



Impressum

Réalisation: Martina Fellay et Léonore Miauton, équipe gestion de projets

Rédaction: Le fin mot Communication, Martigny

Correction: Bernard Carron

Traduction: Dave Brooks, ELCS

Conception graphique: Atelier Grand, Sierre

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

Impression: CRI, Centre rhodanien d'impression SA, Martigny

Tirage: 1800 exemplaires

SOMMAIRE

Messages

Audace, travail et excellence	
Olivier Dumas, président du Conseil de fondation de l'Idiap	2
Hier pôle national, demain toujours en quête d'excellence	
Prof. Hervé Bourlard, directeur de l'Idiap	3

Recherche

Idiap, carte d'identité: Un institut à taille humaine et au rayonnement international	5
Pôle de recherche national IM2, berceau et accélérateur de croissance	10
Valais*Wallis Digital: La mémoire du Valais, par tous et pour tous!	13
Projet SIIP: La voix, signature au cœur de la sécurité	14
Evaluations internationales: Les travaux de l'Idiap salués sur la scène européenne	15

Réseau

Alliance stratégique Idiap-EPFL: Recherche et formation, fructueux partenariat	17
International Create Challenge: Accélération express pour création d'entreprises	18
L'actualité des start-up: Les trois «K» ont le vent en poupe!	19
Infrastructure informatique: Au service du partage mondial de données	21

Visages

Philip Garner, chercheur senior: «La vie n'est pas sérieuse!»	23
Vasil Khalidov, ingénieur de développement senior: «Je suis impressionné par la qualité de l'infrastructure de l'Idiap»	24
David Imseng, postdoctorant: «Pour l'amoureux du Valais que je suis, travailler à l'Idiap est extraordinaire!»	25
Michel Salamin, professeur de français: «Grâce à mes étudiants, je voyage!»	26
Allées et venues	27
Distinctions	28
Thèses achevées	29

Finances

Compte d'exploitation	31
Sources de financement / Charges / Commentaires	32
Bilan	33

Organisation

Organigramme opérationnel	35
Collaborateurs	36
Conseil de fondation 2013	38
Comité d'accompagnement	39
Principaux partenaires	40

Encarté scientifique

Idiap Research Areas: Human and Media Computing	II
Scientific Progress Report	IV
Selection of Idiap's key scientific achievements in 2013	
Main projects in progress	XVIII
Major publications / Conferences	XXVI

MESSAGE DU PRÉSIDENT

AUDACE, TRