Idiap, David Imseng, doctorant, mais surtout les ingénieurs de développement ont mis au point les systèmes de reconnaissance vocale français, allemand et les outils d'indexation. Même des étudiants ont travaillé durant l'été sur ce projet!» Enfin, Mediaparl est un projet mené en partenariat avec le Service parlementaire, en particulier avec Daniel Petitjean, son responsable informatique, et avec Canal 9, télévision régionale.



Mediaparl permettra aux députés mais aussi au grand public, grâce à un formulaire de recherche simple, de retrouver n'importe quelle intervention parlementaire depuis 2009. Un lecteur audio-vidéo permet de naviguer dans la séquence en question. En matière de reconnaissance automatique de la parole, l'Idiap se positionne comme l'un des meilleurs instituts au monde.

Le Service parlementaire, qui emploie une dizaine de personnes, apporte un appui juridique et scientifique aux acteurs du Grand Conseil et assume les travaux d'ordre administratif. Mediaparl leur permettra d'améliorer encore leurs services.

Interview de Claude Bumann, chef du Service parlementaire.

Pourquoi avez-vous eu envie de participer au projet Mediaparl?

Le parlement valaisan a toujours eu un rôle de précurseur vis-à-vis des autres parlements cantonaux. Nous avons été les premiers à être retransmis à la télévision, notre site internet se veut un outil très performant et notre équipe cherche en permanence à améliorer les services offerts aux parlementaires. Mediaparl entre dans ce cadre.

Tout est parti de l'idée d'un parlement sans papier, née sous la présidence de Paul-André Roux. En 2010, alors qu'il était premier vice-président du Grand Conseil et directeur adjoint de l'Idiap, Jean-Albert Ferrez a accéléré le projet. Lors de la sortie annuelle du Service parlementaire, il nous a invités à l'Idiap et nous a montré les travaux en cours dans le domaine de la reconnaissance vocale, évoquant les possibilités d'application au parlement. Quand j'ai vu le système, entièrement en anglais et encore très perfectible, je me suis dit que ça allait prendre encore dix ans avant que ce soit opérationnel. Puis la Loterie et le Canton ont soutenu Mediaparl, et quand nous sommes retournés à l'Idiap deux ans plus tard, les progrès réalisés étaient impressionnants!

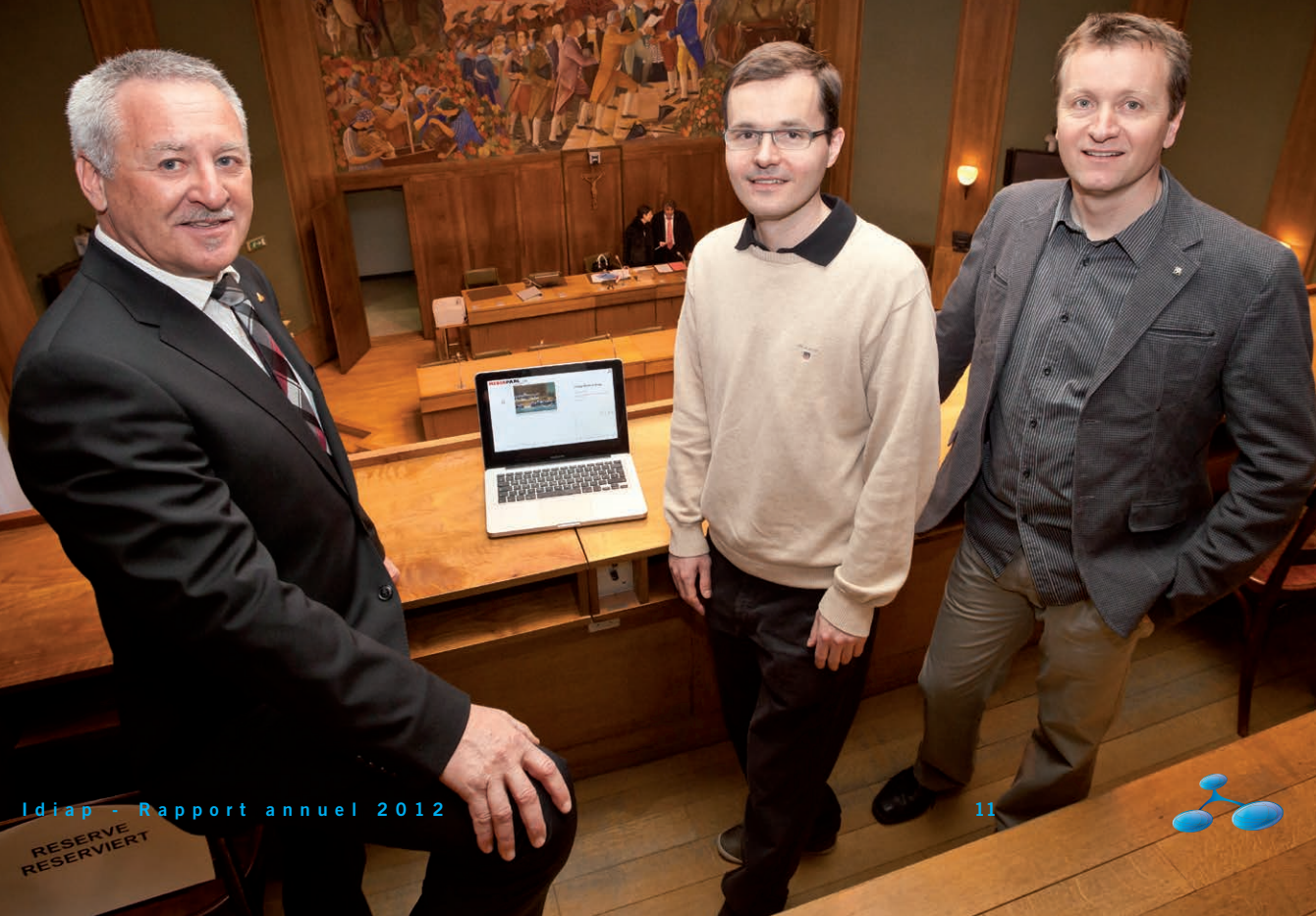
Mediaparl devrait entrer en fonction au printemps 2013, quels en seront les bénéfices?

Aujourd'hui, même si nous disposons d'un moteur de recherche sur internet, ça reste fastidieux de rechercher une intervention faite lors d'une session. Il faut auparavant savoir la date de la session, le jour exact, puis visionner dans les archives de Canal 9 toute la session. Mediaparl permettra de retrouver facilement et rapidement d'anciennes interventions, simplement avec un mot clé. Une vraie révolution! D'ailleurs, quand la presse en a parlé, des collègues nous ont signalé leur intérêt, le parlement vaudois mais aussi le parlement fédéral! S'il se généralise, Mediaparl permettra un précieux échange de réflexions et d'informations entre cantons.

Le second volet de Mediaparl va retranscrire les sessions. La machine remplacerait-elle l'homme?

Je ne crois pas, non. Françoise Manni, qui effectue aujourd'hui ce travail de saisie, gagnera du temps et celui-ci pourra être réinvesti pour améliorer nos prestations à l'intention des députés. Car nous avons encore plein d'idées pour améliorer celles-ci, et la marge de progression existe.

Claude Bumann, chef du Service parlementaire, Alexandre Nanchen, responsable du projet Mediaparl à l'Idiap, et Vincent Bornet, directeur de la télévision régionale Canal 9. (De g. à dr.)



PROJET ARMASUISSE

POUR COMMUNIQUER EN TOUTE SÉCURITÉ

L'Idiap travaille depuis deux ans sur un projet de recherche pour le compte d'armasuisse. Objectif: faciliter et protéger les communications «sensibles».

Il y a environ deux ans, armasuisse* prend contact avec l'Idiap. Son département communication et protection électromagnétique est à la recherche de solutions innovantes pour faciliter et protéger les communications sensibles dans des environnements défavorables. Très vite, la piste de la synthèse vocale, l'un des domaines d'expertise de l'Idiap, est privilégiée et la mission confiée à Philip Garner, chercheur senior. Celui-ci travaille sur ce projet en équipe avec le chercheur Petr Motlicek et l'ingénieur de développement Milos Cernak.

Diminuer le débit en transmettant du texte plutôt que du son

Un exemple permet d'illustrer les enjeux. Dans des cas critiques, la communication à très longue distance se fait en utilisant des radios en bande haute fréquence (HF). «Afin d'assurer la robustesse et la fiabilité de l'échange, explique Phil Garner, il est nécessaire de minimiser la largeur de la bande passante et d'en réserver une part importante pour la répétition de l'information, codée ou non, d'où l'intérêt de diminuer fortement la somme de données à transmettre.» En bref, il faut pouvoir assurer que, même en cas de parasites ou autres problèmes de transmission, le contenu soit encore suffisamment présent pour que le destinataire le comprenne.

C'est ici qu'intervient la reconnaissance vocale: plutôt que de transmettre la parole réelle, on émet principalement du texte. Cette transformation s'opère par le biais d'un appareil développé à l'Idiap, le *Juicer*. A l'arrivée, un synthétiseur vocal – le genre de dispositif qu'utilise pour s'exprimer le célèbre physicien britannique Stephen Hawking – permet de restituer le texte sous forme de parole. Le système peut notamment fonctionner entre deux radios militaires ou via satellite. «Pour l'instant, cette méthode ne permet pas de reproduire les émotions véhiculées par la voix, ni son intensité, précise Phil Garner, mais nous y travaillons.»



Quant au débit, il passe grâce à ce système de plusieurs milliers de bits par seconde (le débit d'une conversation normale) à environ 100 à 200 bits par seconde. La communication s'en trouve facilitée et sécurisée d'autant. «Ce bas débit permet de rendre la communication possible en milieu hostile malgré le bruit ambiant ou malgré des tentatives de brouillage des transmissions.»

Des contingences liées aux particularités helvétiques

Armasuisse reste naturellement discret sur les scénarii d'engagement de ce futur système de communication. «Nous comprenons parfaitement cette attitude, qui d'ailleurs n'handicape en rien nos recherches.»

Dernière particularité du projet: le dispositif devra fonctionner indépendamment de la langue, conformément aux particularités linguistiques helvétiques. Or, l'anglais reste la référence dans le monde de la recherche, à l'Idiap comme dans tout autre institut de rayonnement mondial. «Depuis quelques années, nous essayons de modifier cela et d'utiliser également le français, l'allemand et ses variantes tel le «Walliserdeutsch». C'est un joli défi et ici, à Martigny, nous sommes dans une situation idéale pour le relever!»

Repères

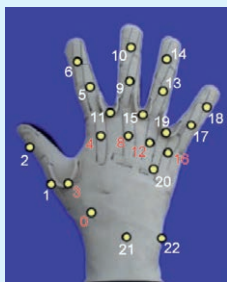
Titulaire d'un Master en ingénierie électronique de l'Université de Southampton, Philip N. Garner rejoint l'Idiap en avril 2007. Ce Britannique de 43 ans, spécialiste du traitement de la parole, avait auparavant travaillé dans l'industrie, pour la firme Canon notamment, dans le Surrey puis à Tokyo. En 2012, il obtient son doctorat à l'Université d'East Anglia (Royaume-Uni).



* Armasuisse est l'un des centres de compétences de la Confédération. Il est notamment chargé de l'acquisition de systèmes et de matériels technologiquement complexes en matière de sécurité.

NOUVELLE GÉNÉRATION DE MAIN ARTIFICIELLE

En associant les progrès de la mécatronique et de la micro-électricité aux algorithmes de l'Idiap, le projet Ninapro développe une nouvelle prothèse de la main, plus habile et plus facile à maîtriser.



Pour la phase de récolte de données, l'Idiap a travaillé avec un «cyberglove», un gant équipé de capteurs pour enregistrer les mouvements quotidiens de personnes valides.



Ce gant (cyberglove) a été développé pour enregistrer et reproduire les mouvements quotidiens de la main.

On imagine les difficultés quotidiennes d'une personne qui doit réapprendre à vivre sans l'une de ses mains... Se laver, s'habiller, conduire, cuisiner: chaque action requiert une aide extérieure et un nouvel apprentissage. Dans la plupart des cas, la main valide compense la perte du membre amputé. Les prothèses actuelles d'électromyographie de surface (contrôlées par un signal électrique émanant des muscles du moignon) se limitent pour l'essentiel à l'ouverture et la fermeture de la main, avec un très faible contrôle. Sans compter que pour les maîtriser, il faut compter des mois d'efforts et d'entraînements répétitifs, souvent démotivants.

Pour capter les mouvements d'une main valide, les électrodes sont placées sur les muscles concernés. En portant ce gant spécial (cyberglove), les testeurs fournissent les données qui permettront ensuite à la main artificielle de faire le travail toute seule. ▼

Comment dès lors améliorer ces mains artificielles? Interpellée à ce sujet par un collègue de l'Institute of Robotics and Mechatronics DLR (Centre aérospatial allemand), Barbara Caputo, chercheuse senior à l'Idiap, décide de relever le défi: «J'ai immédiatement compris que nos travaux de recherche dans le domaine des robots domestiques intelligents pouvaient s'appliquer ici. J'ai aussi été motivée par les enjeux humains du projet, la perspective d'une phase de réhabilitation ludique et rapide pour les patients, avec davantage d'autonomie à la clé!» Le défi: offrir une prothèse qui présente tous les avantages d'un implant, sans ses inconvénients.



Barbara Caputo

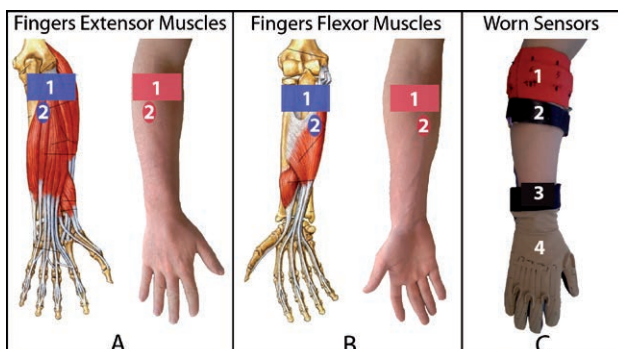
50 volontaires valides pour modéliser leurs mouvements

Lancé en 2011, le projet Ninapro (pour «Non Invasive Adaptive Hand Prosthetics») s'appuie sur les progrès récents de la mécatronique et de la micro-électronique et sur le savoir-faire de l'Idiap en matière de reconnaissance visuelle pour augmenter la dextérité de cette main.

Première étape: la collecte et l'analyse de données. Une cinquantaine de personnes valides y ont participé afin de définir les gestes les plus effectués. «Plus on aura de modèles, plus on aura de chances de succès. Car la portée des gestes dépend des caractéristiques physiques de la personne (âge, taille, force, etc.)», précise Barbara Caputo, qui espère impliquer 1000 volontaires. Suivront les essais avec les personnes amputées. Le projet tentera ensuite d'obtenir des fonds pour trois nouvelles années de recherche, de 2013 à 2016.

Une démo pour convaincre et obtenir des fonds

Pour faire connaître Ninapro, l'Idiap et ses partenaires (DLR et HES-SO) préparent un film de présentation et plusieurs démonstrations interactives. «Les possibilités finales de la prothèse dépendront des besoins des personnes amputées mais aussi de l'industrie des prothèses et de son engagement financier. Nous avons la technologie pour réaliser du sur-mesure comme du low cost.»



RÉSEAU



LA CONFIANCE DES PLUS GRANDS

L'Idiap entretient depuis plusieurs années des relations industrielles avec des géants asiatiques ou européens dans les domaines de la téléphonie mobile et de la sécurité. Une carte de visite pour son secteur de transfert de technologies. Tour d'horizon.

THALES

Vidéosurveillance des lieux publics

Le groupe français Thales a fusionné en 2011 deux de ses filiales pour donner naissance à Thales Communications & Security. Cette nouvelle société devrait lui permettre d'asseoir sa position de numéro un européen des systèmes d'information et de communication sécurisés pour les marchés mondiaux de la défense, de la sécurité et du transport terrestre.

Basée à Gennevilliers (Hauts-de-Seine), Thales Communications & Security emploie 7000 personnes sur neuf sites en France et consacre 20% de son chiffre d'affaires à la recherche et au développement. Partenaire industriel de l'Idiap dans un projet européen lancé en 2010, Thales lui a confié en 2012 un mandat de recherche en lien avec la vidéosurveillance et la détection des bagages oubliés. A l'heure où les caméras se multiplient dans l'espace public, sans forcément d'augmentation parallèle des ressources humaines pour visionner les images, le savoir-faire de l'Idiap en matière de reconnaissance automatique d'image est précieux.

NOKIA Connecting People

Ensemble pour la téléphonie de demain

Pour Nokia et l'Idiap, 2012 a été une année particulière. Leur collaboration a en effet connu l'été dernier un moment important avec l'organisation commune du Mobile Data Challenge (voir page 9), un défi lancé à la communauté mondiale des chercheurs.

Depuis 2009, Nokia et l'Idiap collaborent sur un projet qui vise à récolter des données auprès de 200 volontaires en Suisse romande munis d'un smartphone qui enregistre leur comportement. Aujourd'hui complète, cette base de données constitue un fabuleux outil de recherche. L'Idiap s'est vu confier le droit exclusif de sa distribution pour la communauté scientifique. A noter que les données ont été anonymisées et «watermarkées» (tatouage numérique) pour en assurer la protection et la traçabilité.



Les interactions sociales dans le travail d'équipe

Le Laboratoire de sciences de la communication de NTT (Nippon Telegraph Telecom) travaille avec l'Idiap sur l'analyse des comportements, en particulier des interactions non verbales, dans le travail d'équipe, dont la qualité revêt une grande importance pour les industriels. En utilisant des méthodes probabilistes, le projet NISHA vise à répertorier ces comportements, à analyser leurs propriétés et à utiliser cette connaissance pour caractériser des aspects-clés des relations interpersonnelles.



L'Idiap, seul institut suisse sélectionné

Numéro un des ventes de téléphones portables en Europe, Samsung lance chaque année des appels à projets aux chercheurs du monde entier. En 2011, l'Idiap était le seul institut suisse retenu par le groupe sud-coréen. Il s'était vu confier 100 000 dollars et la mission de développer un nouveau mode de reconnaissance vocale. En 2012, la collaboration s'est poursuivie, l'institut ayant décroché un nouveau projet de recherche autour du thème de la reconnaissance automatique de langage multilingue. Une véritable reconnaissance et une fierté pour l'Idiap.

NEC

Aider la machine à comprendre le sens d'un texte

Le groupe japonais NEC (Nippon Electronic Company), qui possède deux laboratoires aux Etats-Unis, finance une thèse de doctorat à l'Idiap. Le but du projet: faire comprendre à la machine la structure, et par déduction le sens d'un texte. Une mission extrêmement complexe qui débute par de petites tâches simples.



INTERNATIONAL CREATE CHALLENGE

DE L'IDÉE À LA START-UP EN TROIS SEMAINES



L'Idiap a lancé en 2012 un nouveau concours, l'International Create Challenge (ICC). Une occasion unique pour de jeunes chercheurs de se muer en entrepreneurs, et ce sous l'aile de l'Idiap. Entre nouvelles sociétés en création et projets de recherche en cours de négociations, la première édition de l'ICC fut un succès!

Pour un coup d'essai, ce fut un coup de maître: septante dossiers soumis, vingt-huit chercheurs sélectionnés puis rassemblés en équipes, douze projets présentés, six primés, plusieurs start-up en cours de création et de nombreux projets de recherche en cours de négociations. La première édition de l'International Create Challenge (ICC), qui s'est tenue du 1^{er} au 21 septembre 2012, a connu un énorme succès.

L'ICC est né à l'initiative de l'Institut de recherche Idiap, du Pôle de recherche national IM2 et d'IdeArk. Il offre à de jeunes chercheurs une chance unique doublée d'un défi de taille: passer de l'idée à la start-up en l'espace de trois semaines. Le lancement de l'ICC résulte d'un constat, ainsi que l'explique François Foglia, directeur adjoint de l'Idiap et créateur de l'événement. «Les chercheurs sont souvent débordés par leurs travaux ou leurs publications scientifiques, si bien qu'ils ne prennent pas le temps de développer leurs idées de produits commerciaux. Grâce à cet événement, nous créons un environnement adéquat afin qu'ils puissent s'extraire de leur quotidien et se consacrer uniquement à leur projet, en espérant aboutir à un prototype au terme de l'événement.»

Trois semaines d'immersion, encadrés par des spécialistes

L'environnement adéquat? Un break de trois semaines, tous frais payés, à l'Idiap, au Centre du Parc de Martigny. Un programme en immersion complète durant lequel les participants bénéficient de conseils d'experts en matière de création d'entreprise. Ils sont ainsi orientés dans les domaines de la recherche de financement, de la gestion de projet, du modèle d'entreprise et de la propriété intellectuelle, tout en découvrant l'art de présenter un projet en nonante secondes!...

Par cette initiative, l'Idiap entend favoriser le transfert de technologies, l'éclosion de nouvelles start-up et avec elles la création d'emplois. «Cela fait partie de la mission de notre institut de recherche. Cet événement permet aussi de donner une certaine visibilité au Valais technologique.»

L'aide prodiguée l'est aussi sur le plan scientifique, puisque les concurrents disposent de deux jours pour échanger idées et résultats avec une cinquantaine de chercheurs travaillant au sein du Pôle de recherche national IM2. A ces conditions idéales s'ajoutent, pour le vainqueur, l'accès gratuit durant trois ans à l'incubateur The Ark, sur le site d'IdeArk à Martigny (prix The Ark), ainsi qu'une somme de 10000 francs suisses (prix IdeArk).

Une perspective qui a fait mouche. Lancé au début de l'année 2012, l'appel des organisateurs a séduit pas moins de septante personnes travaillant dans les domaines de recherche de l'Idiap. «Nous avons d'abord retenu cinquante-cinq dossiers, puis nous avons mis les candidats en contact sur le réseau social LinkedIn pour qu'ils forment des équipes de projets, précise François Foglia. Au final, nous avons sélectionné douze équipes, soit vingt-huit personnes.»

Un concours international de haut niveau

Les candidats retenus proviennent de l'Idiap, de l'EPFL, de l'EPFZ, de l'Université de Genève, de celle de Fribourg, mais aussi d'autres institutions en Grèce, en Roumanie ou encore en Ecosse. «Nous voulions un concours international, nous avons réussi ce pari», se réjouit celui qui est aussi directeur ad interim d'IdeArk et président de l'association liée au Pôle de recherche IM2.

A l'arrivée, le jury de l'ICC'2012 a récompensé pas moins de six projets, en tête desquels VocaBoca, une application destinée à faciliter l'apprentissage des langues via smartphone ou tablette numérique. La compétition a en outre débouché sur la création de nouvelles start-up. «C'est très satisfaisant, tout comme la qualité d'ensemble des projets présentés. Cela dit, le succès de cette première édition ne se mesure pas seulement au nombre de sociétés créées. L'expérience a aussi permis à de nouveaux réseaux et projets de recherche de se former.»

Autant d'excellentes raisons de reconduire l'International Create Challenge. Ce sera chose faite dès 2013, avec une deuxième édition prévue entre le 21 septembre et le 11 octobre.



APPRENDRE UNE LANGUE AVEC SON SMARTPHONE

Entretien avec Alberto Daccarett Armijo, dont l'application VocaBoca a remporté l'ICC'2012.

Alors qu'ils sont en train de travailler sur leur application, Alberto Daccarett Armijo, son frère Elías et Javier Sierra remarquent sur le web une annonce présentant l'International Create Challenge 2012. Conscient de ce que l'expertise de l'Idiap pourrait apporter à son projet, le trio colombiano-espagnol, formé de deux analystes financiers et d'un développeur, pose sa candidature. Bien lui en a pris puisque son application VocaBoca a remporté le concours.

VocaBoca est une application pour smartphone et tablette permettant de mémoriser de nouveaux mots de manière ludique et de créer ainsi son propre dictionnaire sonore illustré. Il suffit par exemple de photographier un objet pour obtenir, en quelques secondes, sa description, sa définition, sa traduction et sa prononciation. Ou de signaler un mot inconnu, par écrit ou par oral, pour en découvrir le sens et la photo.



«Le concours de l'Idiap est fondamental pour le développement de notre application, il nous permet de gagner du temps», se réjouit Alberto Daccarett Armijo, qui livre ses sentiments sur cette aventure.

Qu'est-ce qui a motivé votre participation à l'ICC?

Le concept nous a beaucoup plu. Mes coéquipiers et moi-même n'habitons pas le même pays – Javier et moi sommes à Genève, Elías vit en Espagne – ce qui nous oblige à travailler souvent par téléphone. Là, on nous offrait l'opportunité de nous focaliser pendant trois semaines sur le projet, sans distractions, sans copains, sans famille. Donc on pouvait travailler tard, jusqu'à minuit, même le samedi et le dimanche!

Quelle était l'ambiance durant ces trois semaines?

Géniale! Même si elles étaient en compétition, les équipes formaient une petite communauté. D'ailleurs, nous avons gardé le contact avec d'autres participants. Durant le concours, on échangeait des idées, on s'entraidait. Nous avons été encadrés par des investisseurs, des coaches et des chercheurs dont les conseils nous ont été très utiles. VocaBoca n'est que notre deuxième application, nous avons encore beaucoup à apprendre.

A quel stade en est aujourd'hui VocaBoca?

Les utilisateurs trouvent l'application attrayante, mais peu sont disposés à l'acheter. Le plus urgent pour nous est de mettre en place un modèle d'affaires, puis d'améliorer notre application en se demandant pour quel type de fonctionnalités les gens seraient prêts à payer. C'est presque le plus difficile pour nous! Nous employons aujourd'hui les outils génériques du web, mais à terme nous espérons utiliser les technologies de l'Idiap, en matière de reconnaissance visuelle et de reconnaissance de la parole notamment. Des milliers de nouvelles applications apparaissent chaque semaine, qui disparaissent après quelques jours; c'est un marché très concurrentiel. Une fois que nous aurons établi le business model, l'incubateur The Ark jouera un rôle important pour la phase d'exécution.

Si vous deviez tirer le bilan de cette expérience?

Dans les domaines des connaissances professionnelles, du réseautage et de la reconnaissance médiatique, ça aide, ça aide beaucoup! J'ai mentionné la deuxième édition du Challenge sur mon mur Facebook et je vois qu'il y a déjà des gens intéressés. Je leur conseille vraiment de participer, parce que c'est une expérience formidable et qu'ils apprendront beaucoup de choses.

www.vocaboca.com



CROISSANCE ET NOUVEAUX ENGAGEMENTS

Nées des technologies de l'Idiap, au cœur de l'incubateur The Ark, sur le site d'IdeArk, les start-up KeyLemon, Klewel et Koemei ont vécu une année 2012 marquée par la croissance. Toutes trois ont élargi leurs rangs. Interviews express des nouveaux arrivés.

Klewel

www.klewel.com

Klewel fournit depuis 2007 des solutions innovantes pour enregistrer, archiver et diffuser sur le web des présentations et conférences. Son outil principal, Triskel, est une solution de webcasting «clé en mains» composée d'une station d'enregistrement audiovisuel qui tient dans une valise. Celle-ci est liée à une plateforme web permettant de référencer, éditer et publier le contenu audio et vidéo de l'orateur et de sa présentation (diapositives projetées). Chacun peut ensuite voir la présentation sur le site internet de l'institution, et y rechercher un contenu précis à l'aide de mots-clés. Sa solution a obtenu en 2008 le label européen d'excellence au forum international du CeBit et a été retenue en 2010 comme finaliste de l'Innovation Award de l'Association internationale des Palais de Congrès.

Klewel assure aujourd'hui les enregistrements des conférences de grandes sociétés telles que Nestlé, la BCV, l'EPFL, l'Unicef ou encore l'EHL. En 2013, Klewel compte intégrer à son outil la reconnaissance vocale et un portail de conférences publiques.



La nouvelle arrivée: Dr Sandy Ingram

29 ans – Libanaise – domiciliée à Martigny (Valais)

Ingénieure Recherche & Développement

«Impliqué à la fois dans un projet de recherche européen et dans un projet fédéral, Klewel avait besoin de ressources supplémentaires en R&D. Je travaille donc depuis décembre 2012 sur le développement d'un portail web permettant d'intégrer différentes technologies mises à disposition par six partenaires européens dont IBM Research et l'Université d'Edimbourg. Je pense que, demain, les solutions Klewel offriront un point d'accès à un pool de présentations de conférences enregistrées, éditées et publiées de manière quasi automatique et intuitive!»

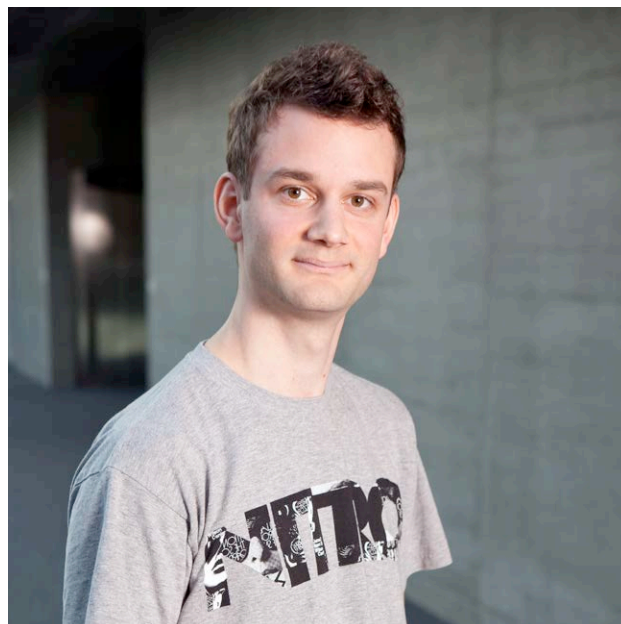


www.keylemon.com

Créée en 2008 par un doctorant de l'Idiap, un économiste et un professeur en entrepreneuriat, KeyLemon a développé un logiciel de reconnaissance faciale permettant d'ouvrir sa session en présentant son visage à son ordinateur. Depuis 2011 l'outil a été adapté aux smartphones.

L'application compte plus de 3 millions d'utilisateurs et KeyLemon a déjà accordé la licence de sa technologie à une société qui revend des solutions bancaires pour toute l'Amérique du Sud et à une société MedTech en Europe.

En 2012, KeyLemon a signé un partenariat avec Fujitsu: sa technologie sera préinstallée sur toute une gamme d'ordinateurs du géant japonais. L'année a également été marquée par l'intégration dans son outil de la reconnaissance vocale et par un projet de déploiement du logiciel sur une plateforme Cloud.



Le nouvel arrivé: Romain Cherix

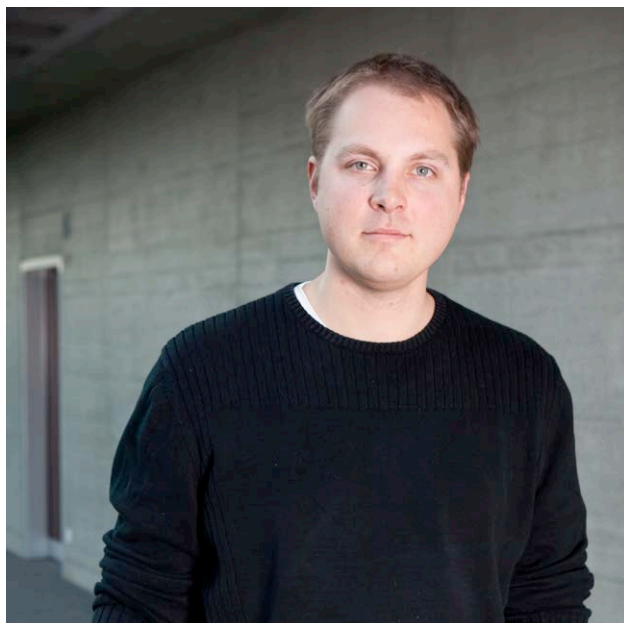
24 ans – Suisse – domicilié à Bex (Vaud) – Ingénieur de développement software

«Je travaille chez KeyLemon depuis le 1^{er} octobre 2012: l'entreprise avait besoin de ressources supplémentaires pour le développement d'une nouvelle plateforme Cloud. Je m'occupe de développement Web ainsi que de la réalisation d'applications de démonstration pour nos clients. Je pense que la technologie KeyLemon, demain, sera installée sur la plupart de nos outils informatiques. C'est un tel gain de sécurité et de temps!»

KOEMEI

www.koemei.com

En 2015, 90% de ce qui se trouvera sur le web le sera sous forme vidéo. Or, aucun référencement n'existe pour ce genre de contenu. Koemei propose une solution de retranscription vocale qui distingue les différents locuteurs et permet aux moteurs de recherche d'y accéder. La solution Koemei est devenue l'alliée des formations supérieures. Elle a été notamment adoptée par l'Université de Genève, l'IMD Business School de Lausanne et l'Université de Californie à Berkeley. Koemei se profile également



sur le marché du sous-titrage pour sourds et malentendants. En 2012, Koemei a inauguré une antenne américaine à San Francisco. Elle a également été sélectionnée pour participer au Demo Spring, au cœur de la Silicon Valley, une prestigieuse manifestation qui vise à faire connaître les nouvelles technologies du domaine de l'ICT ainsi qu'au Disrupt TechCrunch à New York. 800 entreprises avaient déposé leur candidature pour le Demo Spring, et 50 seulement ont été retenues.

Le nouvel arrivé: Sébastien Dupont

30 ans – Belge – Domicilié à Daillon (Valais) – Lead developer

«Afin de mettre à disposition les fonctionnalités d'indexation et de sous-titrage qui sont rendues possibles par sa technologie de retranscription, Koemei a construit une plateforme web. Mon rôle consiste à faire évoluer ce service avec l'aide de mon équipe: le rendre robuste, facile d'utilisation et capable de s'adapter à la demande. A terme, Koemei offrira une solution permettant d'améliorer l'accessibilité de tous les médias audiovisuels, dont il est encore aujourd'hui difficile de tirer parti.»



VISAGES



LAKSHMI SAHEER, PREMIÈRE DOCTEUR INDIENNE DE L'IDIAP

«MES COLLÈGUES ICI SONT ÉPATANTS.»

Depuis son arrivée en Suisse, voici cinq ans, Lakshmi Saheer est devenue docteur, a fondé une famille et s'apprête à lancer sa propre start-up. Rencontre avec une jeune femme qui aime gravir les montagnes.

Vous êtes née en Inde il y a trente-deux ans. Pourquoi avoir choisi la Suisse pour poursuivre vos études?

Lorsque je faisais mon Master à l'Indian Institute of Technology, à Madras, je connaissais l'EPFL et l'Idiap de réputation. L'école doctorale Idiap-EPFL m'a paru idéale pour mon doctorat. Le traitement de la parole est un domaine qui m'intéresse depuis toujours, et l'Idiap est l'un des meilleurs instituts au monde dans ce domaine. Les cours à l'EPFL, la recherche à l'Idiap, c'était un mix idéal! (Ndlr: Lakshmi a obtenu son doctorat en avril 2013) Et puis avant de venir, j'ai travaillé une année en Suède pour le compte de Sony Ericsson, et j'ai réalisé que je me sentais bien en Europe.

Qu'est-ce qui vous a le plus surpris en arrivant ici?

Je n'étais jamais venue en Suisse. On m'avait dit que ce pays était très beau, mais je n'avais pas imaginé à quel point. Jamais je n'avais vu une nature aussi magnifique! Au début, j'ai été un peu surprise par les horaires hebdomadaires de travail, bien plus longs qu'en Suède. Mais aujourd'hui je pense que cette durée explique le haut niveau de productivité et d'efficacité que l'on rencontre ici.

Quel souvenir gardez-vous de votre tout premier jour à l'Idiap?

Je ne m'attendais pas à ce que l'institut soit aussi grand. L'accueil a été très chaleureux, peut-être aussi parce que j'étais la première chercheuse indienne à l'intégrer. En entrant dans mon bureau, j'ai été emballée par la vue qu'il offrait, ce magnifique panorama de montagnes. Avec mon mari, nous avons fait plusieurs marches, notamment à la Pierre Avoi et à la cabane Panossière, où nous avons passé une nuit. J'en garde un souvenir formidable et j'ai hâte de pouvoir y retourner en famille. Ça se présente bien puisque ma fille, âgée de 16 mois, aime déjà beaucoup marcher!

Comment la femme que vous êtes se sent-elle dans le monde de la recherche?

Les femmes étant moins nombreuses, il y a peut-être moins d'amitiés à partager. Cependant mes collègues ici sont époustouflants, je ne subis ni discrimination ni pression. Cela dit, lorsque j'ai reçu la Bourse d'études Anita Borg de Google (ndlr: réservée aux étudiantes), à Zurich, en 2011, j'ai quand même trouvé cette communauté de femmes extrêmement stimulante.

Et vous êtes non seulement femme, mais maman...

Effectivement! Mon mari me soutient énormément, mais je dois avouer que c'est la culture «family friendly» de la Suisse et de l'Idiap en particulier qui m'a permis d'achever mon doctorat et d'oser lancer mon entreprise. Mes collègues et connaissances m'aident beaucoup.

Où envisagez-vous votre avenir: en Europe, en Inde, ailleurs dans le monde?

A long terme, je veux retourner en Inde auprès de ma famille et de mes amis. Je tiens à donner quelque chose en retour à la société d'où je viens et je trouve important que ma fille connaisse ce pays. Mais pour ces prochaines années, nous avons envie de rester en Suisse, c'est un endroit idéal pour nous.

Pour terminer, histoire de tester votre intégration: connaissez-vous la raclette?!

Oui, bien sûr, et la fondue aussi! En Suisse, pour une végétarienne qui aime le fromage, il est très facile de se nourrir.



De l'émotion dans les messages vocaux

Lakshmi Saheer s'apprête à lancer sa propre start-up. Son projet, baptisé Emogen (pour *Emotion Generation*), permet d'ajouter des émotions et expressions dans tout type de message vocal, de synthèse en particulier. Cette technologie pourrait enrichir considérablement le discours d'une personne privée de la parole et s'exprimant par l'intermédiaire d'un synthétiseur vocal.

Présenté à l'International Create Challenge 2012 (voir pages 16-17), Emogen a été distingué dans plusieurs catégories (The pitch challenge Award, Venture Kick, EuroFin Ventures, IMD Start-up Competition).



ÉQUIPE GESTION DE PROJETS

LES MANAGERS DE LA RECHERCHE

Formulaires, contacts, délais, contrats, budgets... Rien ne leur fait peur! L'équipe de gestion de projets soutient les chercheurs dans leurs démarches administratives pour décrocher fonds et partenariats.

En 2012, l'Idiap a géré 42 projets scientifiques. Un record! Et la progression devrait se poursuivre: fin janvier 2013, pas moins de 15 projets avaient déjà été soumis. Répondre aux appels d'offres, coordonner les projets avec les partenaires, suivre les mandats obtenus: ces tâches prennent de plus en plus d'importance et contribuent directement au succès de l'Idiap. C'est pourquoi l'équipe de gestion de projets a été renforcée en 2012. Ce sont ainsi désormais trois personnes qui secondent les équipes de recherche dans leurs démarches administratives. Léonore Miauton, Christophe Ecoeur et Martina Guetl forment un team aux compétences très complémentaires. Ils veillent à la soumission, au démarrage et au suivi de chaque projet. A l'affût des virgules, des centimes, des guide-lines et des dead-lines, rien ne leur échappe. «Nous sommes impliqués à toutes les étapes d'un projet, c'est passionnant!», s'enthousiasme Christophe Ecoeur.

Rédaction de rapports, répartition des fonds, etc.

Par exemple, lorsque l'Idiap répond à un appel d'offres de l'Union européenne, il doit se conformer à une foule de critères. Les chercheurs qui relèvent le défi ne disposent souvent que de quelques semaines pour rédiger les documents et rassembler les pièces nécessaires à l'élaboration d'un dossier qui peut aller jusqu'à 200 pages. Certains mandats exigent l'impli-

cation d'autres partenaires en Europe, avec lesquels il faut négocier et signer des contrats. Une fois l'argent obtenu, la gestion de projets commence. «L'administration européenne est très pointilleuse. Nous devons produire un rapport annuel, redistribuer l'argent aux partenaires, justifier nos budgets au centime près», relève Léonore Miauton. Il faut également s'adapter aux chercheurs et à leur culture, à leurs outils informatiques, leurs méthodes de travail. Ici, l'administration ne rime pas avec routine. Chaque projet est une aventure unique, avec ses acteurs, ses règles, ses enjeux propres.

Participation aux événements régionaux

Avec sa connaissance parfaite des dossiers, la nouvelle équipe est bien placée pour valoriser et vulgariser les travaux de recherche de l'institut. C'est elle qui gère la participation de l'Idiap aux différents événements régionaux. Exemples: les démonstrations technologiques dans le cadre de «Oser tous les métiers» et l'animation des ateliers scientifiques du festival «Hérisson sous gazon». C'est aussi au trio de la gestion de projets que revient la tâche d'organiser les semaines de l'ICC (International Create Challenge, lire pages 16-17), un programme qui offre une opportunité unique aux entrepreneurs et chercheurs du monde entier de développer leur produit, voire de lancer une start-up.



Trois gestionnaires de projets au service des chercheurs:
Léonore Miauton, Christophe Ecoeur et Martina Guetl.



Christophe Ecoeur, l'économiste
«Je n'ai pas hésité une seconde!»

Lorsque la Direction de l'Idiap a proposé à l'assistant du service comptabilité de rejoindre la gestion de projets, la réponse a fusé: «J'y vais!» Christophe Ecoeur ne croit pas au hasard. Cette opportunité s'est présentée alors qu'il recherchait un nouveau défi professionnel. «Je suivais déjà les budgets des projets, mais j'ai désormais davantage de responsabilités et de contacts avec les gens. Exactement ce dont je rêvais!» Christophe Ecoeur connaît bien la maison. Il y est entré en 2010 après un Bachelor en économie d'entreprise. Il apprécie particulièrement la bonne ambiance qui règne au bureau. «Nous nous complétons parfaitement, c'est un plaisir de travailler ensemble!»



Martina Guetl, la littéraire
«Je découvre un monde nouveau»

Rien ne la prédestinait à travailler dans un milieu scientifique. Cette Autrichienne de 26 ans est arrivée en Valais par amour, après des études en lettres et en administration publique à Lausanne. Sa maîtrise des langues et son aisance dans la communication ont séduit la Direction de l'Idiap, qui l'a engagée en novembre 2012. «J'appréhendais de m'immerger dans un monde aussi spécialisé. Mais en fait, je m'y plais beaucoup! J'apprends une foule de choses nouvelles et j'apprécie le côté multinational de l'institut.» Ses compétences de rédactrice et de traductrice s'avèrent indispensables dans la gestion de projets.



Léonore Miauton, la scientifique
«On nous laisse une grande liberté»

Physicienne formée à l'EPFL et à l'EPFZ, Léonore Miauton connaît parfaitement les rouages des projets scientifiques. Elle a travaillé dans la gestion des inventions aux Etats-Unis à Baltimore (Johns Hopkins University) pendant cinq ans avant de fonder sa propre entreprise d'informatique à Singapour. Son dernier poste à l'EPFL l'a amenée à conduire de grands projets pour les pôles de recherche nationaux. Elle a signé à l'Idiap en septembre 2012, heureuse de se lancer dans une nouvelle aventure au service des chercheurs. «Dans cette structure à taille humaine, la Direction nous encourage à amener nos idées, à faire nos propositions. C'est une chance!» Léonore profite du contact avec les chercheurs pour se tenir au courant des dernières découvertes. Son expérience du terrain est très précieuse.



ANTOINE DORSAZ, AIDE-COMPTABLE

«LES FRUITS DE LA PASSION»

Ancien sportif d'élite, Antoine Dorsaz partage désormais son temps entre l'Idiap et la HES-SO Valais, où il poursuit ses études en économie. Rencontre avec un jeune homme perfectionniste et curieux de tout.

Il porte sur son poignet droit un tatouage représentant les anneaux olympiques. Une inscription discrète, témoignage de son existence antérieure de sportif d'élite. A 24 ans, Antoine Dorsaz a déjà vécu plusieurs vies. Hier sur les patinoires du monde entier, aujourd'hui à l'Idiap, en tant qu'assistant comptable. Un poste à temps partiel qui lui permet de poursuivre en parallèle des études d'économie à la HES-SO Valais.

Le quotidien d'Antoine a changé mais la passion demeure intacte. «J'ai envie d'être bon dans tout ce que j'entreprends.» Premier domaine dans lequel le Valaisan exprime sa soif d'excellence, le patinage artistique. «J'ai commencé à 10 ans, ce qui est assez tard, si bien que j'ai dû travailler très dur pour combler mon retard sur les autres. J'ai adoré ça!»



Un rêve: travailler au sein d'une entreprise internationale

Cette discipline exigeante, qui le conduira avec sa partenaire Anaïs Morand jusqu'aux Jeux Olympiques de Vancouver, en 2010, constitue une formidable école de vie. «Il faut être très perfectionniste, persévérant, ne pas avoir peur de tomber et de tomber encore. Entre un saut réussi et une chute, il n'y a pas grand-chose.»

Après une 15^e place aux J.O., Antoine Dorsaz met un terme à sa carrière sportive. «J'avais réalisé mon rêve. Je n'avais plus la motivation pour m'entraîner quotidiennement au plus haut niveau.» Surtout, le Valaisan a envie de se sentir connecté à la «vraie vie». «La compétition, les voyages, c'est une vie géniale, mais on est un peu dans une bulle.»

Parce qu'il est «curieux de tout», Antoine veut reprendre des études et travailler. Amoureux des voyages et des langues – il pratique l'anglais, l'allemand et le russe – il rentre dans son Valais natal avec un rêve: travailler au sein d'une entreprise internationale.

«Les chercheurs? Des «cerveaux» abordables»

Lorsqu'il apprend que l'Idiap cherche un assistant comptable à 50%, Antoine postule sans attendre: «Je connaissais l'institut de réputation. C'est un formidable atout pour le Valais!» Il s'y sent aujourd'hui comme un poisson dans l'eau. «Je côtoie

des gens de toutes cultures, j'aime parler avec eux. Ces chercheurs sont à la fois des «cerveaux» et des gens détendus et abordables: c'est très impressionnant. J'ai un grand respect pour ce qu'ils font. Quant à mon travail, j'ai la chance d'être dans une équipe sympathique, bien organisée et efficace. Rien ne traîne, j'apprécie beaucoup!» Grâce à ce temps partiel à l'Idiap, le jeune Valaisan peut mener en parallèle ses études à la HES-SO Valais de Sierre, où il vise un Bachelor en économie d'entreprise, dévorant des ouvrages sur l'économie.

«C'est intéressant, surtout dans le contexte actuel, de comprendre les enjeux, les erreurs commises...»

Comme si ses journées n'étaient pas assez remplies, Antoine trouve encore le temps de donner des cours à de jeunes patineurs. «Transmettre, c'est gratifiant. Et cela me permet de garder un pied dans le monde du patinage.» Un univers pour lequel il lui arrive, avoue-t-il, d'éprouver un soupçon de nostalgie.

Palmarès sportif

Principaux résultats du couple A. Dorsaz/A. Morand:

Saison 2009-2010

- Jeux Olympiques, Vancouver: 15^{es}
- Championnats du monde, Turin: 13^{es}
- Championnats d'Europe, Tallin: 8^{es}

Saison 2008-2009

- Championnats du monde, Los Angeles: 14^{es}
- Championnats du monde juniors, Sofia: 10^{es}
- Championnats d'Europe, Helsinki: 12^{es}

ALLÉES ET VENUES

En 2012, on enregistre 18 départs pour 21 arrivées. Parmi les nouveaux venus, on compte deux program managers, un aide-comptable, sept postdoctorants et onze doctorants.

ILS SONT ARRIVÉS EN 2012

Prénom, nom, fonction, origine, domicile

Chindansh Bhatt, postdoctorant, Inde	Joël Legrand, doctorant, France
Olivier Canévet, doctorant, France	Adolfo Lopez Mendez, postdoctorant, Espagne
Antoine Dorsaz, aide-comptable, Suisse, Fully	Léonore Miauton, senior program manager, Suisse, Chexbres
Nesli Erdogmus, postdoctorante, Turquie	Youssef Oualil, doctorant, Maroc
Paul Gay, doctorant, France	Nikolaos Pappas, doctorant, Grèce
Martina Guetl, junior program manager, Autriche	Novi Patricia, doctorante, Indonésie
Manuel Günther, postdoctorant, Allemagne	Darshan Santani, doctorant, Inde
Pierre-Edouard Honnet, doctorant, France	Gyorgy Szaszak, postdoctorant, Hongrie
Ilja Kuzborskij, doctorant, Lituanie	Romain Tavenard, postdoctorant, France
Alexandros Lazaridis, postdoctorant, Grèce	Raphaël Ullmann, doctorant, Suisse, Lausanne
Rémi Leuret, doctorant, France	

ILS REPARTENT

Prénom, nom, fonction, origine, année d'arrivée à l'Idiap, nouvel employeur

Constantin-Cosmin Atanasoaei, doctorant, Roumanie, 2008, Pix4D, Ecublens
Venkatesh Bala Subburaman, doctorant, Inde, 2007, Multitel Mons, Belgique
Cheng Chen, postdoctorant, Chine, 2010, ISTB, Université de Berne, Berne
Valérie Devanthéry, program manager, Suisse, 2008
Stefan Duffner, postdoctorant, Allemagne, 2008
Jean-Albert Ferrez, directeur adjoint, Suisse, 2001, Energie de Sion Région, Sion
Pierre Ferrez, program manager, Suisse, 2004, HES-SO Valais, Sion
Danil Korchagin, postdoctorant, Russie, 2008, Aicue Sàrl, Martigny
Gwenolé Lecorvé, postdoctorant, France, 2011, Université de Rennes, Rennes, France
Hui Liang, doctorant, Chine, 2008, ETHZ, Zurich
Christopher McCool, postdoctorant, Australie, 2008, NICTA, Brisbane, Australie
Youssef Oualil, doctorant, Maroc, 2012, Universität des Saarlandes, Saarbrücken, Allemagne
Edgar Francisco Roman Rangel, doctorant, Mexique, 2008, Université de Genève, Genève
Serena Soldo, doctorante, Italie, 2009
Gokul Thattaguppa Chittaranjan, doctorant, Inde, 2010
Fabio Valente, chercheur, Italie, 2005
Jagannadan Varadarajan, doctorant, Inde, 2008, Illinois at Singapore Pte LTD, Singapour
Roy Geoffrey Wallace, postdoctorant, Australie, 2010, Zap Technology, Brisbane, Australie



DISTINCTIONS

PRIX IDIAP

Interne

Chaque année, l'Idiap décerne deux prix destinés à ses doctorants. Le premier récompense une recherche, le second une publication. Pour l'attribution du prix Idiap de la Recherche, le candidat est évalué par une commission interne sur la base de cinq critères: ses publications, sa collaboration dans l'équipe, son implication dans le projet, son sens de la communication et son autonomie. Pour le prix de la Publication, une première sélection est effectuée par les seniors de l'institut parmi les travaux dont l'auteur principal est un doctorant Idiap. Les membres du Comité d'accompagnement notent ensuite, séparément et de façon anonyme, les écrits choisis.

En 2012, le prix de la Recherche a été attribué à **David Imseng** pour la qualité de ses publications et l'excellence de ses recherches traitant de la reconnaissance multilingue de la parole. Celui de la publication a été remis à **Lakshmi Saheer** pour son excellent article scientifique intitulé «Vocal Tract Length Normalization for Statistical Parametric Speech Synthesis».



David Imseng



Lakshmi Saheer

Externe

Cette année, l'Idiap tient à relever la très brillante participation de ses chercheurs lors des conférences internationales. La qualité de leur recherche a été récompensée par plusieurs prix.

Daniel Gatica-Perez et Kate Farrahi

Best Paper Award, IEEE Int. Symposium on Wearable Computers, juin 2012

Extracting Mobile Behavioral Patterns with the Distant N-Gram Topic Model

Kenneth Funes

Best Student Paper award, CVPR workshop on Gesture Recognition, juillet 2012

Gaze Estimation from Multimodal Kinect Data

Anindya Roy

European Biometrics Industry Award 2012, septembre 2012

A Fast Parts-Based Approach to Speaker Verification Using Boosted Dice Classifiers

Dinesh Babu Jayagopi et Dairazalia Sanchez-Cortes

Outstanding Paper Award at ICMI 2012, octobre 2012

Linking Speaking and Looking Behavior Patterns with Group Composition, Perception, and Performance

Daniel Gatica-Perez

Paper Award, ACM Int. Conf. on Multimodal Interaction, octobre 2012

Linking Speaking and Looking Behavior Patterns with Group Composition, Perception, and Performance

Eric Malmi, Trinh-Minh-Tri Do et Daniel Gatica-Perez

Best Paper Paper Award Nomination at MUM 2012, décembre 2012

Checking In or Checked In: Comparing Large-Scale Manual and Automatic Location Disclosure Patterns

THÈSES ACHÉVÉES

Sept étudiants ont achevé leur thèse en 2012: Venkatesh Bala Subburaman, Constantin-Cosmin Atanasoaei, Jagannadan Varadarajan, Hui Liang, Tatiana Tommasi, Edgar Francisco Roman-Rangel, Lakshmi Babu Saheer

- ***Alternative Search Techniques for Face Detection Using Location Estimation and Binary Features***

Venkatesh Bala Subburaman, 14 février 2012

Directeurs de thèse: Prof. Hervé Bourlard, Dr Sébastien Marcel

Membres du jury: Dr Hazim Kemal Ekenel, Prof. Abdenour Hadid, Prof. Jean-Philippe Thiran

- ***Boosting Multivariate Look-up Tables for Face Processing***

Constantin-Cosmin Atanasoaei, 27 avril 2012

Directeurs de thèse: Prof. Hervé Bourlard, Dr Sébastien Marcel

Membres du jury: Prof. Josef Kittler, Prof. Timothy Cootes, Dr Vincent Lepetit

- ***Sequential Topic Models for Mining Recurrent Activities and their Relationships: Application to Long Term Video Recordings***

Jagannadan Varadarajan, 9 juillet 2012

Directeur de thèse: Dr Jean-Marc Odobez

Membres du jury: Prof. Shaogang Gong, Prof. Bernt Schiele, Prof. Pascal Frossard

- ***Data-Driven Enhancement of State Mapping-Based Cross-Lingual Speaker Adaptation***

Hui Liang, 30 octobre 2012

Directeurs de thèse: Prof. Hervé Bourlard, Dr John Dines

Membres du jury: Prof. Keiichi Tokuda, Prof. Tanja Schultz, Dr Jean-Marc Vesin

- ***Learning to Learn by Exploiting Prior Knowledge***

Tatiana Tommasi, 7 novembre 2012

Directeurs de thèse: Prof. Hervé Bourlard, Dr Barbara Caputo

Membres du jury: Prof. Jean-Philippe Thiran, Prof. Vittorio Ferrari, Prof. Jim Little

- ***Statistical Shape Descriptors for Ancient Maya Hieroglyphs Analysis***

Edgar Francisco Roman-Rangel, 28 novembre 2012

Directeurs de thèse: Dr Daniel Gatica-Perez, Dr Jean-Marc Odobez

Membres du jury: Prof. Jean-Philippe Thiran, Prof. Stéphane Marchand-Maillet, Dr Changhu Wang

- ***Unified Framework of Feature Based Adaptation for Statistical Speech Synthesis and Recognition***

Lakshmi Babu Saheer, 28 novembre 2012

Directeur de thèse: Prof. Hervé Bourlard

Membres du jury: Prof. Alan W. Black, Dr Ralf Schlüter, Prof. Jean-Philippe Thiran



FINANCES



COMPTE D'EXPLOITATION

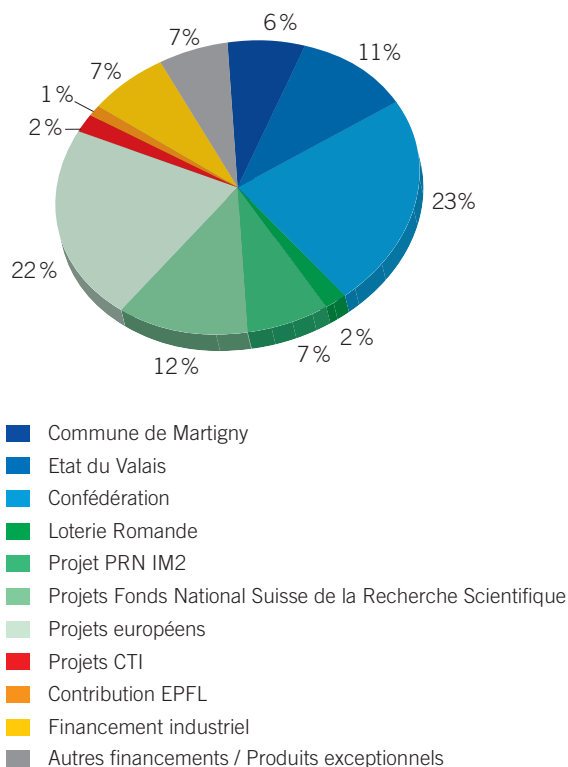
(En francs suisses)

	2011	2012	%
PRODUITS			
Commune de Martigny	650 000	650 000	6,29%
Etat du Valais	900 000	1 120 000	10,83%
Confédération	2 357 000	2 393 000	23,14%
Loterie Romande	-	250 000	2,42%
Projet PRN IM2	533 900	764 071	7,39%
Projet Fonds National Suisse de la Recherche Scientifique	1 072 445	1 203 153	11,64%
Projets européens	2 390 855	2 219 526	21,47%
Projets CTI	105 135	251 168	2,43%
Contribution EPFL	72 000	72 000	0,70%
Financement industriel	943 656	746 307	7,22%
Autres financements / Produits exceptionnels	903 443	669 450	6,47%
TOTAL DES PRODUITS	9 928 434	10 338 675	100,00%
CHARGES			
Frais de personnel	7 520 332	7 608 490	73,59%
Formation et déplacements	473 803	528 042	5,11%
Partenaires externes	119 968	92 469	0,89%
Informatique : matériel et maintenance	237 787	249 146	2,41%
Frais administratifs	159 423	178 872	1,73%
Promotion et communication / 20 ^e anniversaire (2011)	241 641	77 007	0,75%
Loyer et charges	872 967	889 306	8,60%
Amortissement	290 622	281 777	2,72%
Charges exceptionnelles	40 000	421 340	4,08%
TOTAL DES CHARGES	9 956 543	10 326 449	99,88%
RÉSULTAT D'EXPLOITATION	-28 109	12 226	0,12%



SOURCES DE FINANCEMENT / CHARGES / COMMENTAIRES

Répartition des sources de financement



Commentaires sur les comptes 2012

En 2012, l'Ildiap continue sur sa lancée et boucle ses comptes sur un bénéfice de Fr. 12 226.- et une belle progression de 4,13% par rapport à l'exercice précédent. La barre mythique des 10 millions est franchie pour la première fois depuis la création de l'institut en 1991. La part des subventions publiques se situe toujours dans l'objectif stratégique avec, cette année, un pourcentage de 40,3%. Le canton du Valais renforce sa contribution financière de manière importante. Le nouveau montant alloué à l'Ildiap s'élève à Fr. 1 120 000.-. Par rapport à 2011, l'ensemble des subventions publiques croît globalement de 6,55%. Toutes les provisions en cours (Fr. 301 000.-) sont dissoutes. La contrepartie se retrouve dans les charges exceptionnelles.

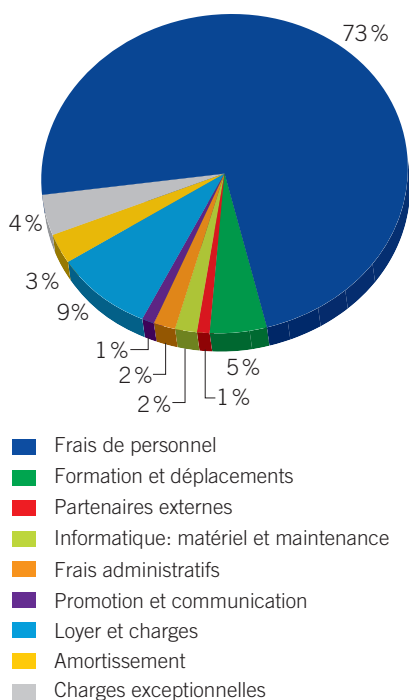
Subventions Confédération, Canton, Commune

(En milliers de francs suisses)

ANNÉE	2010	2011	2012	2013*
Confédération	1795	2357	2393	2428
Canton	900	900	1120	1220
Commune	600	650	650	700

* Budget

Répartition des charges



BILAN

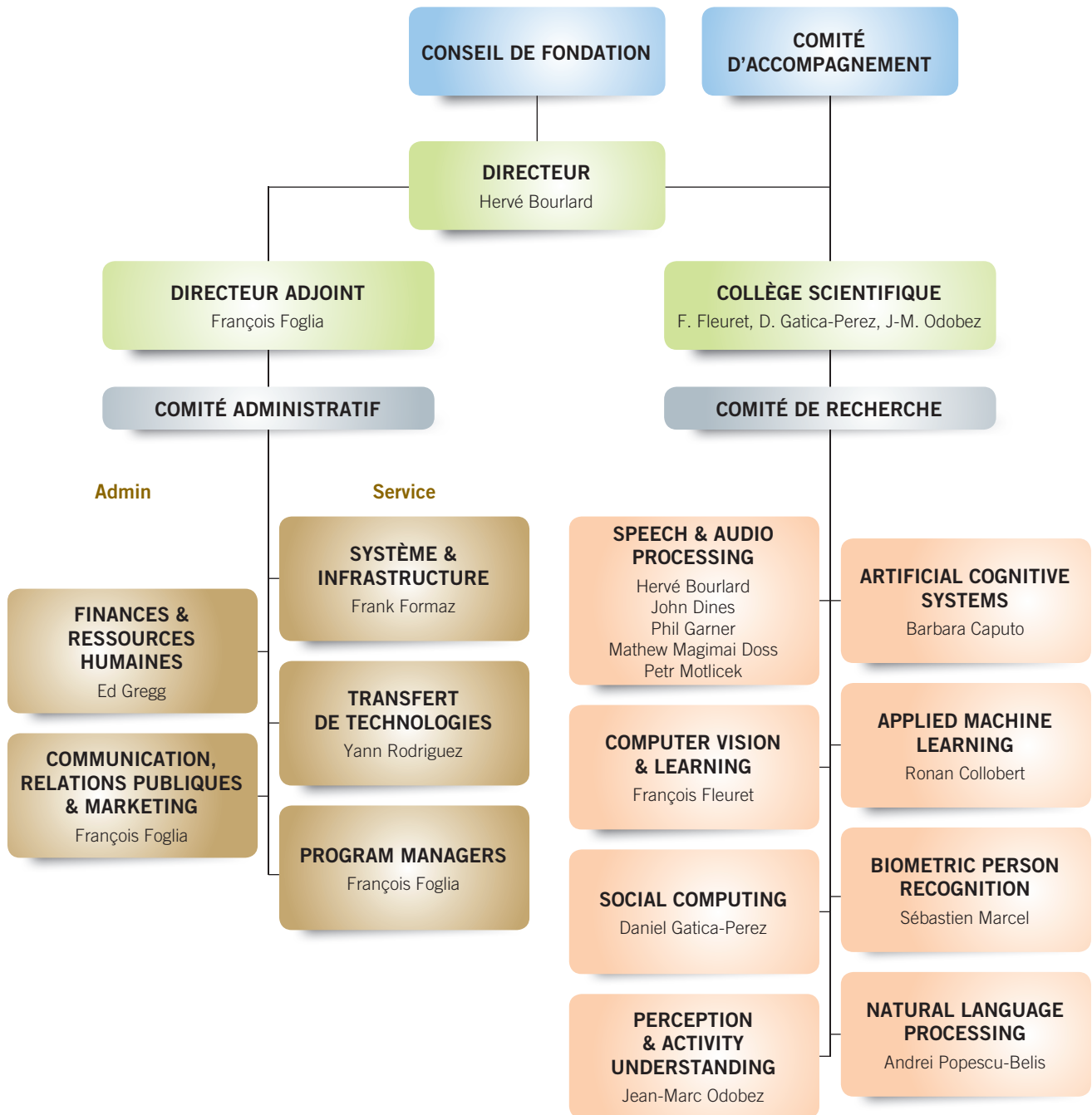
(En francs suisses)

	31.12.2011	31.12.2012
ACTIFS		
Liquidités	2 443 047,24	2 104 427,07
Débiteurs	499 887,55	287 382,35
Actifs de régularisation et divers	3 016 530,83	2 536 046,69
TOTAL ACTIFS CIRCULANTS	5 959 465,62	4 927 856,11
Mobilier et matériel informatique	523 402,89	504 960,99
Participations financières	10 000,00	10 000,00
TOTAL ACTIFS IMMOBILISÉS	533 402,89	514 960,99
TOTAL ACTIFS	6 492 868,51	5 442 817,10
PASSIFS		
Créanciers	461 687,65	365 838,69
Passifs de régularisation	4 583 667,75	3 918 238,84
Provisions	301 000,00	0,00
TOTAL FONDS ÉTRANGERS	5 346 355,40	4 284 077,53
Capital	40 000,00	40 000,00
Réserve spéciale	1 000 000,00	1 000 000,00
Résultat reporté	134 622,17	106 513,11
Résultat net	-28 109,06	12 226,46
TOTAL FONDS PROPRES	1 146 513,11	1 158 739,57
TOTAL PASSIFS	6 492 868,51	5 442 817,10



ORGANISATION







COLLABORATEURS

Personnel scientifique

Prénom, nom, fonction, origine, domicile, année d'arrivée à l'Idiap

Oya Aran Karakus, postdoctorante, Turquie, 2009
Afsaneh Asaei, doctorante, Iran, 2008
Chindansh Bhatt, postdoctorant, Inde, 2012
Joan Isaac Biel, doctorant, Espagne, 2008
Hervé Bourlard, directeur, Suisse, Saxon, 1996
Olivier Canévet, doctorant, France, 2012
Barbara Caputo, chercheuse senior, Italie, 2005
Ivana Chingovska, doctorante, Macédoine, 2011
Ronan Collobert, chercheur, France, 2010
John Dines, chercheur, Australie, 2003
Trinh-Minh-Tri Do, postdoctorant, Vietnam, 2009
Charles Dubout, doctorant, Suisse, Renens, 2009
Laurent El Shafey, doctorant, France, 2010
Rémi Emonet, postdoctorant, France, 2010
Nesli Erdogmus, postdoctorante, Turquie, 2012
Marc Ferras Font, postdoctorant, Espagne, 2011
François Fleuret, chercheur senior, France, 2007
Marco Fornoni, doctorant, Italie, 2010
Kenneth Funes Mora, doctorant, Costa Rica, 2011
Philip Garner, chercheur senior, Angleterre, 2007
Daniel Gatica-Perez, chercheur senior, Mexique, 2002
Paul Gay, doctorant, France, 2012
Arjan Gijsberts, postdoctorant, Pays-Bas, 2011
Manuel Günther, postdoctorant, Allemagne, 2012
Maryam Habibi, doctorant, Iran, 2011
Najeh Hajlaoui, postdoctorant, France, 2011
Alexandre Heili, doctorant, France, 2010
Pierre-Edouard Honnet, doctorant, France, 2012
David Imseng, doctorant, Suisse, Rarogne, 2009
Dinesh Babu Jayagopi, doctorant, Inde, 2007
Samuel Kim, postdoctorant, Corée du Sud, 2011
Ilja Kuzborskij, doctorante, Lituanie, 2012
Alexandros Lazaridis, Grèce, 2012
Rémi Lebret, doctorant, France, 2012
Leonidas Lefakis, doctorant, Grèce, 2010
Riwal Lefort, postdoctorant, France, 2011
Joël Legrand, doctorant, France, 2012
Adolfo Lopez-Mendez, postdoctorant, Espagne, 2012
Mathew Magimai Doss, chercheur, Inde, 2007
Sébastien Marcel, chercheur senior, France, 2000
Thomas Meyer, doctorant, Suisse, Martigny, 2010
Gelareh Mohammadi, doctorante, Iran, 2009
Petr Motlicek, chercheur, République tchèque, 2005
Laurent Nguyen, doctorant, Suisse, Lausanne, 2011
Jean-Marc Odobez, chercheur senior, France / Suisse, Clarens, 2001
Dimitri Palaz, doctorant, Suisse, Martigny, 2011
Nikolaos Pappas, doctorant, Grèce, 2012
Novi Patricia, doctorante, Indonésie, 2012
Andrei Popescu-Belis, chercheur senior, France / Roumanie, 2007
André Rabello Dos Anjos, postdoctorant, Brésil, 2010
Ramya Rasipuram, doctorante, Inde, 2010
Lakshmi Saheer, doctorante, Inde, 2008
Dairazalia Sanchez-Cortes, doctorante, Mexique, 2009
Darshan Santani, doctorant, Inde, 2012
Ashtosh Sapru, doctorant, Inde, 2011
Samira Sheikhi, doctorante, Iran, 2010
Nicolae Suditu, doctorant, Roumanie, 2008
Gyorgy Szaszak, postdoctorant, Hongrie, 2012
Mohammad Javad Taghizadeh, doctorant, Iran, 2011
Romain Tavenard, postdoctorant, France, 2012
Tatiana Tommasi, doctorante, Italie, 2008
Raphaël Ullmann, doctorant, Suisse, Lausanne, 2012
Alessandro Vinciarelli, chercheur senior, Italie, 1999
Majid Yazdani, doctorant, Iran, 2008
Sree Harsha Yella, doctorant, Inde, 2010

Ingénieurs de développement

Philip Abbet, ing. de développement senior, Suisse, Conthey, 2006
Olivier Bornet, ing. de développement senior, Suisse, Nendaz, 2004
Milos Cernak, ing. de développement senior, Slovaquie, 2011
Vasil Khalidov, ing. de développement senior, Russie, 2010
Salim Kayal, ing. de développement, Suisse, Vevey, 2011
Christine Marcel, ing. de développement, France, 2007
Florent Monay, ing. de développement, Suisse, Choëx, 2008
François Moulin, ing. de développement, Suisse, Vollèges, 2009
Alexandre Nanchen, ing. de développement, Suisse, Martigny, 2008
Flavio Tarsetti, ing. de développement, Suisse, Martigny, 2008

Personnel administratif

Prénom, nom, fonction, origine, domicile, année d'arrivée à l'Idiap

Céline Aymon Fournier, relations publiques, Suisse, Fully, 2004
Antoine Dorsaz, aide-comptable, Suisse, Fully, 2012
Christophe Ecoeur, junior program manager, Suisse, Collombey, 2010
François Foglia, directeur adjoint, Suisse, Saxon, 2006
Edward-Lee Gregg, responsable financier, Etats-Unis, 2004
Martina Guetl, junior program manager, Autriche, 2012
Léonore Miauton, senior program manager, Suisse, Chexbres, 2012
Sylvie Millius, assistante administrative, Suisse, Vétroz, 1996
Yann Rodriguez, transfert de technologies, Suisse, Martigny, 2006
Nadine Rousseau, assistante administrative, Suisse, Saxon, 1998

Ingénieurs système

Bastien Crettol, ingénieur système, Suisse, Sion, 2005
Norbert Crettol, ingénieur système, Suisse, Martigny, 2002
Cédric Dufour, ingénieur système, Suisse, Verbier, 2007
Frank Formaz, administrateur système, Suisse, Fully, 1998
Louis-Marie Plumel, ingénieur système, France, 2011
Vincent Spano, webmaster, Suisse, Martigny-Combe, 2004

Stagiaires

Prénom, nom, origine, institution d'origine

Les stagiaires de l'Idiap passent généralement entre trois et dix mois dans l'institut de recherche. Certains sont étudiants à l'EPFL et effectuent ce stage dans le cadre de leur travail de diplôme. D'autres arrivent dans le cadre de programmes d'échange d'étudiants mis en place dans les projets européens auxquels participe l'Idiap.

Aniruddha Adiga, Inde, Florida Tech University, Floride (USA)
Philippe Ballestraz, Suisse, HES-SO Valais, Sion
Okal Billy, Kenya, Jacobs University, Bremen (DEU)
Holger Caesar, Allemagne, Karlsruhe Institute of Technology (DEU)
Aleksandra Cerekovic, Croatie, University of Zagreb, Zagreb (HRV)
Thiyagarajan Chockalingam, Inde, Colorado State University (USA)
Tiago de Freitas Pereira, Brésil, Univ. de Campinas, Sao Paulo (BRA)
Michal Drozdal, Pologne, Autonomous University of Barcelona (ESP)
Manon Fournier, Suisse, Lycée-Collège de l'Abbaye, St-Maurice
Javier Galbally, Espagne, Autonomous University of Madrid (ESP)
Huseyn Gasimov, Azerbaïdjan, EPFL, Lausanne
Hu Hainan, Chine, EPFL, Lausanne
Mohsen Kaboli, Iran, KTH University, Stockholm (SWE)
Yunus-Emre Kara, Turquie, Bogazici University, Istanbul (TUR)
Alp Kindiroglu, Turquie, Bogazici University, Istanbul (TUR)
Ilya Kuzborskij, Lituanie, The University of Edinburgh (GRB)
Kexing Li, Chine, EPFL, Lausanne
Jukka Määttä, Finlande, University of Oulu (FIN)
Eric Malmi, Finlande, Aalto University (FIN)
Miranti Mandasari, Indonésie, Radboud University (NLD)
Alvaro Marcos, Espagne, University of Alcalá (ESP)
Braid Regula Meyer, Suisse, ISME, St-Gall
Peter Naoki, Suisse, EPFL, Lausanne
Pedro Henrique Oliveira Pinheiro, Brésil, EPFL, Lausanne
Jérémy Rappaz, Suisse, HES-SO Valais, Sion
Sandrine Revaz, Suisse, Université de Fribourg, Fribourg
Bin Shen, Chine, Purdue University, Indiana (USA)
Yang Sun, Chine, Radboud University (NLD)
Romain Tavenard, France, Université de Rennes I, IRISA-INRIA (FRA)
Lucia Teijeiro Mosquera, Espagne, University of Vigo (ESP)
Kelly Tiraboschi, Suisse, Lycée-Collège de l'Abbaye, St-Maurice
Aleksei Triastcyn, Russie, EPFL, Lausanne
Vagia Tsiminaki, Grèce, EPFL, Lausanne
Na Xingyu, Chine, Beijing Institute of Technology, Beijing (CHN)

Visiteurs

Prénom, nom, institution d'origine

Chercheurs ou industriels, les visiteurs ne passent que quelques jours ou quelques semaines à l'institut, les uns pour renforcer les liens interinstitutionnels, les autres pour prendre la mesure des travaux qui s'effectuent dans l'institut.

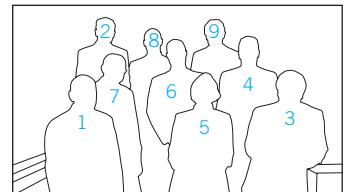
Elie El Khoury, Liban, Université de Toulouse, Toulouse, France
Filiberto Pla, Espagne, Jaume I University, Castelló de la Plana, Espagne
Adrian Penate-Sanchez, Espagne, Polytechnic University of Catalonia, Barcelone, Espagne



CONSEIL DE FONDATION



Le Conseil de fondation assume la responsabilité de la gestion économique et financière de l'institut de recherche, définit ses structures, nomme son directeur, et de manière plus générale veille au bon développement de la fondation en défendant ses intérêts.



- M. Olivier Dumas, Président**
Directeur d'Electricité d'Emosson SA
- M. Jean-Daniel Antille, Vice-président**
Responsable de l'Antenne Régions Valais romand
- Prof. Philippe Gillet, Vice-président**
Vice-président pour les affaires académiques
Ecole polytechnique fédérale de Lausanne (EPFL)
M. Marc-André Berclaz (*absent sur la photo*)
Président du comité directeur de la HES-SO Valais Wallis
- M. Stefan Bumann**
Chef du service de la formation tertiaire, Etat du Valais
- Mme Adrienne Corboud Fumagalli**
Vice-présidente pour l'innovation et la valorisation
Ecole polytechnique fédérale de Lausanne (EPFL)
- M. Marc-Henri Favre**
Président de la Ville de Martigny
- M. Patrick Furrer**
Vice-directeur d'Euresearch
M. Jean-René Germanier (*absent sur la photo*)
Conseiller national
- M. Jordi Montserrat**
Directeur Romandie de «venturelab»/«venture kick»
Prof. Christian Pellegrini (*absent sur la photo*)
Doyen du Conseil de fondation de l'Idiap
Professeur honoraire à la Faculté des sciences de l'Université de Genève
- M. Walter Steinlin**
Swisscom relations universitaires
Président de la Commission pour la technologie et l'innovation (CTI)

Trois nouveaux membres ont fait leur entrée dans le Conseil de fondation de l'Idiap. Trois personnalités qui, par leur fonction, leur expérience et leur vision, offrent à l'Idiap des perspectives nouvelles. Interviews express.

- 1) *Qu'est-ce qui vous relie à l'Idiap?*
- 2) *Quelle est votre vision de l'institut, de ses forces?*
- 3) *Comment voyez-vous l'Idiap de demain?*

«L'institut a opéré dès ses débuts des choix thématiques visionnaires.»



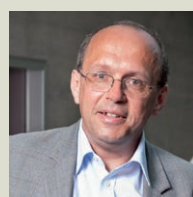
Marc-André Berclaz
Président du comité directeur de la HES-SO Valais Wallis, directeur opérationnel du Pôle EPFL Valais Wallis et vice-président de Swiss universities

1) La HES-SO est très engagée dans la recherche appliquée et le développement et, à ce titre, collabore régulièrement avec l'Idiap. Nous voulons contribuer à ce que ces relations se développent et se renforcent encore.

2) Développer un vrai pôle de compétence en Valais, sans université, ce n'était pas gagné d'avance! L'institut a opéré dès ses débuts des choix thématiques assez visionnaires pour l'époque. Il a su se faire une place très importante sur la scène scientifique internationale. C'est la preuve d'une belle énergie et d'une solide capacité entrepreneuriale.

3) Les mouvements actuels privilégient les grandes structures, c'est pourquoi l'Idiap doit aujourd'hui se développer sur des alliances, en particulier avec l'EPFL. Il lui faut aussi songer à assurer sa pérennité en préparant la relève.

«Je vois l'Idiap de demain comme un centre d'excellence européen, voire mondial.»



Patrick Furrer
Vice-directeur d'Euresearch et point de contact national pour les technologies de l'information et de la communication (ICT)

1) Nous entretenons depuis une dizaine d'années des liens étroits avec l'Idiap, que nous conseillons en matière de stratégie d'intégration dans les projets de l'Union européenne.

2) En vingt ans, le «petit institut» qu'était l'Idiap au départ a réalisé quelque chose de magique en Valais, en y croissant et en y créant des entreprises innovantes. Cette success story, qui réjouit le Valaisan d'origine que je suis, doit beaucoup au directeur de l'Idiap mais aussi au soutien de la Commune de Martigny, du Canton et de la Confédération.

3) L'institut possède des compétences uniques dans les domaines de la reconnaissance de la parole et de l'image, de la biométrie ou de la gestion de conférences. Je vois l'Idiap de demain comme un centre d'excellence européen, voire mondial. Le savoir-faire est là, il manque encore un peu de visibilité.

«Un bel institut qui sait se positionner.»



Jordi Montserrat
Directeur Romandie de «venturelab» et «venture kick»

1) La valorisation de la recherche est devenue aujourd'hui un axe fondamental et c'est dans ce cadre que nous collaborons depuis quelques années avec l'Idiap, en participant au financement et en accompagnant certaines de ses spin-off.

2) L'Idiap a montré sa capacité à gérer des projets de très haut niveau. Ses domaines de recherche, comme l'interaction homme-machine ou l'intelligence artificielle, sont très porteurs. J'aime les instituts de ce type, aux domaines de recherche bien précis, bien menés et qui se profilent au niveau international.

3) Le défi des prochaines années consiste à demeurer à la pointe. Dans cette optique, le rapprochement avec l'EPFL et la CTI (Commission pour la technologie et l'innovation) est essentiel. Il s'agit aussi d'accentuer la dimension de valorisation entrepreneuriale: davantage de spin-off et davantage de rayonnement international pour un bel institut qui sait se positionner!



COMITÉ D'ACCOMPAGNEMENT

Le comité d'accompagnement se compose de personnalités du monde scientifique choisies par la Direction de l'Idiap pour leurs compétences exceptionnelles et leur vision avant-gardiste. Bien que leur rôle soit strictement consultatif, leur soutien et leurs conseils sont fréquemment sollicités et se révèlent précieux dans la prise de décision en matière de recherche, de formation et de transfert de technologies.

Dr Jordan Cohen

Independent Consultant, Spelamode
Half Moon Bay, CA, USA

Prof. Dr Donald Geman

Professor of Mathematics, Johns Hopkins University
Baltimore, MD, USA

Dr John Makhoul

Chief Scientist, Speech and Signal Processing, BBN Technologies
Cambridge, MA, USA

Prof. Dr Nelson Morgan

Deputy director (and former director) of the International Computer Science Institute (ICSI)
Berkeley, CA, USA

Dr David Nahamoo

Senior Manager, Human Language Technologies, IBM Research
Yorktown Heights, NY, USA

Prof. Gerhard Sagerer

Rector, Bielefeld Universität
Germany

Dr Roelof van Zwol

Senior Research Scientist, Multimedia, Audience Sciences,
Yahoo! Research
Santa Clara, CA, USA

Prof. Dr Bayya Yegnanarayana

Professor and Microsoft Chair, International Institute of Information Technology (IIIT)
Hyderabad, India



PRINCIPAUX PARTENAIRES

VILLE DE MARTIGNY

CANTON DU VALAIS

CONFÉDÉRATION SUISSE

Secrétariat d'Etat à la formation, à la recherche et à l'innovation (SEFRI)



www.loterie.ch



www.swisscom.com



Swiss Power Group.

www.groupemutuel.ch



www.epfl.ch



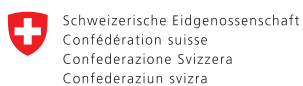
www.theark.ch



www.ideark.ch



www.snf.ch



Agence pour la promotion de l'innovation CTI

www.bbt.admin.ch/kti

HASLERSTIFTUNG

www.haslerstiftung.ch



ENCARTÉ SCIENTIFIQUE

IDIAP RESEARCH AREAS: HUMAN AND MEDIA COMPUTING

To face its continuous growth and diversification in different, related, research directions, while still fostering internal multi-disciplinary collaborations, Idiap reorganized its internal structuring of its research themes along the following dimensions. Idiap has thus changed/adapted the way it presents itself and describes its current activities, to take into account the new areas of development not only towards human-computer interaction but also toward human-to-human interaction, collaboration, behavior, and innovation. Thus, after several (13) years of positioning itself under the general theme of “Multimodal human-computer interaction”, Idiap decided to officially cover a larger research domain, now referred to as “**Human and Media Computing**”.

Articulated around our current activities, “Human and Media Computing” now covers the following research themes:

- **Perceptual and cognitive systems:** Speech processing; Natural language understanding and translation; Document and text processing; Vision and scene analysis; Multimodal processing; Cognitive sciences and Robotics.

Idiap combines its multi-disciplinary expertise to advance the understanding of human perceptual and cognitive systems, engaging in research on multiple aspects of human-computer interaction with computational artefacts such as natural language understanding and translation, document and text processing, vision and scene analysis (with a particular emphasis on human sensing), multimodal interaction, computational cognitive systems, robotics, and methods for automatically training such systems.

As evidenced in the present report, new major trends and developments at Idiap are now directed towards multilingual speech processing (multilingual speech recognition and synthesis), text-level machine translation and, possibly later, speech-to-speech translation.

Building on Idiap’s expertise in audio-visual processing, machine learning and robotics, there is also more and more emphasis on new activities in monitoring and autonomous robots able to build semantic spatial representations for indoor environments, updating continuously such representation taking advantage of incoming data as well as of prior experience from other robots. Furthermore, we are making significant contributions in the emerging area of advanced prostheses, leveraging on Idiap’s expertise in robotics, brain-computer interface and machine learning.

- **Social/human behavior:** Social media; Mobile media; Social interaction analysis; Social signal processing.

This area spans methods for analyzing human and social behavior from a variety of information sources with the goals of understanding social phenomena and developing human-centered applications. Idiap has investigated new approaches for behavioral analysis in face-to-face communication, online interaction in social media sites like YouTube, and smartphone-mediated interaction.

Continued projects include LS-CONTEXT (a long-term collaboration on smartphone sensing and data mining with Nokia Research that resulted in two Paper Awards in 2011); SONVB (a SNSF Sinergia interdisciplinary project featured in the Digital Life issue of the official SNSF Research Magazine); and VlogSense (funded by the NCCR IM2). New projects that started in 2011 include NISHA (in collaboration with NTT Communication Science Labs, Japan) and SOBE (a SNSF Ambizione Fellowship Grant). Idiap also worked together with the University of Geneva on an NCCR proposal towards creating a new type of Social Computer.

- **Information interfaces and presentation:** Multimedia information systems, User interfaces; System evaluation.

Information processing by computers must be accompanied by human-computer interfaces that present information and receive input in an efficient and usable way, possibly acquiring information from users in a non-disruptive way. Current research directions at Idiap focus on multimedia information systems, search and recommendation, and interactive information retrieval, and several new projects are currently being initiated in those directions.

- **Biometric person recognition:** Face recognition (detection-localization-identification-verification); Speaker identification and verification; Multimodal biometric person recognition; Countermeasures to spoofing attacks.

Biometric person recognition refers to the process of automatically recognizing a person using distinguishing behavioral patterns (gait, signature, keyboard typing, lip movement, hand-grip) or physiological traits (face, voice, iris, fingerprint, hand geometry, EEG, ECG, ear shape, body odor, body salinity, vascular).



Idiap is currently a recognized leader in that field, fully exploiting our multidisciplinary expertise in image processing, computer vision, pattern recognition and machine learning. After the end of the MOBIO EU project, this leadership Idiap–EPFL Joint Development Plan – 2011 Activity Report (PART 1) page 9/35 was confirmed in 2011 by the granting of new large EU projects, like the ICT project TABULA RASA and the SECURITY project BEAT. In 2011, Idiap and EPFL also worked together to initiate a joint NCCR proposal for a “Swiss Center for Biometric Security”.

- **Machine learning:** Statistical and neural network based machine learning; Computational efficiency, targeting real-time applications; Very large datasets; Online learning.

Research in machine learning aims at developing computer programs able to learn from examples. Instead of relying on a careful tuning of parameters by human experts, machine-learning techniques use statistical methods to directly estimate the optimal setting, which can hence have a complexity beyond what is achievable by human experts.

Today, Idiap is also recognized as a key leader in that field with new trends towards “collaborative machine learning”, deep neural network architectures, and large-scale distributed learning algorithms. Real-life applications include end-to-end spoken-term detection (see the DeepSTD project), Natural Language Processing with a particular interest in semantic analysis (SemTex project, in collaboration with NEC Research, USA) and large-scale distributed web search (SODS).



1. Speech and Audio Processing

Speech processing has been one of the mainstays of Idiap's research portfolio for many years. Today it is still the largest group within the institute, and Idiap continues to be recognised as a leading proponent in the field. The group has certain core technology, built around the connectionist approach to automatic speech recognition (ASR), and continues to contribute in that area. However, the expertise of the group has diverged to encompass hidden Markov model technology, text to speech synthesis (TTS) and generic audio processing. Much of the group's recognition has come from contributions to the community; these include the AMI corpus¹ and the Juicer ASR decoder². More recently, the MediaParl database has been released³. This highlights that, although the group has traditionally worked with English speech, many recent projects are multi-lingual in nature. This has in turn become a focus of the group; one that we are well placed to capitalise on given our geographical location.

Automatic Speech Recognition

In recent years, at Idiap the ASR research activities have been expanded from mono-lingual to cross-/multi-lingual processing. More specifically, in addition to focusing on "majority" languages other than English such as, French, German (funded through EU project D-Box⁴), Idiap is actively carrying research in several ASR directions, including:

- **Robust parametrization and acoustic modeling.**
We are still investigating new features (e.g., posterior-based features) and new acoustic models (new forms of hidden Markov models, such as KL-HMM, or artificial neural networks) that are more robust to noise and acoustic environments, as well as to speaker variability (e.g., accented speech). Over the last 12 months, we mainly focused on several novel acoustic modeling approaches that can effectively exploit acoustic and linguistic resources of a different language or multiple languages (funded through SNSF projects like ICS-2010⁵, and FlexASR⁶).

- **Cross-lingual and multi-lingual speech recognition.**
In addition to the above, and in the context of an Industry sponsored project (DAUM⁷), we are also investigating new fast acoustic model adaptation techniques (currently referred to as "subspace gaussian modeling") in cross-lingual and multi-lingual scenarios.
- **Swiss-languages.**
As part of a project with the Valais (and probably Vaud) Parliament (MediaParl project), we are developing state-of-the-art ASR systems for Swiss languages. We have made very good progress towards Swiss German and Swiss French speech recognition.
- **Grapheme-based ASR.**
State-of-the-art ASR systems typically use phones as the standard subword units. However, phones are not always well defined, and phonetic transcription of the lexicon words do not always match the actual pronunciation (especially in the case of accented speech). We are thus investigating the use of alternate subword unit representations such as graphemes (funded through the SNSF project FlexASR and HASLER Foundation project AddG2SU⁸).
- **Template-based ASR.**
Given the availability of very large databases, and the need to relax some of the HMM statistical assumptions (e.g., to better model the temporal dynamics of the speech signal), there is renewed interest in "template-based ASR" (also referred to as "episodic model"), replacing the training of stochastic models by simply storing full speech templates and performing recognition through large scale dynamic programming (dynamic time warping). In the context of the EU Marie-Curie project SCALE⁹ or SNSF projects like AMUSE¹⁰, we have recently shown that using posterior-based features and the appropriate local distance (based on KL-divergence) was yielding similar or better performance than full fletched HMM systems.

These research directions are particularly addressing practical challenges such as, non-native or accented speech recognition, rapid development of ASR systems for new or under-re-

1 <http://www.idiap.ch/mmm/corpora/ami>

2 <http://juicer.amiproject.org/juicer>

3 <http://www.idiap.ch/dataset/mediaparl>

4 <http://www.idiap.ch/scientific-research/projects/d-box-a-generic-dialog-box-for-multilingual-conversational-applications>

5 <http://www.idiap.ch/scientific-research/projects/interactive-cognitive-systems>

6 <http://www.idiap.ch/scientific-research/projects/flexible-grapheme-based-automatic-speech-recognition>

7 <http://www.idiap.ch/scientific-research/projects/domain-adaptation-using-sub-space-models>

8 <http://www.idiap.ch/scientific-research/projects/flexible-acoustic-data-driven-grapheme-to-subword-unit-conversion>

9 <http://www.idiap.ch/scientific-research/projects/speech-communication-with-adaptive-learning>

10 <http://www.idiap.ch/scientific-research/projects/adaptive-multilingual-ASR-TTS>



sourced languages and/or domains. In addition in continuous speech recognition, our ASR research also focuses on the development and improvement of state-of-the-art keyword spotting systems for audio mining. Keyword spotting (spoken term detection) is also addressed in the Applied machine Learning group, as discussed in Section 6.

Speech Synthesis

- Text-to-speech synthesis (TTS).

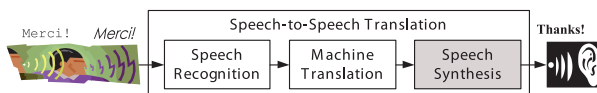
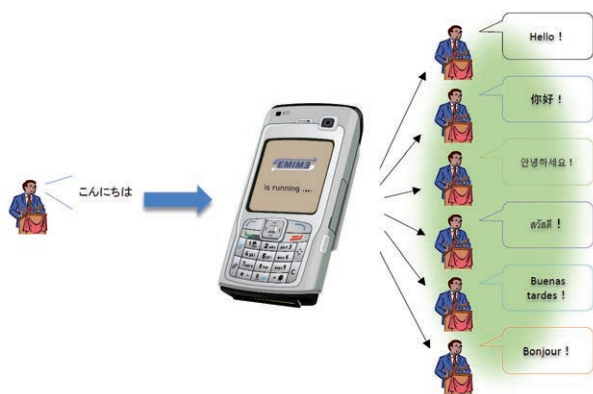


Figure 1 : Typical speech-to-speech translation system architecture developed at Idiap.

TTS is a relatively new venture for the speech group at Idiap, but two events recently marked its maturity: The first was the graduation of the first two PhD students in TTS; the second was the beginning of the SIWIS¹¹ project (Spoken Interaction With Interpretation in Switzerland). The TTS work at Idiap is largely aligned with the trend in the field towards statistical synthesis, which uses the same technology as ASR, placing us on a firm technical footing. Initial work was done under the EU FP7 EMIME¹² project, a project in speech to speech translation where the synthetic translated voices were adapted cross language to sound like the original speakers. This theme is continued in SIWIS which aims to do the same in Swiss languages, including the concept of accents and (to some extent) dialects. Another theme of SIWIS is adaptation of speech prosody, i.e., spoken emotion. This will ensure that, e.g., questions and stressed phrases remain as such after translation.

11 <http://www.idiap.ch/project/siwis>

12 <http://www.emime.org/>

SIWIS and EMIME highlight the requirement for a collaborative approach: Speech to speech translation requires a large infrastructure that is difficult to maintain at a single site. To achieve this, SIWIS brings together partners at ETHZ and the Universities of Geneva and Edinburgh.

- Speech coding. Another research area requiring TTS is speech coding, where very low bit rates can be achieved by using a speech recogniser followed by a synthesiser. Funded by armasuisse¹³, the technology lends itself to clear and secure communication. Innovative aspects draw on those of SIWIS: the system will be multi-lingual, and will model prosody to give acceptable latency; i.e., the synthetic speech will be produced as the speaker is talking word by word rather than sentence by sentence.

Audio Processing

Besides ASR and TTS, Idiap has been active in other areas of (pre-)processing generic audio signals, including:

- Sparse structured representations. Idiap continues in theoretical studies of sparse structured representations of audio signals, more specifically on model-based sparse component analysis, with different practical applications, such as source separation and localisation based on a small number of microphones, room acoustic modelling, and new beamforming techniques to improve recognition of speech recorded from far-field microphones. This work is mainly carried out in the context of the EU Marie-Curie project SCALE¹⁴.
- Microphone arrays. Idiap further performs studies on audio recorded from microphone arrays, especially those organized in “ad-hoc” topologies. Recently, we mainly focused on the automatic calibration of arbitrary microphone array geometries (using techniques like matrix completion and multi-dimensional scaling) in order to perform localization required for high quality data acquisition. In this context, we also consider environmental ambiguities such as reverberation and overlapping speech, thus requiring multi-source localization using ad-hoc microphones in reverberant rooms. This work is currently pursued in the context of the NCCR IM2¹⁵.

13 <http://www.ar.admin.ch/internet/armasuisse/en/home.html>

14 <http://www.idiap.ch/scientific-research/projects/speech-communication-with-adaptive-learning> and <http://www.scale.uni-saarland.de/cms/>

15 <http://www.im2.ch>



- **Speaker diarization.**

In speaker diarization — the task of annotating temporal regions of audio recordings with labels indicating “who spoke when” — Idiap focuses on processing of spontaneous overlapping speech through the use of multiple sources of information investigated in the context of various projects, including the EU FP7 project inEvent¹⁶ and the SNSF project RODI¹⁷.

- **Audio indexing**

In the context of the Hasler Foundation project SESAME¹⁸, Idiap aims at advancing the state-of-the-art in speech processing and spoken language understanding for automatically structuring data from spoken cultural heritage archives. Technical challenges include speaker segmentation and clustering, role recognition as well as topic/story segmentation and tracking over very heterogeneous

- **Very low bit-rate speech coding.**

Recently, Idiap has become active in the domain of very low bit-rate speech coding through a collaboration with armasuisse, in the context of the RECOD project¹⁹. The technique exploits the opportunity of having technology in both ASR and TTS, allowing information to be transmitted at the word or phoneme level rather than acoustic level, whilst still being perceived as speech.

- **Perceptual background noise analysis for telecommunication systems.**

Today, communication noises in mobile networks are often suppressed to improve speech intelligibility. But newer speech network generations can now transmit much higher quality signals, hence the potential to better exploit the information about a talker’s environmental context. New international ITU standards will soon be defined along that trend. In the CTI project PANDA²⁰, and in collaboration with SwissQual AG, we aim at exploiting leading edge speech processing technique to automatically extract and perceptually classify environmental noises towards those new standards.

2. Computer Vision and Learning

The main objective of the Computer Vision and Learning group is the development of novel machine-learning techniques, with a strong emphasis on their algorithmic efficiency, and applications mostly in the processing of images and video streams.

Our current research follows two main axes. The first is the study of novel generic learning methods to cope with large dimension feature spaces or large training sets. The second is the development of new algorithms for scene understanding, which involves practical problems such as object detection in natural scenes, or pedestrian tracking with a multi-camera setup.

Generic Learning

- **Learning in large dimension feature space.**

In the context of the EU FP7 MASH project²¹, we are investigating the learning of classifiers in large features space, ideally to cope with hundreds of families of feature extractors, each of tens of thousands of dimensions.

We have proposed two variants of AdaBoost to cope with these difficulties. The first one consists of sampling a few features from each family before the learning starts, and to use this features to estimate at every Boosting step the most promising feature family, so that we can bias the sampling accordingly. The second one models the loss reduction as a function of the number of features looked at, and the number of samples used to estimate edges. This model allows to optimize the trade-off between the two.

- **Efficient sample sampling and Reservoir learning.**

To improve the extraction and the use of samples in the context of large scale learning, two issues are at hand: The extraction of novel samples, usually incurring a cost linear with the number of samples, and the processing of the said samples, also of linear cost.

The SNSF DASH project addresses the former. It can be seen as a form of unsupervised active learning: identify good “sources” of samples (web sites, data-bases, videos) according to their ability to improve a predictor, and focus the extraction of additional training data from there, leading to a sub-linear cost. The latter is addressed currently in the context of the MASH project, in which we are developing novel techniques between batch and on-line learning by keeping a limited amount of samples in a *reservoir*, that optimizes the reduction of the loss on the full sample population.

- **High-dimension density similarity measures.**

We are investigating the clustering and classification of distributions of points in high-dimension spaces. To leverage the remarkable ability of random forests to cope with such

16 <http://www.inevent-project.eu/>

17 <http://www.idiap.ch/scientific-research/projects/role-based-speaker-diarization>

18 <http://www.idiap.ch/scientific-research/projects/searching-swiss-audio-memories>

19 <http://www.idiap.ch/project/armasuisse/recod/index.html>

20 <http://www.idiap.ch/scientific-research/projects/perceptual-background-noise-analysis-for-the-newest-generation-of-telecommunication-systems-1>

21 <http://mash-project.eu>



situations, we have proposed a novel metric, which looks at the Entropy of the empirical distribution of a group of samples in the leaves of a forest build for another group of points.

If the two groups follow the same distribution, this entropy should be similar to that of the initial group, hence maximum. We applied this metric to the study of the growth dynamic of neurons, and its modulation by their genotypes in the framework of a collaboration with the University of Geneva, the University of Basel, and EPFL.

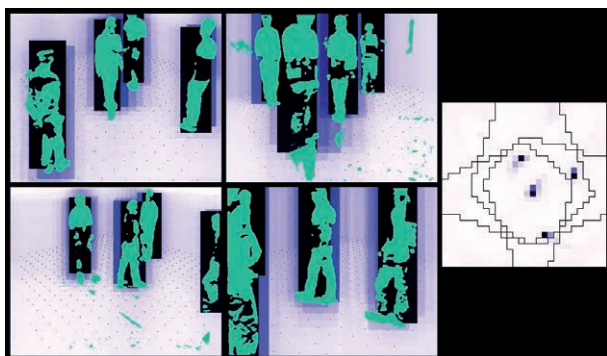


Figure 2 : Our fast version of the Deformable Part Model object detector (left) is close to one order of magnitude faster than other existing implementations. The Probabilistic Occupancy Map (right) estimates probabilities of presence on the ground, given the result of a background subtraction procedure in individual views.

Scene understanding

- Fast object detection.

In the SNSF project VELASH, we are investigating the use of very large feature spaces for image analysis, and in particular the speed-up of state-of-the-art object detectors based on the use of linear operators at multiple scales and locations. To speed-up this process, we proposed a careful organization of the computation to allow the use of the classical Fast Fourier Transform. Our implementation of the image parsing is close to one order of magnitude faster than the best pre-existing algorithms, and available under the open-source GPL3 license.

We have also developed a Boosting variant to train jointly a cascade of classifiers for detection. This procedure relies on a stochastic interpretation of the classifier responses. The resulting algorithm pushes all the classifiers to respond properly on the positive samples, and pushes the classifiers which are “already good” to get even better on each negative samples.

- Playground learning with pose-indexed features.

Related to our work on the stationary features, we have started a project in collaboration with the Universitat Politècnica de Catalunya, whose underlying philosophy is to leverage high-quality multi-million pixels training image, to build models able to cope with extremely degraded test images of a few hundreds pixels. The key idea is to use the high-quality signal to build strong priors, namely 3-D models, and to design pose-indexed features automatically from these models. Ultimately, we analytically define measurements in the images, parametrized by the pose of the object we want to check, and train a predictor from these measurements. In practice, instead of dealing only with scale and location at the algorithmic level, we let the system similarly cope with any out-of-plan rotation, by making the image measurements follow the pixel displacements.

- Multi-camera tracking.

We have a long-lasting collaboration with the CVLab at EPFL around the design of a robust multi-camera tracking system. It is composed of the Probabilistic Occupancy Map, and a flow-based convex multi-target tracker. The former estimates in individual time frames marginal probabilities of presence of people by minimizing the Kullback-Leibler divergence between the “true” posterior under a joint generative model, and a product law. The latter connects these detections properly by optimizing a flow in a spatio-temporal graph. Both components are extremely efficient computationally and process in real-time multiple camera streams to produce state-of-the-art results. Open-source implementations of both components are available under the GPL3 license.

We used the same flow-based tracking procedure to minimize the exponential loss by optimizing the locations of training examples, alternatively with a standard AdaBoost to train an image-based model for detection. The resulting FlowBoost procedure reduces by more than an order of magnitude the required amount of labelled examples, and was used with success for pedestrian tracking in monocular sequences, and neuron tracking in microscopic videos.



3. Social Computing

Social computing is an interdisciplinary domain that integrates theory and models from mobile and ubiquitous computing, multimedia, machine learning, and social sciences to analyze and interpret human and social behavior in everyday life, and to create devices and systems that support interaction and communication. The Social Computing group in 2012 was composed of 1 head of group, 2 postdoctoral researchers, 1 SNSF Ambizione Research Fellow, 5 PhD students, and 8 visiting researchers (1 faculty-on-sabbatical, 1 postdoc, and 6 MS/PhD students). Research lines that were investigated in 2012 included analysis of ubiquitous face-to-face interaction, behavioral analysis of online conversational video, and urban data mining using smartphones and location-based social networks.

Face-to-face interaction

Collaborating with social psychologists, we are studying computational models to analyze dyadic and small-group interactions in workplace settings using a variety of sensors (cameras, Kinect devices, microphone arrays, smartphones, and wearables), and automatically infer key social variables for psychologists and recruiters. In the dyadic case, in the context of organizational behavior situations (job interviews and job implementation), we have investigated connections between automatically measured audio and visual nonverbal cues and stress, hireability, and performance.



Figure 3 : Typical scenario of the NISHA project, an NTT-Idiap Social Behavior Analysis Initiative funded by NTT Communication Science Laboratories, Japan.

In the small-group case, we are studying links between behavioral cues and variables like emergent leadership and personality traits, using collective classification and transfer learning approaches. This research was conducted in the framework of two SNSF projects (SONVB²²: Sensing and Analyzing Or-

²² <http://www.idiap.ch/project/sonvb>

ganizational Behavior; SOBE²³, an Ambizione Research Fellowship) and the NISHA²⁴ project (NTT-Idiap Social Behavior Analysis Initiative) funded by NTT Communication Science Laboratories, Japan. One of our works on small-group analysis was awarded an Outstanding Paper Award at the ACM Int. Conf. on Multimodal Interaction 2012. Our research in this domain has also been featured in New Scientist, ACM Tech-News, and the SNSF Horizons Research Magazine.

Online conversational video

The work in this line develops around VlogSense, an approach to characterize video bloggers (vloggers) in social media sites like YouTube through the automatic extraction of behavioral features including prosody, body motion, and facial expressions, which enables high-level social analysis (Figure 4). More specifically, we are examining the problem of interpersonal perception in vlogging, i.e., how vloggers are perceived by their online audiences. We have shown several significant associations between automatically extracted nonverbal and verbal cues and first impressions of personality traits, attractiveness, and mood.

As a new element, video crowdsourcing with platforms like Amazon Mechanical Turk has been used to collect interpersonal judgments; this is a powerful technique that can scale up to large amounts of data while relying on demographically diverse viewers. This research was funded by the SNSF National Center of Competence in Research on Interactive Multimodal Information Management (NCCR IM2 through the VlogSense project (Modeling Conversational Behavior in Social Media).

Urban data mining

Our work in this domain spans two lines: large-scale smartphone sensing and location-based social networks (Figure 4). For the first line, jointly with Nokia Research, we led the Mobile Data Challenge 2012, an initiative in which data from the Lausanne Data Collection Campaign was released for research purposes to the mobile computing community.

The Lausanne data is perhaps the largest mobile sensing academic effort in Europe to date, involving close to 200 volunteers in French-speaking Switzerland and smartphone data for up to 18 months. The Challenge attracted participation from several hundred researchers around the world. Regarding modeling, our work focused on automatic place recognition and probabilistic location prediction using a variety of location sensor data (GPS, Wifi, etc).

Regarding location-based social networks, supported by the

²³ <http://www.idiap.ch/scientific-research/projects/multimodal-computational-modeling-of-nonverbal-social-behavior-in-face-to-face-interaction-1>

²⁴ <http://www.idiap.ch/project/nisha/>



SNSF HAI²⁵ project (Human Activity and Interactivity), we are investigating the use of FourSquare data for two goals: the development of transfer learning techniques to boost smartphone-based location prediction models, and the characterization of phenomena at large geographic scales like cities, including characterization of urban places according to language usage, and sampling effects in location-based social networks.

Our work in this domain received three Conference Paper Awards between 2010 and 2012, and was featured in 2012 in New Scientist, PhysOrg, and Les Echos.

4. Perception and Activity Understanding

The Perception and Activity Understanding group conducts research in human activities analysis from multi-modal data. This entails the investigation of fundamental tasks like the representation, detection, segmentation and tracking of people, the characterization of their state, and the modeling of sequential data and their interpretation in forms of gestures, activities, behavior or social relations, through the design of principled algorithms extending models from computer vision, statistical learning, or multimodal signal processing. Some of the main research carried out this year are described below.

Tracking and person representation

In the context of the EU FP7 VANAHEIM project, we have continued our investigation towards enhanced behavior recognition in surveillance context. Different research directions have been investigated this year on this topic. For the human tracking task, we have proposed several approaches such as using the tracking context (crowd density) to adapt the association parameters of human detections, or learn them from intermediate tracklet results, leading to state-of-the-art results on PETS benchmark data and on the Torino metro project data. On the learning side, we have proposed a novel algorithm designed for the joint adaptation of several classifiers whose output are coupled. It was applied to the joint estimation of head and body orientations, which are coupled by physical constraints and are also related to the motion direction when people are walking. Samples results are shown in the top of Fig. 5.

Non-verbal behavior extraction

Our long-standing work on human interaction modelling has been continued in the context of several projects including the EU FP7 HUMAVIPS project, and the TRACOME and the SONVB funded by SNSF. HUMAVIPS seeks to endow humanoid robots with basic social skills necessary to deal with small

groups of people. In this project, we have improved our multi-person realtime head tracking and pose estimation algorithm to handle low quality sensors, sensor motion (when the robot nods, turns the head etc) and variable situations encountered in the Human-Robot interface (HRI) context, and are working on attention modeling by studying different (non-linear, switching) gaze models relating sequences of Visual Focus of Attention to head pose sequences. Research in the project focused on the maintaining of persons identity over time and use of stereo information for the tracking part, and on the use of dialog sensitive contextual information for attention modeling.

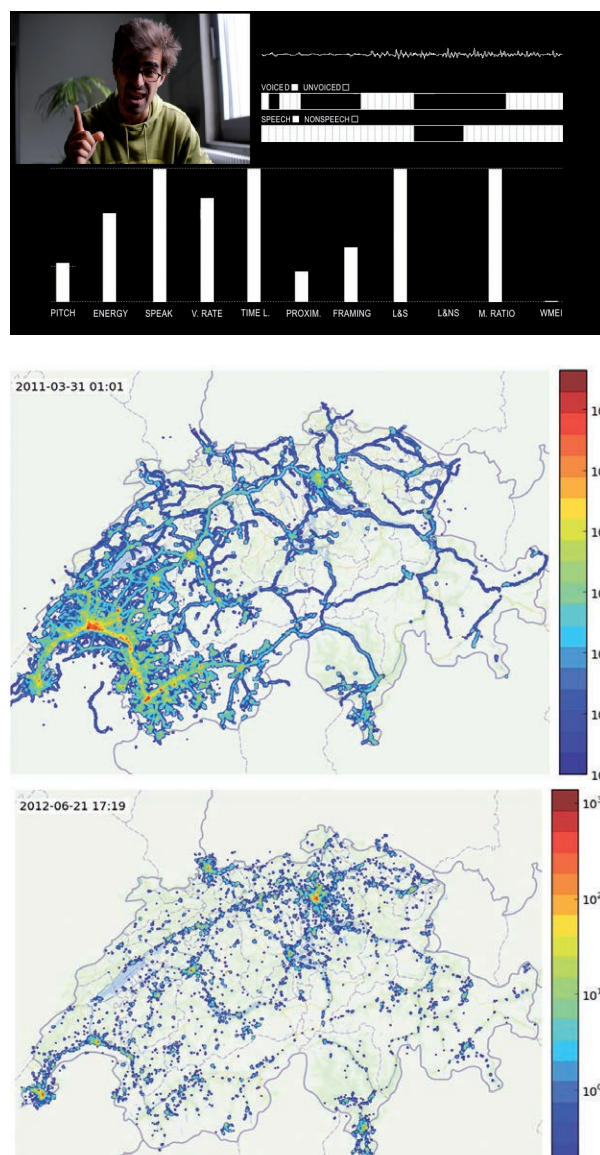


Figure 4 : Top: The VlogSense concept: Online vloggers are characterized by a variety of nonverbal behavioral features. Center: Heatmap of the Lausanne smartphone location data in Switzerland. Bottom: Heatmap of check-ins in Switzerland using six months of Foursquare data.

25 <http://www.idiap.ch/scientific-research/projects/human-activity-and-interactivity-modeling>



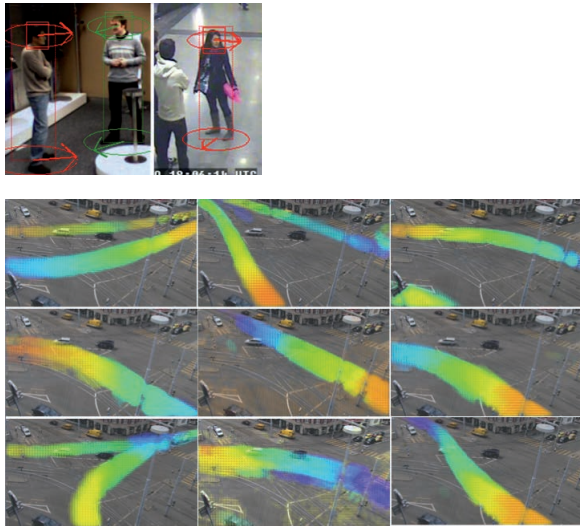


Figure 5 : Top: sample results of head and body orientation estimation. Bottom: automatic activity discovery from traffic videos. Each image shows an extracted activity motif: a sequence of observations that regularly co-occur over time. The color gradient indicates the timing since the start of the motif (from blue to red).

In SNSF project TRACOME we have investigated multimodal gaze tracking using depth sensor data. Indeed, the advent of cheap depth cameras like Kinect changed the HRI research landscape in recent years, and we are exploring the use of such device for non-verbal behavior extraction. This year, we proposed a framework to estimate gaze directions under free head motion. The depth information was exploited to accurately track the head pose using 3D deformable models, and render the face and eyes as if they were frontal, while the RGB image of the eye/pupil is used to estimate the gaze direction, as illustrated in Fig. 6. Despite the relatively low resolution of the image, errors from 10 to 15 degrees under complex head pose were achieved. The work received the best student paper award at the CVPR Kinect Gesture challenge workshop. Finally, in the context of the SONVB project that investigates new analysis approaches for automatic social inference in face-to-face organizational scenarios, we have worked with the social computing group on head gesture recognition in natural conversations, developing a multimodal nodding detector that leverages on the audio- based self-context making use of findings in psychology establishing that the dynamics of head gestures are conditioned on whether people are speaking or not.



Figure N°6 : Top: 3D rendered mesh from depth and RGB image. Left top: using the 3D tracker output, eye regions from the RGB image can be cropped and rendered as if the head pose was frontal. Bottom: sample of recognized gaze under free head movements. Green lines materialize the ground-truth gaze directions, while the red ones represent estimated gaze.

Activity analysis

In the context of the EU FP7 VANAHEIM project and SNSF HAI project, we have continued our efforts in the design of mining algorithms for the unsupervised discovery of recurrent activity pattern in multivariate time series, where observed values are caused by the superposition of multiple phenomena that can occur concurrently and with no synchronization (see 5). This is a typical situation where multiple sensor record the activities of multiple objects/people, like multimodal sensors (proximity, water, light sensors, etc.) in home automation (domotics) applications. This year, we proposed a graphical model to discover, from such multi-dimensional event time series, both global states and local rules that governs the occurrence of events. It was demonstrated in traffic scenes where global states accounts for traffic signal state (phase of a traffic cycle) while local rules could account for right of way or series of street segments a car can follow. These models can be used for abnormality detection and with ethologists from University of Vienna, we have investigated the selection of video streams to be displayed in control rooms of large public spaces and direct the attention of operators towards interesting information. The comparison of the the the rankings of video streams according to the 'unusualness' or interestingness of their content made by the humans with those resulting from the automatic abnormality rating validate the approach. Finally, in the context of the SNSF project PROMOVAR project funded by SNF,



we are exploring the use of these pattern discovery techniques for the classification of action from video inputs. The difficulty there lies in the exploitation of labeled information in the learning of these patterns for improving class discrimination.

5. Artificial Cognitive Systems

The Artificial Cognitive Systems group works on the development of multi-modal learning algorithms for enabling artificial agents to act autonomously in realistic setting, with a special emphasis on the ability to autonomously detect knowledge gaps and fill them autonomously with open ended learning strategies. The focus of the work is on designing algorithms that are principled, computationally efficient and that provides robust performance in very realistic settings while having at the same time theoretical guarantees on the expected behavior. Over the last years, the research activity has focused on two key projects: semantic spatial understanding of indoor places and advanced hand prostheses.

Semantic Spatial Understanding of Indoor Places

The work on semantic spatial understanding of indoor places, started in 2005, has been continued in the context of several projects, including the and ICS projects, funded by SNSF. The VISION@HOME project aims at providing models and methods to detect, recognise and categorise the 3D shape of everyday objects and their affordances in homes. It proposes the innovative Situated Vision paradigm, developing 3D visual perception capabilities from the view of the robot, its task and the environment it operates in.

The Situated Vision approach is inspired by recent work in cognitive science, neuroscience and interdisciplinary work in EU projects: it fuses qualitative and quantitative cues to extract and group 3D shape elements and relate them to affordance categories. Task-oriented and 3D attention mechanisms let the robot execute primitive actions to exploit the affordances perceived. Perception integrates quantitative and qualitative shape information from multiple 2D and 3D measurements. The analysis of the shapes is used to find instances of semantic 3D concepts, such as providing support to objects, enclosing space, etc. that can be used to those spatial concepts to find semantic entities, such as table surfaces, cupboards, closets, drawers and to learn which perceived affordances belong to which object category. Complementary to this, the main goal of the SNSF ICS project is to develop a system able to visually learn semantic concepts that characterize rooms and indoor environment, such as names referring to the activities normally performed in them (the fitness room) and the objects they contain (the bedroom). This would make it

possible then to take advantage of such knowledge also in working scenarios which differ from the original ones. Taking inspiration from biological models of human perception, we have identified two main components for the representation of indoor scenes: (1) a description of the global appearance of the image in term of image features, and (2) a description of the local landmarks present in some regions of the image. From a computational point of view the two representations could be regarded as a global appearance description of the scene, for example by means of statistics of visual features, and as a statistical representation of the co-occurrence of local concepts and scene categories. The design and integration of computational models of these two perceptual components, suitable for indoor place categorization, constitutes the core of our research.

Over the last year, we casted the semantic spatial modeling problem into that of learning from multiple cues. Our contributions have been a principled online Multi Kernel Learning algorithm able to combine optimally multiple features while providing theoretical guarantees on the expected performance, and a global feature representation encoding at the same time task-driven and data driven spatial information. The combination of these two contributions has led us to obtain the state of the art in the field, as measured on reference benchmark databases. We also developed a new online transfer learning algorithm for leveraging over prior semantic spatial models in a dynamic, open-ended fashion, that allows to modulate the contributions of different prior sources in a principled manner.

Advanced Hand Prostheses

The work on advanced hand prostheses has been initiated in 2011 with the SNSF funded NINAPRO²⁶ project. Its goal is to develop a family of algorithms able to significantly augment the dexterity, and reduce the training time, for sEMG controlled prosthesis. Indeed, the state of the art in hand prosthetics does not offer more than 2-3 degrees of freedom and a very coarse control of the force, as there is no haptic feedback. Patients interface with the prosthesis via surface electromyography (sEMG), recorded using surface electrodes. Learning how to control the device through many input sEMG channels is a long and difficult process for most patients, that therefore settles for limited and very simplified movements (open/close). This contrasts with recent advances in mechatronics, thanks to which mechanical hands gifted with many degrees-of-freedom and force control are being built. There is a need for prosthetic hands able to naturally reproduce a wide amount of movements and forces, while at the same time requiring

²⁶ <http://www.idiap.ch/project/ninapro/>



a lower effort in learning how to control hand postures. This goes beyond mechatronic dexterity: the real challenge is how to provide patients with a cheap, easy and natural way of controlling the prosthesis. By testing its findings on a very large collection of data, NINAPRO has the ambition to pave the way for a new generation of prosthetic hands. The work is organized along four themes: data acquisition and analysis (theme 1), augmented dexterity: posture classification (theme 2), augmented dexterity: natural control (theme 3) and adaptive learning (theme 4).

Dr. B. Caputo is the coordinator of the project, and the principal investigator in the themes 2 and 4. The objective of theme 2 (augmented dexterity) is to push the current state of the art in prosthetic hand posture classification from handling a maximum of 12 postures up to 40-50. Over the last year, we have conducted a thorough benchmark evaluation using the most popular and effective feature descriptors proposed so far in the literature, used as input to the current off the shelf state of the art classifiers available. The evaluation was conducted on the NINAPRO database, that for the number of subjects and hand postures it contains, represents a very significant advance with respect to any previous study of this type. Our results show clearly that existing solutions that worked well on more limited settings are not able to provide the accuracy required by amputated subjects. At the same time, combining different features with advanced multi cue methods seem to obtain promising results, especially when the cues are acquired by different modal sensors.

The goal of theme 4 (adaptive learning) is to develop learning algorithms to better interpret the sEMG signals acquired from the patients, with the ultimate goal of boosting the learning process necessary for the patient to effectively use the prosthesis. Our plan is to build pre-trained models of various data postures, on the data acquired in theme 1, and adapt these general models to the needs of individual users as new data will become available using adaptive online learning methods. Over the last year, we pursue this vision in the large-margin classifiers framework, developing a transfer learning algorithm across multiple subjects that assumes that all prior models and the new models to be learned by the new subject all contain the same number and type of postures. In such conditions, it is possible to show that leveraging over priors significantly boost performance with a highly reduced number of repetitions by the new user. This work has been accepted for publication in the international journal *IEEE Transaction of Robotics*.

6. Applied Machine Learning

Our group is interested in computer algorithms which can *learn* a behavior to achieve a given task of interest, in contrast to algorithms whose behavior is constrained by hand-crafted rules. Our research is driven from real-world applications involving a large amount of data. Domains of interest include natural language processing, computer vision, and audio processing fields. A particular emphasis is given to generic machine learning tools which require minimum a priori knowledge on the data (such as deep learning techniques), as well as unsupervised learning techniques which can leverage inherent semantic from large-scale structured data.

Deep Learning

Many state-of-the-art Natural Language Processing, Computer Vision or Audio Processing systems address a task by applying simple statistical models to carefully hand-crafted features. In other words, *the researchers themselves discover intermediate data representations* relevant for the task of interest, by engineering *task-specific* features. In contrast, deep learning algorithms aim at *learning* several layers of features representing raw data, with an increasing level of abstraction. Our group investigates both new generic deep learning algorithms, as well as algorithms specific to the following applications:

- **Natural Language Processing (NLP).**

In an ongoing collaboration with NEC Laboratories America²⁷, we investigate new generic machine learning algorithms for various NLP tasks. We have a particular interest in *learning* generic semantic word representations, chunk of word representations, or sentence representations trained on *large unlabeled text corpora*. Recently, these semantic representations allowed us to design a *single architecture* which outputs tags for several NLP tasks ranging from syntax analysis (Part of Speech tagging, chunking, syntactic parsing) to semantic analysis (Name Entity Recognition, Semantic Role Labeling). Our architecture is state-of-the-art both in accuracy and speed performance.

- **Scene Parsing.**

Scene parsing is a technique that consist on giving a label to all pixels in an image according to the class they belong to. Most systems rely on costly graphical models to take a global decision from local classifiers. We currently investi-

²⁷ <http://ml.nec-labs.com>



gate new efficient end-to-end systems based on recurrent convolutional neural networks (see Figure 7), alleviating the need of any hand-crafted feature and allowing the modeling of complex spatial dependencies with a low inference cost

- **Spoken Term Detection.**

The DeepSTD project, funded by the Swiss Hasler Stiftung²⁸, is concerned about applying new deep learning techniques to audio processing, with a particular interest in fast spoken term detection. The year 2012 was devoted to estimating phoneme class conditional probabilities directly from raw speech signal, using convolutional neural networks.



Figure 7: Our recurrent scene parser over original image (a) segments the image as shown in (b). Due to its recurrent nature, it can be fed again with its own prediction (b) and the original image (a), which leads to (c): most mistakes are corrected.

Distributed Semantic Search

Centralized search raises several issues, including ethical problems, inefficiencies due to the need of “copying” the web, and poor scaling capabilities. In the Hasler Foundation project SODS²⁹, we wish to develop a new type of search engine distributed over available web servers (in the same spirit than YaCy³⁰), in contrast to most existing search engines centralized at a single company site. Our web search engine will leverage state-of-the-art Natural Language Processing techniques, as well as distributed machine learning algorithms, in order to organize web pages, and guarantying an efficient search.

28 <http://www.haslerstiftung.ch>

29 <http://www.haslerstiftung.ch>. The SODS project is part of the SmartWorld special call.

30 <http://yacy.de>

Technology Transfer

Our group puts a strong emphasis on the reproducibility of all our research findings, such that our tools can be integrated easily in more complex systems or products. Tools we developed include:

- **Face Technologies.**

This year we started a collaboration with the KeyLemon³¹ startup, through the CLEAR CTI project. We are in charge of developing state-of-the-art algorithms for multi-pose face detection, head pose detection, facial feature detection, and gender detection. In 2012, we delivered a first version of our multi-pose face and gender detection system (see Figure 8), which runs in real time with good performance even in adverse conditions.

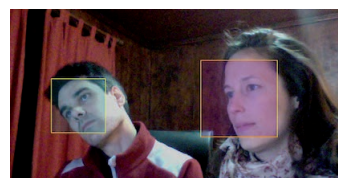


Figure 8: Our real time face detector detects multi-pose faces and labels gender (pink for female, bluish for male) in adverse conditions.

- **SENNA, a Natural Language Processing (NLP) tagger.**

SENNA³² is a standalone software distributed under a non-commercial license, that we develop through our ongoing collaboration with NEC Laboratories America³³. It outputs a host of Natural Language Processing predictions, at blazing speed while holding state-of-the-art accuracy performance. Predictions include part-of-speech tags, chunking, name entity recognition, semantic role labeling and syntactic parsing.

- **TORCH, an Open-Source Machine Learning Platform.**

We are maintaining Torch³⁴ an open-source (BSD license) fast and scalable Matlab-like system, which serve us as an efficient platform to implement all our machine learning algorithms. Torch leverages Luajit, an extremely fast scripting language, which allows us to quickly develop complex demos and prototypes. Torch is widely spread, present in many international academic and private institutions. Thanks to its open-source nature, external contributors wrote various packages for a large number of application domains.

31 <https://www.keylemon.com>

32 <http://ml.nec-labs.com/senna>

33 <http://ml.nec-labs.com>

34 <http://www.torch.ch>



7. Biometric Person Recognition

The Biometric Person Recognition group develops novel algorithms for face recognition (2D, 3D and Near-Infrared), speaker recognition, anti-spoofing (presentation attack detection) and emerging biometric modes (EEG and vascular). The group is geared towards reproducible research using its own signal-processing and machine learning toolbox.

Face Processing

The group investigated the use of a new range of visual descriptors based on Local Binary Patterns (LBP) or on primitive binary features for the tasks of face detection, facial feature localization and face recognition. We proposed initially to consider the histogram of LBP code as a probability distribution that can be adapted through a learning mechanism, but we also proposed a simple and effective approach to consider LBP as a pre-processing that filters out illumination variations from the face image (Figure 9).

We also explored an alternative search technique to the typical sliding window approach for the task of face detection. The proposed alternative search technique, based on primitive binary features, can improve the detection speed without decreasing for the detection rate.

More recently, we proposed a novel unified boosting framework, applied to several face processing tasks (face detection, facial feature localization and pose classification), based on the same boosting algorithm and the same pool of features (Local Binary Patterns). This is in contrast with the standard approaches that make use of a variety of features and models, for example AdaBoost, cascades of boosted classifiers and Active Appearance Models

Face and Speaker Recognition

We leveraged prior work on distribution modeling for part-based face recognition (Figure 10) using session variability modeling techniques. Session variability modeling aims to explicitly model and suppress detrimental within-class (inter-session) variation. We examined two techniques to do this, inter-session variability modeling (ISV) and joint factor analysis (JFA), which were initially developed for speaker recognition.



Figure 9: Illustration of the LBP feature map (b) for the original image (a).

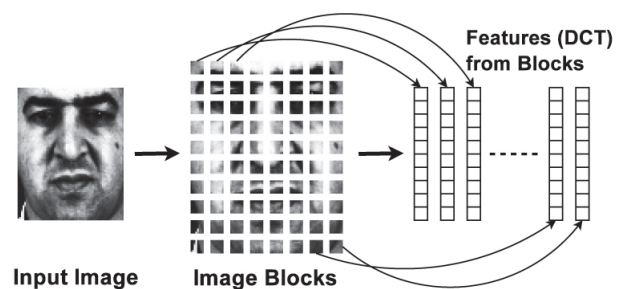


Figure 10: The concept of a parts-based approach: dividing the face into blocks and obtaining a feature vector from each block.

We also developed recently a scalable formulation of Probabilistic Linear Discriminant Analysis (PLDA). PLDA is a probabilistic model that has been shown to provide state-of-the-art performance for both face and speaker recognition. However, it has one major drawback, at training time estimating the latent variables requires the inversion and storage of a matrix whose size grows quadratically with the number of samples for the identity (class). We developed a scalable derivation which is theoretically equivalent to the non-scalable solution and so obviates the need for a variational approximation. The main advantage of the proposed approach is a gain in performance when using more training samples per identity (class), which is made possible by the proposed scalable formulation of PLDA.

Anti-Spoofing

One important aspect of biometric systems is their reliability not only when assaulted by impostors, but also under different types of attacks. One possible security threat is spoofing attacks: an action of outwitting a biometric sensor by presenting a counterfeit biometric evidence of a valid user. It is a direct attack to the sensory input of the biometric system and the attacker does not need previous knowledge about the recognition algorithm. Most of the biometric modalities are not resistant to spoofing attacks: a biometric system is usually designed to only recognize identities without concern whether the sample comes from a live person or not. Despite the existence of very sophisticated biometric systems nowadays, the



task of implementing anti-spoofing schemes for them has attracted much less attention.

Recently, we advanced the state-of-the-art in anti-spoofing to printed photo attacks in face recognition by proposing public benchmarks and effective solutions based on texture modeling, motion detection or spatio-temporal pattern recognition. We also organized the first International Competition on the topic.

Reproducible Research

Over the last year the group developed various software libraries to support its research activities. Two of these libraries have been released open source starting with Torchvision (<http://torch3vision.idiap.ch/>) and finally with Bob (<http://www.idiap.ch/software/bob/>) a major evolution.

Bob is a free signal processing and machine learning toolbox. The toolbox is designed to meet the needs of researchers by reducing development time and efficiently processing data. Firstly, Bob provides a researcher-friendly Python environment for rapid development. Secondly, efficient processing of large amounts of multimedia data is provided by fast Cpp implementations of identified bottlenecks. The Python environment is integrated seamlessly with the Cpp library, which ensures the library is easy to use and extensible. Thirdly, Bob supports reproducible research through its integrated experimental protocols for several databases. Finally, a strong emphasis is placed on code clarity, documentation, and thorough unit testing. Bob is thus an attractive resource for researchers due to this unique combination of ease of use, efficiency, extensibility and transparency. Bob is an open-source library and an ongoing community effort.

Biometrics Evaluation and Testing

Within the EU BEAT project³⁵, coordinated by Idiap, we aim at building a biometry-independent platform for Biometrics research, development and certification. By making use of such a system, academic or industrial partners enable users to easily compare results from distinct algorithms or/and parameterizations with minimal interaction using one or potentially many biometric traits. Here are three practical examples:

1. A teacher asks students to implement a specific illumination normalization algorithm in a pre-configured face recognition toolchain. The teacher would like to receive the examination results from his group of students;

2. A researcher (PhD student or Professor) wishes to test new ideas for pre-processing (i.e. to solve translation, rotation, and scaling problems due to different positions and downward pressure of the finger) in a fingerprint recognition system. She/He wishes to upload the new algorithm to an existing fingerprint-recognition toolchain and measure its impact on the overall system performance;

3. A scientific challenge is organized in which users contribute with features that will be used in a pre-configured key-point localization algorithm, for face recognition. The user wishes to upload his/her own feature-extraction algorithm to the contest and observe how that performs.

8. Natural Language Processing

The Idiap Natural Language Processing group carries out research in semantic-level language analysis with two principal applications: machine translation (MT) and information retrieval (IR). The NLP group currently includes seven people, with additional support from engineers and interns. The overall goal of the NLP group is to improve text analysis performance by using robust semantic and pragmatic analyses in combination with state-of-the-art, low-level, data-driven techniques. These baseline techniques are, respectively, phrase-based statistical machine translation, and various content-based or collaborative filtering algorithms for computing similarity in networked data.

Machine Translation

Work on this topic, mainly within the COMTIS SNSF Sinergia project³⁶ (started in 2010 and coordinated by Idiap) focuses on a problem that is not so often targeted in the current statistical machine translation (SMT) paradigm: the translation of words that establish relationships between sentences. In collaboration with two teams in Geneva, from linguistics and computational learning, we have analyzed large collections of examples of various types of dependencies between sentences, such as rhetorical relations signaled by discourse connectives, which are problematic for current MT engines. We have proposed computational methods for the automatic disambiguation of connectives, using classifiers such as maximum entropy or decision trees based on surface features, but also on more semantic features such as synonymy or word polarity. The accuracy of our classifiers for discourse connectives

³⁵ <http://www.beat-eu.org/>

³⁶ <http://www.idiap.ch/comtis>



tives (e.g. distinguishing between temporal and causal uses of 'since') is competitive against the state of the art. The classifiers have been combined with phrase-based and hierarchical SMT systems, using methods such as concatenated labels or factored models. The systems were trained on large corpora (e.g. Europarl) that were automatically labeled with connectives, with English as the source language, and French, German, Italian and Arabic as target languages. We have demonstrated that the combination of connective classifiers and SMT brought improvement to the translation of discourse connectives, without affecting the rest of the words. In a similar vein, we are dealing with pronouns and verb tenses, which also depend on dependencies across sentences, and are a hot topic in SMT research. Following an internship on pronouns in 2011, we collected data and started classification experiments on the translation of English past tenses into French, in collaboration with the University of Geneva. Data is a central issue and annotation has been carried out. An evaluation metric intended for discourse connectives has been designed, under the name of ACT, for Accuracy of Connective Translation. The metric compares the correct translation with the one generated by an SMT, in terms of identical, synonymous and incompatible translations, managing also the cases when a connective is skipped in one of the translations.

Similarity and Retrieval in Graph-based Multimedia Repositories

The NLP group develops methods for multimedia information retrieval based on the analysis of content, including sentiment analysis, and taking advantage of the networked structure of social media repositories. This research is grounded in work on multimodal meeting analysis and meeting browsers, in the context of the AMI³⁷ and AMIDA³⁸ EU integrated projects, and the NCCR IM2³⁹. We also contribute to the recent InEvent⁴⁰ EU project, and receive funding for technology transfer in collaboration with Klewel SA.

In theoretical work, we focused on similarity metrics over networked data, including multimedia data represented mainly through its metadata fields. A probabilistic random walk model was used to define the concept of Visiting Probability from a graph node to another, and a distance measure was defined based on it. Several similarity learning models were proposed over graphs, and were applied to natural language processing tasks (such as word similarity or document clustering) and information retrieval ones, including learning-to-rank. The Automatic Content Linking Device (ACLD) serves here as a general applicative framework. The ACLD is a multimedia re-

trieval system with spoken queries (explicit or implicit) and running in real-time (e.g. to enrich a conversation) or offline (e.g. to enrich a past lecture). Several retrieval mechanisms have been explored to increase the robustness of the ACLD against ASR errors. In particular, a method based on Visiting Probability has been shown to outperform keyword-based retrieval. Evaluation metrics for this task are mainly based on binary comparisons of results by human subjects recruited using a crowdsourcing platform. Current work attempts to derive, in the ACLD setting, the optimal subset of keywords from the spoken conversation, by maintaining multiple hypotheses about the topics of the information needs of the users, in order to maximize the diversity *and* relevance of ACLD results when only a few of them can be displayed.

The NLP group currently approaches the problem of retrieving multimedia networked data from the perspective of recommender systems. We have used as a benchmark the TED dataset of about 1100 lectures, with lists of user favorites and ground truth generic recommendations, to compare a number of approaches: content-based, collaborative filtering, and combined ones, and identified the most useful metadata features. We have shown that performing sentiment analysis on user comments improves the accuracy of recommendations, and we are now considering aspect-based sentiment analysis, as well as attempting to recommend segments rather than entire lectures.

37 <http://www.amiproject.org>

38 <http://www.amidaproject.org>

39 <http://www.im2.ch>

40 <http://www.inevent-project.eu>



MAIN PROJECTS IN PROGRESS

ACRONYM NAME, NAME

PARTNERS

EUROPEAN PROJECTS

BBFOR2

Bayesian Biometrics for Forensics

Radboud University Nijmegen
Universidad Autonoma de Madrid
Politecnico di Torino
Universiteit Twente
University of York
Katholieke Universiteit Leuven
Högskolan i Halmstad
Netherlands Forensic Institute
Agnitio Voice Biometrics
Netherlands organisation for applied scientific research (TNO)

BEAT

Biometrics Evaluation and Testing

University of Surrey
Universidad Autónoma de Madrid
TÜViT
TUBITAK Informatics and Information Security, Advanced Technologies Research Center
Sagem Sécurité SA
Katholieke Universiteit Leuven
École Polytechnique Fédérale de Lausanne (EPFL)
Commissariat à l'Energie Atomique

HUMAVIPS

Humanoids with Auditory and Visual Abilities in Populated Spaces

Institut National de Recherche en Informatique et Automatique
The Czech Technical University
Aldebaran Robotics
Bielefeld University

InEvent

Accessing Dynamic Networked Multimedia Events

Radvision Ltd
IBM Israël
Klewel SA
Fraunhofer Heinrich-Hertz-Institute
University of Edinburgh

MASH

Massive Sets of Heuristics for Machine Learning

Centre National de la Recherche Scientifique
Universität Potsdam
Institut National de Recherche en Informatique et en Automatique
The Czech Technical University

PASCAL2

Pattern Analysis, Statistical Modelling and Computational Learning

56 sites in the network

SCALE

Speech Communication with Adaptive Learning

Universität des Saarlandes
University of Edinburgh
University of Sheffield
Radboud University Nijmegen
RWTH Aachen
Motorola Limited UK
Philips
Eurice



DURATION (MONTH/YEAR)**WEB****COORDINATOR****CONTACT**

01.10 – 12.13

<http://lands.let.ru.nl/bbfor2>

Stichting Katholieke Universiteit

Dr. Sébastien Marcel

03.12 – 02.16

www.beat-eu.org

Idiap Research Institute

Dr. Sébastien Marcel

02.10 – 01.13

<http://humavips.inrialpes.fr/>Institut de Recherche
en Informatique et
en AutomatiqueDr. Daniel Gatica-Perez
Dr. Jean-Marc Odobez

11.11 – 10.14

www.inevent-project.eu

Idiap Research Institute

Prof. Hervé Bourlard

01.10 – 12.12

<http://mash-project.eu>

Idiap Research Institute

Dr. François Fleuret

03.08 – 02.13

www.pascal-network.org

University of Southampton

Dr. François Fleuret

01.09 – 12.12

www.scale.uni-saarland.de

Universität des Saarlandes

Prof. Hervé Bourlard



ACRONYM NAME, NAME**PARTNERS****SSPnet**

Social Signal Processing Network

Imperial College of Science, Technology and Medicine, London
University of Edinburgh
University of Twente
Università Di Roma Tre
Queen's University Belfast
DFKI
INRIA
Université de Genève
Technische Universiteit Delft

TA2

Together Anywhere, Together Anytime

EURESCOM - European Institute for Research and Strategic Studies in Telecommunications GmbH
British Telecommunications plc
Alcatel-Lucent Bell NV
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
Goldsmiths' College
Netherlands Organisation For Applied Scientific Research – TNO
The Interactive Institute II Aktiebolag
Stichting Centrum voor Wiskunde en Informatica
Ravensburger Spieleverlag GmbH
Philips Consumer Electronics BV
Limbic Entertainment GmbH
Joanneum Research Forschungsgesellschaft GmbH
Brno University of Technology

TA2-EEUTogether Anywhere, Together Anytime
- Enlarged European Union

Brno University of Technology
Joanneum Research
Forschungsgesellschaft mbH
Goldsmiths College, University of London

TABULA RASATrusted Biometrics
under Spoofing Attacks

University of Oulu
Universidad Autonoma de Madrid
University of Southampton
University of Cagliari
EURECOM
Chinese Academy of Sciences
Starlab Barcelona S.L.
MORPHO
KeyLemon SA
BIOMETRY.com AG
Centre for Science, Society and Citizenship
Sagem Sécurité SA

VANAHEIMVideo/Audio Networked Surveillance
System Enhancement through
Human-Centered Adaptive Monitoring

Multitel ASBL
Institut National de Recherche en Informatique et Automatique
Thales Communications France
Thales Italia
Gruppo Torinese Trasporti
Régie Autonome des Transports Parisiens
University of Vienna



DURATION (MONTH/YEAR)	WEB	COORDINATOR	CONTACT
02.09 – 01.14	www.sspnet.eu	Idiap Research Institute	Dr. Alessandro Vinciarelli
02.10 – 03.12	www.ta2-project.eu	Eurescom	Dr. Phil Garner
02.10 – 03.12	www.ta2-project.eu	Eurescom	Dr. Phil Garner
11.10 – 06.14	www.tabularasa-euproject.org	Idiap Research Institute	Dr. Sébastien Marcel
02.10 – 07.13	www.vanaheim-project.eu	Multitel ASBL	Dr. Jean-Marc Odobez



ACRONYM NAME, NAME**PARTNERS****SNSF PROJECTS****AMSP**

Auditory-motivated signal processing and applications to robust speech enhancement and recognition

A-MUSE

Adaptive Multilingual Speech Processing

AROLES

Automatic Recommendation of Lectures and Snippets

Kleweil

CODICES

Automatic Analysis of Mexican Codex Collections

COMTIS

Improving the Coherence of Machine Translation Output by Modeling Intersentential Relations

University of Geneva/Department of Linguistics
University of Geneva/Department of Computer Science

DASH

Object Detection with Active Sample Harvesting

DIMHA

Diarizing Massive Amounts of Heterogeneous Audio

Koemei SA

FlexASR

Flexible Grapheme-Based Automatic Speech Recognition

HAI-2010

Human Activity and Interactivity Modeling

ICS-2010

Interactive Cognitive Systems

NCCR IM2

Interactive Multimodal Information Management

Ecole Polytechnique Fédérale de Lausanne (EPFL)
University of Geneva
University of Fribourg
Swiss Federal Institute of Technology in Zurich (ETHZ)

LOBI

Low Complexity Binary Features for Robust-to-Noise Speaker Recognition

KeyLemon SA

MCM-FF

Multimodal Computational Modeling of Nonverbal Social Behavior in Face to Face Interaction

MULTIO8EXT

Multimodal Interaction and Multimedia Data Mining

NINAPRO

Non-Invasive Adaptive Hand Prosthetics

Institute of Robotics and Mechatronics
HES-SO Valais

PROMOVAR

Probabilistic Motifs for Video Action Recognition

RODI

Role based speaker diarization

Institute of Robotics and Mechatronics
HES-SO Valais



DURATION (MONTH/YEAR)	WEB	COORDINATOR	CONTACT
01.12 – 07.12		Indian Institute of Science	Dr. Mathew Magimai-Doss
10.12 – 09.15		Idiap Research Institute	Prof. Hervé Bourlard
06.12 – 05.14		Idiap Research Institute	Dr. Andrei Popescu-Belis
08.08 – 07.12		Idiap Research Institute	Dr. Daniel Gatica-Perez
03.10 – 02.13	www.idiap.ch/comtis	Idiap Research Institute	Dr. Andrei Popescu-Belis
09.12 – 08.15		Idiap Research Institute	Dr. François Fleuret
06.12 – 05.14		Idiap Research Institute	Dr. Fabio Valente
03.10 – 04.13		Idiap Research Institute	Dr. Mathew Magimai-Doss
10.10 – 09.12		Idiap Research Institute	Dr. Jean-Marc Odobez
10.10 – 09.12		Idiap Research Institute	Prof. Hervé Bourlard
01.02 – 12.13	www.im2.ch	Idiap Research Institute	Prof. Hervé Bourlard
06.12 – 05.14		Idiap Research Institute	Dr. Sébastien Marcel
11.11 – 10.14		Idiap Research Institute	Dr. Oya Aran Karakus
10.10 – 09.12		Idiap Research Institute	Prof. Hervé Bourlard
01.11 – 12.13	www.idiap.ch/project/ninapro	Idiap Research Institute	Dr. Barbara Caputo
02.12 – 07.13		Idiap Research Institute	Dr. Jean-Marc Odobez
11.11 – 10.14		Idiap Research Institute	Dr. Fabio Valente



ACRONYM NAME, NAME**PARTNERS****SIWIS**

Spoken Interaction with Interpretation in Switzerland

University of Geneva
University of Edinburgh
Eidgenössische Technische Hochschule Zürich

SONVB

Sensing and Analysing Organizational Nonverbal Behavior

University of Neuchâtel
Dartmouth College

TRACOME

Robust Face Tracking, Feature Extraction and Multimodal Fusion for Audio-Visual Speech Recognition

UBM

Understanding Brain Morphogenesis

Ecole Polytechnique Fédérale de Lausanne (EPFL)
University of Basel

VELASH

Very Large Sets of Heuristics for Scene Interpretation

VISION@HOME

Situated Vision to Perceive Object Shape and Affordances

SNSF PROJECTS (INDO-SUISSE)**CCPP**

Cross Cultural Personality Perception

University of Geneva
International Institute of Information Technology, India

HASLER FOUNDATION**ADDG2SU**

Flexible Acoustic Data-Driven Grapheme to Subword Unit Conversion

CLAS3

Cross-Lingual Adaptation for Text to Speech Synthesis

DEEPSTD

Universal Spoken Term Detection with Deep Learning

SESAME

SEarching Swiss Audio MEMories

SODS

Semantically Self-Organized Distributed Web Search

UBSL

User-Based Similarity Learning for Interactive Image Retrieval

V-FAST

Vocal-tract based Fast Adaptation for Speech Technology

ARMASUISSE**RECOD 2012**

Very low bit-rate speech coding

In addition to the above projects a number of industrial projects (CTI, The Ark) and grants are ongoing at Idiap.



DURATION (MONTH/YEAR)	WEB	COORDINATOR	CONTACT
12.12 – 11.15		Idiap Research Institute	Dr. Phil Garner
06.10 – 05.13	www.idiap.ch/project/sonvb	Idiap Research Institute	Dr. Daniel Gatica-Perez
01.11 – 12.13		Ecole Polytechnique Fédérale de Lausanne (EPFL)	Dr. Jean-Marc Odobez
01.11 – 12.13		University of Basel	Dr. François Fleuret
09.09 – 08.12		Idiap Research Institute	Dr. François Fleuret
10.11 – 09.14		Idiap Research Institute	Dr. Barbara Caputo
05.09 – 04.12	www.idiap.ch/project/ccpp	Idiap Research Institute	Dr. Alessandro Vinciarelli
01.13 – 12.15		Idiap Research Institute	Dr. Mathew Magimai-Doss
11.11 – 08.12		Idiap Research Institute	Dr. John Dines
12.11 – 11.14		Idiap Research Institute	Dr. Ronan Collobert
01.11 – 12.13		Idiap Research Institute	Dr. Fabio Valente
01.12 – 12.14		Idiap Research Institute	Dr. Ronan Collobert
08.12 – 03.13		Idiap Research Institute	Dr. François Fleuret
03.12 – 12.12		Idiap Research Institute	Dr. Phil Garner
07.12 – 10.12		Idiap Research Institute	Phil Garner



MAJOR PUBLICATIONS / CONFERENCES

This selection, from among the many publications of Idiap, illustrates the diversity of our research.

BOOKS, BOOK CHAPTERS AND JOURNAL PAPERS

Multimodal Signal Processing: Human Interactions in Meetings

Steve Renals, Hervé Bourlard, Jean Carletta and Andrei Popescu-Belis
Cambridge University Press, 2012

Together Anywhere, Together Anytime, Technologies for Intimate Interactions

Dick C. A. Bulterman, Petr Motlicek, Stefan Duffner and Danil Korchagin
Centrum Wiskunde & Informatica, 2012

Deep Learning via Semi-Supervised Embedding

Jason Weston, Frédéric Ratle, Hossein Mobahi and Ronan Collobert
in: In Neural Networks: Tricks of the Trade, Springer, 2012

Evaluation of Meeting Support Technology

Simon Tucker and Andrei Popescu-Belis
in: Multimodal Signal Processing: Human Interactions in Meetings, pages 237-252, Cambridge University Press, 2012

From Nonverbal Cues to Perception: Personality and Social Attractiveness

Alessandro Vinciarelli, Hugues Salamin, Anna Polychroniou, Gelareh Mohammadi and Antonio Origlia
in: LNCS Proceedings on Cognitive Behavioural Systems, Springer, 2012

Implementing Neural Networks Efficiently

Ronan Collobert, Koray Kavukcuoglu and Clément Farabet
in: Neural Networks: Tricks of the Trade, Springer, 2012

Multimodal Signal Processing for Meetings: an Introduction

Andrei Popescu-Belis and Jean Carletta
in: Multimodal Signal Processing: Human Interactions in Meetings, pages 1-11, Cambridge University Press, 2012

Sampling techniques for audio-visual tracking and head pose estimation

Jean-Marc Odobez and Oswald Lanz
in: Multimodal Signal Processing: Human Interactions in Meetings, pages 84-102, Cambridge University Press, 2012

Sparsity in Topic Models

Jagannadan Varadarajan, Remi Emonet and Jean-Marc Odobez
in: Practical Applications of Sparse Modeling: Biology, Signal Processing and Beyond, MIT Press, 2012

Speaker Diarization

Fabio Valente and Gerald Friedland
in: Multimodal Signal Processing: Human Interactions in Meetings, Cambridge University Press, 2012

User Requirements for Meeting Support Technology

Denis Lalanne and Andrei Popescu-Belis
in: Multimodal Signal Processing: Human Interactions in Meetings, pages 210-221, Cambridge University Press, 2012

A Fast Parts-based Approach to Speaker Verification using Boosted Slice Classifiers

Anindya Roy, Mathew Magimai.-Doss and Sébastien Marcel
in: IEEE Transactions on Information Forensics and Security, 7(1):241-254, 2012

A Method, Apparatus and Computer Program for Determining the Location of a Plurality of Speech Source

Afsaneh Asaei, Hervé Bourlard and Volkan Cevher,
in: 2012US-13/654055, 2012

A Probabilistic Approach to Mining Mobile Phone Data Sequences

Katayoun Farrahi and Daniel Gatica-Perez
in: Personal and Ubiquitous Computing, 2012

A real-time deformable detector

Karim Ali, Francois Fleuret, David Hasler and Pascal Fua
in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012

A Sequential Topic Model for Mining Recurrent Activities from Long Term Video Logs

Jagannadan Varadarajan, Remi Emonet and Jean-Marc Odobez
in: International Journal of Computer Vision, 2012

A Simple Continuous Pitch Estimation Algorithm

Philip N. Garner, Milos Cernak and Petr Motlicek
in: IEEE Signal Processing Letters, 2012

A Unified Structured Sparse Coding Framework for Spatio-Spectral Information Recovery

Afsaneh Asaei, Hervé Bourlard, Bhiksha Raj and Volkan Cevher
in: IEEE Transactions on Audio, Speech and Language Processing, 2012

Assessing Sparse Coding Methods for Contextual Shape Indexing of Maya Hieroglyphs

Edgar Roman-Rangel, Jean-Marc Odobez and Daniel Gatica-Perez
in: Journal of Multimedia, 7(2):179--192, 2012

Audiovisual Diarization Of People In Video Content

Elie Khoury, Christine Sénac and Philippe Joly
in: Multimedia Tools and Applications, 2012

Automatic Attribution of Personality Traits Based on Prosodic Features

Gelareh Mohammadi and Alessandro Vinciarelli
in: IEEE Transactions on Affective Computing, 2012

Automatic Role Recognition in Multiparty Conversations: an Approach Based on Turn Organization, Prosody and Conditional Random Fields

Hugues Salamin and Alessandro Vinciarelli
in: IEEE Transactions on Multimedia, 2012



**Bridging the Gap Between Social Animal and Unsocial Machine:
A Survey of Social Signal Processing**

Alessandro Vinciarelli, Maja Pantic, Dirk Heylen, C. Pelachaud, I. Poggi,
F. D'Errico and M. Schroeder
in: IEEE Transactions on Affective Computing, 2012

**Conversation Analysis at Work: Detection of Conflict in Competitive
Discussions through Automatic Turn-Organization Analysis**

A. Pesarin, M. Cristani, V. Murino and Alessandro Vinciarelli
in: Cognitive Processing, 2012

**Cross-pollination of normalisation techniques from speaker to face
authentication using Gaussian mixture models**

Roy Wallace, Mitchell McLaren, Chris McCool and Sébastien Marcel
in: IEEE Transactions on Information Forensics and Security,
7(2):553 -- 562, 2012

Discovering Places of Interest in Everyday Life from Smartphone Data

R. Montoliu, J. Blom and Daniel Gatica-Perez
in: Multimedia Tools and Applications, 2012

**Emergent leaders through looking and speaking: from audio-visual
data to multimodal recognition**

Dairazalia Sanchez-Cortes, Oya Aran, Dinesh Babu Jayagopi,
Marianne Schmid Mast and Daniel Gatica-Perez
in: Journal on Multimodal User Interfaces, 2012

Finding Information in Multimedia Records of Meetings

Andrei Popescu-Belis, Denis Lalanne and Hervé Bourlard
in: IEEE Multimedia, 19(2):48-57, 2012

Human Interaction Discovery in Smartphone Proximity Networks

Trinh-Minh-Tri Do and Daniel Gatica-Perez
in: Personal and Ubiquitous Computing, 2012

**Improving Control of Dexterous Hand Prostheses Using Adaptive
Learning**

Tatiana Tommasi, Francesco Orabona, Claudio Castellini and
Barbara Caputo
in: IEEE TRANSACTIONS ON ROBOTICS, 2012

**Jointly Data-Driven and Phonological Knowledge-Guided Enhance-
ment of State Mapping-Based Cross-Lingual Speaker Adaptation**

Hui Liang and John Dines
in: IEEE Transactions on Audio, Speech and Language Processing,
2012

Mining Large-Scale Smartphone Data for Personality Studies

Gokul Chittaranjan, Jan Blom and Daniel Gatica-Perez
in: Personal and Ubiquitous Computing, 2012

**Multistream speaker diarization of meetings recordings
beyond MFCC and TDOA features**

Deepu Vijayasenan, Fabio Valente and Hervé Bourlard
in: Speech Communication, 54(1), 2012

**Phase AutoCorrelation (PAC) features for noise robust speech
recognition**

Shajith Ikbal, Hemant Misra, Hynek Hermansky and
Mathew Magimai.-Doss
in: Speech Communication, 54(7):867-880, 2012

**Real-time model learning using Incremental Sparse Spectrum
Gaussian Process Regression**

Arjan Gijsberts and Giorgio Metta
in: Neural Networks, 2012

Regularized Bundle Methods for Convex and Non-Convex Risks

Trinh-Minh-Tri Do and Thierry Artieres
in: Journal of Machine Learning Research, 13:3539-3583, 2012

**Structured Sparsity Models for Multiparty Speech Recovery
from Reverberant Recordings**

Afsaneh Asaei, Mohammad Golbabaee, Hervé Bourlard and Volkan
Cevher
in: IEEE Transaction on Audio, Speech and Language Processing,
2012

The ICSI RT-09 Speaker Diarization System

Gerald Friedland, Adam Janin, David Imseng, Xavier Anguera, Luke
Gottlieb, Marijn Huijbregts, Mary Tai Knox and Oriol Vinyals
in: IEEE Transactions on Audio, Speech, and Language Processing,
20(2):371--381, 2012

**The TA2 Database – A Multi-Modal Database From Home
Entertainment**

Stefan Duffner, Petr Motlicek and Danil Korchagin
in: International Journal of Computer and Electrical Engineering,
4(5):670-673, 2012

**The YouTube Lens: Crowdsourced Personality Impressions
and Audiovisual Analysis of Vlogs**

Joan-Isaac Biel and Daniel Gatica-Perez
in: IEEE Transactions on Multimedia, 2012

Transcribing meetings with the AMIDA systems

Thomas Hain, Lukas Burget, John Dines, Philip N. Garner, Frantisek
Grezl, Asmaa El Hannani, Marijn Huijbregts, Martin Karafiat,
Mike Lincoln and Vincent Wan
in: IEEE Transactions on Audio, Speech, and Language Processing,
20(2):486--498, 2012

**Vocal Tract Length Normalization for Statistical Parametric
Speech Synthesis**

Lakshmi Saheer, John Dines and Philip N. Garner
in: IEEE Transactions on Audio, Speech and Language Processing,
2012

**Wordless Sounds: Robust Speaker Diarization using
Privacy-Preserving Audio Representations**

Sree Hari Krishnan Parthasarathi, Hervé Bourlard and
Daniel Gatica-Perez
in: IEEE Transactions on Audio, Speech, and Language Processing,
2012



A Multiple Hypothesis Gaussian Mixture Filter for Acoustic Source Localization and Tracking

Youssef Oualil, Friedrich Faubel and Dietrich Klakow
in: 13th International Workshop on Acoustic Signal Enhancement, pages 233-236, 2012

A TDOA Gaussian Mixture Model for Improving Acoustic Source Tracking

Youssef Oualil, Friedrich Faubel, Mathew Magimai.-Doss and Dietrich Klakow
in: 20th European Signal Processing Conference, 2012

A tree-based distance between distributions: application to classification of neurons

Rival Lefort and Francois Fleuret
in: ICASSP 2012 : IEEE International Conference on Acoustics, Speech and Signal Processing, 2012

Acoustic Data-driven Grapheme-to-Phoneme Conversion using KL-HMM

Ramya Rasipuram and Mathew Magimai.-Doss
in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, 2012

An Agent-Based Focused Crawling Framework for Topic- and Genre-Related Web Document Discovery

Nikolaos Pappas, Georgios Katsimpras and Efstathios Stamatatos
in: 24th IEEE International Conference on Tools with Artificial Intelligence, Athens, Greece, IEEE, 2012

An Open Source Framework for Standardized Comparisons of Face Recognition Algorithms

Manuel Günther, Roy Wallace and Sébastien Marcel
in: Computer Vision - ECCV 2012. Workshops and Demonstrations, Idiap Research Institute, Heidelberg, pages 547-556, Springer Berlin, 2012

Annotation and Recognition of Personality Traits in Spoken Conversations from the AMI Meetings Corpus

Fabio Valente, Samuel Kim and Petr Motlicek
in: Proceedings of Interspeech 2012, 2012

Assessing the Impact of Language Style on Emergent Leadership Perception from Ubiquitous Audio

Dairazalia Sanchez-Cortes, Petr Motlicek and Daniel Gatica-Perez
in: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, Ulm, Germany, 2012

Automatic detection of conflict escalation in spoken conversations

Samuel Kim, Sree Harsha Yella and Fabio Valente
in: INTERSPEECH, ISCA, Portland, Oregon, USA., 2012

Automatic detection of conflicts in spoken conversations: ratings and analysis of broadcast political debates

Samuel Kim, Fabio Valente and Alessandro Vinciarelli
in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, 2012

Automatic Speaker Role Labeling in AMI Meetings: Recognition of Formal and Social Roles,

A. Sapru and Fabio Valente
in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, 2012, 2012

Baseline Multimodal Place Classifier for the 2012 Robot Vision Task, Jesus Martinez-Gomez, Ismael

Garcia-Varea and Barbara Caputo
in: Working Notes of the ImageCLEF 2012 Laboratory, 2012

Beyond Dataset Bias: Multi-task Unaligned Shared Knowledge Transfer

Tatiana Tommasi, Novi Quadrianto, Barbara Caputo and Christoph H. Lampert
in: Asian Conference on Computer Vision, 2012

Bi-Modal Authentication in Mobile Environments Using Session Variability Modelling

Petr Motlicek, Laurent El Shafey, Roy Wallace, Chris McCool and Sébastien Marcel
in: Proceedings of the 21st International Conference on Pattern Recognition, 2012

Bi-Modal Person Recognition on a Mobile Phone: using mobile phone data

Chris McCool, Sébastien Marcel, Abdenour Hadid, Matti Pietikainen, Pavel Matejka, Jan Cernocky, Norman Poh, J. Kittler, Anthony Larcher, Christophe Levy, Driss Matrouf, Jean-François Bonastre, Phil Tresadern and Timothy Cootes
in: IEEE ICME Workshop on Hot Topics in Mobile Multimedia, 2012

Bob: a free signal processing and machine learning toolbox for researchers

André Anjos, Laurent El Shafey, Roy Wallace, Manuel Günther, Chris McCool and Sébastien Marcel
in: Proceedings of the ACM Multimedia Conference, 2012

Boosting localized binary features for speech recognition

Anindya Roy, Mathew Magimai.-Doss and Sébastien Marcel
in: Symposium on Machine Learning in Speech and Language Processing (MLSLP), 2012

Boosting under-resourced speech recognizers by exploiting out of language data - Case study on Afrikaans

David Imseng, Hervé Bourlard and Philip N. Garner
in: Proceedings of the 3rd International Workshop on Spoken Languages Technologies for Under-resourced Languages, Cape Town, pages 60-67, 2012

Bridging the Past, Present and Future: Modeling Scene Activities From Event Relationships and Global Rules

Jagannadan Varadarajan, Remi Emonet and Jean-Marc Odobez
in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, Providence, Rhode Island, USA, 2012

Building the NinaPro Database: a Resource for the Biorobotics Community

Manfredo Atzori, Arjan Gijsberts, Simone Heynen, Anne-Gabrielle Mittaz Hager, Olivier Deriaz, Patrick van der Smagt, Claudio Castellini, Barbara Caputo and Henning Müller
in: Proceedings of the Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics, 2012

Checking In or Checked In: Comparing Large-Scale Manual and Automatic Location Disclosure Patterns

Eric Malmi, Trinh-Minh-Tri Do and Daniel Gatica-Perez
in: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, Ulm, Germany, 2012



Collecting data for socially intelligent surveillance and monitoring approaches: the case of conflict in competitive conversations

Alessandro Vinciarelli, Samuel Kim, Fabio Valente and Hugues Salamin
in: International Symposium on Communications, Control, and Signal Processing, 2012

Combination of Sparse Classification and Multilayer Perceptron for Noise Robust ASR

Yang Sun, Mathew Magimai.-Doss, Jort F. Gemmeke, B. Cranen, Louis ten Bosch and Lou Boves
in: Proceedings of Interspeech, 2012

Combining Acoustic Data Driven G2P and Letter-to-Sound Rules for Under Resource Lexicon Generation

Ramya Rasipuram and Mathew Magimai.-Doss
in: Proceedings of Interspeech, Portland, Oregon, 2012

Combining Cepstral Normalization and Cochlear Implant-Like Speech Processing for Microphone Array-Based Speech Recognition

Cong-Thanh Do, Mohammad J. Taghizadeh and Philip N. Garner
in: Proceedings of the IEEE Workshop on Spoken Language Technology, 2012

Combining transcription-based and acoustic-based speaker identifications for broadcast news

Elie Khoury, Antoine Laurent, Sylvain Meignier and Simon Petitrenaud
in: IEEE International Conference on Acoustics, Speech, and Signal Processing, 2012

Combining Vocal Tract Length Normalization with Hierarchical Linear Transformations

Lakshmi Saheer, Junichi Yamagishi, Philip N. Garner and John Dines
in: Proceedings in International conference on Speech and Signal processing, Kyoto, Japan, pages 4493-4496, IEEE SPS (ICASSP), 2012

Comparing different acoustic modeling techniques for multilingual boosting

David Imseng, John Dines, Petr Motlicek, Philip N. Garner and Hervé Bourlard
in: Proceedings of Interspeech, Portland, Oregon, 2012

Computational Methods For Structured Sparse Component Analysis of Convolutional Speech Mixtures

Afsaneh Asaei, Michael E. Davies, Hervé Bourlard and Volkan Cevher
in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, 2012

Contextual Conditional Models for Smartphone-based Human Mobility Prediction

Trinh-Minh-Tri Do and Daniel Gatica-Perez
in: Proceedings of the 14th ACM International Conference on Ubiquitous Computing, 2012

Conversion of Recurrent Neural Network Language Models to Weighted Finite State Transducers for Automatic Speech Recognition

Gwénoél Lecorvé and Petr Motlicek
in: Proceedings of Interspeech, Portland, Oregon, USA, pages to appear, 2012

Crowdsourcing Micro-Level Multimedia Annotations: The Challenges of Evaluation and Interface

Sunghyun Park, Gelareh Mohammadi, Ron Artstein and Louis-Philippe Morency
in: Proceedings of International ACM Workshop on Crowdsourcing for Multimedia, 2012

Detecting and Labeling Folk Literature in Spoken Cultural Heritage Archives using Structural and Prosodic Features

Fabio Valente and Petr Motlicek
in: IEEE Content Based Multimedia Indexing, 2012

DiarTk : An Open Source Toolkit for Research in Multistream Speaker Diarization and its Application to Meetings Recordings

Deepu Vijayaseenan and Fabio Valente
in: Proceedings of Interspeech, 2012

Discourse-level Annotation over Europarl for Machine Translation: Connectives and Pronouns

Andrei Popescu-Belis, Thomas Meyer, Jeevanthi Liyanapathirana, Bruno Cartoni and Sandrine Zufferey
in: Proceedings of the eighth international conference on Language Resources and Evaluation (LREC), pages 5, 2012

Empirical validations of multilingual annotation schemes for discourse relations

Sandrine Zufferey, Liesbeth Degand, Andrei Popescu-Belis and Ted Sanders
in: 8th Joint ACL-ISO Workshop on Interoperable Semantic Annotation, 2012

Exact Acceleration of Linear Object Detectors

Charles Dubout and Francois Fleuret
in: Proceedings of the European Conference on Computer Vision, 2012

Experiences in the Creation of an Electromyography Database to Help Hand Amputated Persons

Manfredo Atzori, Arjan Gijsberts, Simone Heynen, Anne-Gabrielle Mittaz Hager, Claudio Castellini, Barbara Caputo and Henning Müller
in: Proceedings of the 24th European Medical Informatics Conference, 2012

Extracting Directional and Comparable Corpora from a Multilingual Corpus for Translation Studies

Bruno Cartoni and Thomas Meyer
in: Proceedings of the eighth international conference on Language Resources and Evaluation (LREC), Istanbul, TR, pages 6, 2012

Extracting Informative Textual Parts from Web Pages Containing User-Generated Content

Nikolaos Pappas, Georgios Katsimpras and Efstathios Stamatatos
in: 12th International Conference on Knowledge Management and Knowledge Technologies, ACM ICPS, Graz, Austria, pages 4:1-4:8, ACM, 2012

Extracting Mobile Behavioral Patterns with the Distant N-Gram Topic Model

Katayoun Farrahi and Daniel Gatica-Perez
in: Proceedings of the IEEE International Symposium on Wearable Computers, Newcastle, 2012

Face Recognition with Disparity Corrected Gabor Phase Differences

Manuel Günther, Dennis Haufe and Rolf P. Würtz
in: Artificial Neural Networks and Machine Learning, Heidelberg, pages 411-418, Springer Berlin, 2012

Face Verification using Gabor Filtering and Adapted Gaussian Mixture Models

Laurent El Shafey, Roy Wallace and Sébastien Marcel
in: Proceedings of the 11th International Conference of the Biometrics Special Interest Group, Darmstadt, Germany, pages 397-408, GI-Edition, 2012



FaceTube: predicting personality from facial expressions of emotion in online conversational video

Joan-Isaac Biel, Lucia Teijeiro-Mosquera and Daniel Gatica-Perez
in: Proceedings International Conference on Multimodal Interfaces (ICMI-MLMI), 2012

From Speech to Personality: Mapping Voice Quality and Intonation into Personality Differences

Gelareh Mohammadi, Antonio Origlia, Maurizio Pili and Alessandro Vinciarelli
in: in Proceedings of ACM Multimedia 2012, 2012

Gaze Estimation From Multimodal Kinect Data

Kenneth Funes and Jean-Marc Odobez
in: IEEE Conference in Computer Vision and Pattern Recognition, Workshop on Gesture Recognition, Providence, RI, USA, 2012

Impact du degré de supervision sur l'adaptation à un domaine d'un modèle de langage à partir du Web

Gwénolé Lecorvé, John Dines, Thomas Hain and Petr Motlicek
in: Actes de la conférence conjointe JEP-TALN-RECITAL 2012, Grenoble, France, pages 193-200, ATALA/AFCP, 2012

Improving Acoustic Based Keyword Spotting Using LVCSR Lattices

Petr Motlicek, Fabio Valente and Igor Szoke
in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, Japan, pages 4413-4416, 2012

Indoor Scene Recognition using Task and Saliency-driven Feature Pooling

Marco Fornoni and Barbara Caputo
in: Proceedings of the British Machine Vision Conference, Guildford, UK, 2012

Investigating the Midline Effect for Visual Focus of Attention Recognition

Samira Sheikhi and Jean-Marc Odobez
in: Int Conf. on Multimodal Interaction (ICMI), Santa Monica, 2012

Iterative Relevance Feedback with Adaptive Exploration/Exploitation Trade-off

Nicolae Suditu and Francois Fleuret
in: Proceedings of the 21st ACM Conference on Information and Knowledge Management, pages 1323-1331, 2012

Joint Detection and Localization of Multiple Speakers using a Probabilistic Interpretation of the Steered Response Power

Youssef Oualil, Mathew Magimai.-Doss, Friedrich Faubel and Dietrich Klakow
in: Statistical and Perceptual Audition Workshop, 2012

LBP-TOP based countermeasure against face spoofing attacks

Tiago de Freitas Pereira, André Anjos, José Mario De Martino and Sébastien Marcel
in: International Workshop on Computer Vision With Local Binary Pattern Variants - ACCV, pages 12, 2012

Leveraging over prior knowledge for online learning of visual categories

Tatiana Tommasi, Francesco Orabona, Mohsen Kaboli and Barbara Caputo
in: Proceedings of the British Machine Vision Conference, 2012

Linking Speaking and Looking Behavior Patterns with Group Composition, Perception, and Performance

Dinesh Babu Jayagopi, Dairazalia Sanchez-Cortes, Kazuhiro Otsuka, Junji Yamato and Daniel Gatica-Perez
in: Proceedings of the International Conference on Multimodal Interaction (ICMI), Santa Monica, USA, 2012

Machine Translation of Labeled Discourse Connectives

Thomas Meyer, Andrei Popescu-Belis, Najeh Hajlaoui and Andrea Gesmundo
in: Proceedings of the Tenth Biennial Conference of the Association for Machine Translation in the Americas (AMTA), pages 10, 2012

Macro-Action Discovery Based on Change Point Detection and Boosting

Leonidas Lefakis and Francois Fleuret
in: International Conference on Machine Learning and Applications, 2012

MediaParl: Bilingual mixed language accented speech database

David Imseng, Hervé Boudlard, Holger Caesar, Philip N. Garner, Gwénolé Lecorvé and Alexandre Nanchen
in: Proceedings of the 2012 IEEE Workshop on Spoken Language Technology, pages 263--268, 2012

Microphone Array Beampattern Characterization for Hands-free Speech Applications

Mohammad J. Taghizadeh, Philip N. Garner and Hervé Boudlard
in: IEEE 7th Sensor Array and Multichannel Signal Processing Workshop(SAM), Hoboken, NJ, USA, pages 473-476, 2012

Modeling dominance effects on nonverbal behaviors using granger causality

Kyriaki Kalimeri, Bruno Lepri, Oya Aran, Dinesh Babu Jayagopi, Daniel Gatica-Perez and Fabio Pianesi
in: Proceedings of International Conference on Multimodal Interaction, ICMI 2012, Santa Monica, CA, 2012

Multimodal Cue Detection Engine for Orchestrated Entertainment

Danil Korchagin, Stefan Duffner, Petr Motlicek and Carl Scheffler
in: Proceedings International Conference on MultiMedia Modeling, Klagenfurt, Austria, 2012

On Speaker-Independent Personality Perception and Prediction from Speech

Polzehl Tim, Schoenenberg Katrin, Moller Sebastian, Metz Florian, Gelareh Mohammadi and Alessandro Vinciarelli
in: Proceedings of INTERSPEECH 2012, 2012

On the Challenge of Classifying 52 Hand Movements from Surface Electromyography

Ilija Kuzborskij, Arjan Gijsberts and Barbara Caputo
in: 34th Annual Conference of the IEEE Engineering in Medicine & Biology Society, 2012

On the Effectiveness of Local Binary Patterns in Face Anti-spoofing

Ivana Chingovska, André Anjos and Sébastien Marcel
in: Proceedings of the 11th International Conference of the Biometrics Special Interest Group, 2012

Overview of the ImageCLEF 2012 Robot Vision Task

Jesus Martinez-Gomez, Ismael Garcia-Varea and Barbara Caputo
in: Working Notes of the ImageCLEF 2012 Laboratory, 2012



Predicting the Conflict Level in Television Political Debates: an Approach Based on Crowdsourcing, Nonverbal Communication and Gaussian Processes

Samuel Kim, Maurizio Filippone, Fabio Valente and Alessandro Vinciarelli
in: ACM Multimedia, 2012

Reading Companion: The Technical and Social Design of an Automated Reading Tutor

Arthur Kantor, Milos Cernak, Jiri Havelka, Sean Huber, Jan Kleindienst and Doris B. Gonzalez
in: Workshop on Child, Computer and Interaction, Portland, Oregon, U.S.A., 2012

Recognizing the Visual Focus of Attention for Human Robot Interaction

Samira Sheikhi, Vasil Khalidov and Jean-Marc Odobez
in: IEEE International Conference on Intelligent Robots and Systems (IROS) - Human Behavior Understanding Workshop (IROS-HBU), 2012

Robot-to-group Interaction in a Vernissage: Architecture & Dataset for Multi-party Dialog

David Klotz, Johannes Wienke, Britta Wrede, Sebastian Wrede, Samira Sheikhi, Dinesh Babu Jayagopi, Vasil Khalidov and Jean-Marc Odobez
in: Proceedings of 5th International Conference on Cognitive Systems, 2012

Robust triphone mapping for acoustic modelling

Milos Cernak, David Imseng and Hervé Bourlard
in: Proceedings of Interspeech, Portland, Oregon, 2012

Socio-Technical Network Analysis from Wearable Interactions

Katayoun Farrahi, Remi Emonet and Alois Ferscha
in: International Symposium on Wearable Computers, 2012

Speaker Diarization and Linking of Large Corpora

Marc Ferras and Hervé Bourlard
in: Proceedings of the IEEE Workshop on Spoken Language Technology, 2012

Speaker Diarization of Meetings based on large TDOA feature vectors

Deepu Vijayasenan and Fabio Valente
in: Proceedings of International Conference on Acoustic, Speech and Signal Processing, 2012

Speaker diarization of overlapping speech based on silence distribution in meeting recordings

Sree Harsha Yella and Fabio Valente
in: INTERSPEECH, Portland, Oregon, USA, 2012

StressSense: Detecting Stress in Unconstrained Acoustic Environments using Smartphones

Hong Lu, Mashfiqui Rabbi, Gokul Chittaranjan, Denise Frauendorfer, Marianne Schmid Mast, Andrew T. Campbell, Daniel Gatica-Perez and Tanzeem Choudhury
in: Ubicomp'12, Pittsburgh, 2012

Structured Sparse Coding for Microphone Array Location Calibration

Afsaneh Asaei, Bhiksha Raj, Hervé Bourlard and Volkan Cevher
in: SAPA-SCALE Conference, The 5th ISCA workshop on statistical and perceptual audition, 2012

Structured Sparse Component Analysis of Compressive Acoustic Measurements

Afsaneh Asaei, Hervé Bourlard and Volkan Cevher
in: Proceeding of International Speech Communication Association, 2012

Sub-Band Based Log-Energy and its Dynamic Range Stretching for Robust In-Car Speech Recognition

Weifeng Li and Hervé Bourlard
in: Proceedings of the 13th Annual Conference of the International Speech Communication Association (InterSpeech), Portland, Oregon, 2012

Supervised and unsupervised Web-based language model domain adaptation

Gwénoél Lecorvé, John Dines, Thomas Hain and Petr Motlicek
in: Proceedings of Interspeech, Portland, Oregon, USA, pages to appear, 2012

Synthetic References for Template-based ASR using Posterior Features

Serena Soldo, Mathew Magimai.-Doss and Hervé Bourlard
in: Proceedings of Interspeech, Portland, Oregon, USA, 2012

Template-based ASR using Posterior features and synthetic references: comparing different TTS systems

Serena Soldo, Mathew Magimai.-Doss and Hervé Bourlard
in: SAPA-SCALE Conference, International Speech Communication Association, 2012

The Good, the Bad, and the Angry: Analyzing Crowdsourced Impressions of Vloggers

Joan-Isaac Biel and Daniel Gatica-Perez
in: Proceedings of AAAI International Conference on Weblogs and Social Media, 2012

The I4U Submission to the 2012 NIST Speaker Recognition Evaluation

Kong Aik Lee, Rahim Saedi, Tawfik Hasan, Tomi Kinnunen, Benoit Fauve, Pierre-Michel Bousquet, Elie Khoury, Pablo Luis Sordo Martinez, Tharmarajah Thiruvaran, Changhuai You, Padmanabhan Rajan, Van Leeuwen, Seyed Omid Sadjadi, Driss Matrouf, Laurent El Shafey, John Mason, Eliathamby Ambikairajah, Hanwu Sun, Anthony Larcher, Bin Ma, Ville Hautamäki, Cernat HaniICI, Billy Braithwaite, Gonzalez-Hautamäki Rosa, Gang Liu, Hynek Boril, Navid Shokouhi, John Hansen, Jean-François Bonastre and Sébastien Marcel
in: NIST Speaker Recognition Conference, 2012

The Idiap Speaker Recognition Evaluation System at NIST SRE 2012
Elie Khoury, Laurent El Shafey and Sébastien Marcel
in: NIST Speaker Recognition Conference, NIST, Orlando, USA, 2012

The INTERSPEECH 2012 Speaker Trait Challenge

Björn Schuller, Stefan Steidl, Anton Batliner, Elmar Nöth, Alessandro Vinciarelli, Felix Burkhardt, Rob Van Son, Felix Wenginger, Florian Eyben, Tobias Bocklet, Gelareh Mohammadi and Benjamin Weiss
in: in Proceedings of INTERSPEECH, 2012

The Mobile Data Challenge: Big Data for Mobile Computing Research

J. K. Laurila, Daniel Gatica-Perez, I. Aad, Blom J., Olivier Bornet, Trinh-Minh-Tri Do, O. Dousse, J. Eberle and M. Miettinen
in: Pervasive Computing, Newcastle, 2012



**Translating English Discourse Connectives into Arabic:
a Corpus-based Analysis and an Evaluation Metric**

Najeh Hajlaoui and Andrei Popescu-Belis

in: Fourth Workshop on Computational Approaches to Arabic Script-based Languages at Proceedings of the Tenth Biennial Conference of the Association for Machine Translation in the Americas (AMTA), 2012

**Unsupervised Activity Analysis and Monitoring algorithms for
Effective Surveillance Systems**

Jean-Marc Odobez, C. Carincotte, Remi Emonet, E. Jouneau, Sofia Zaidenberg, Bertrand Raverra, Francois Bremond and Andrea Grifoni

in: European Conference on Computer Vision, 2012

**Using Crowdsourcing to Compare Document Recommendation
Strategies for Conversations**

Maryam Habibi and Andrei Popescu-Belis

in: RecSys, Recommendation Utility Evaluation (RUE 2012), 2012

**Using KL-divergence and multilingual information to improve ASR
for under-resourced languages**

David Imseng, Hervé Bourlard and Philip N. Garner

in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, pages 4869--4872, 2012

**Using Self-Context for Multimodal Detection of Head Nods
in Face-to-Face Interactions**

Laurent Son Nguyen, Jean-Marc Odobez and Daniel Gatica-Perez

in: Proceedings of the 14th ACM International Conference on Multimodal Interaction, 2012

**Using Sense-labeled Discourse Connectives for Statistical
Machine Translation**

Thomas Meyer and Andrei Popescu-Belis

in: Proceedings of the EACL2012 Workshop on Hybrid Approaches to Machine Translation (HyTra), Avignon, FR, pages 129--138, 2012

**Using Sparse Classification Outputs as Feature Observations
for Noise Robust ASR**

Yang Sun, B. Cranen, Jort F. Gemmeke, Lou Boves, Louis ten Bosch and Mathew Magimai.-Doss

in: Proceedings of Interspeech, 2012

**We are not Contortionists: Coupled Adaptive Learning for Head
and Body Orientation Estimation in Surveillance Video**

Cheng Chen and Jean-Marc Odobez

in: IEEE International Conference on Computer Vision and Pattern Recognition, 2012

*The complete list, abstracts and full texts are available on the Idiap web site at the following address:
<http://publications.idiap.ch>*





Centre du Parc, rue Marconi 19, case postale 592, CH-1920 Martigny
T +41 27 721 77 11 F +41 27 721 77 12 info@idiap.ch www.idiap.ch

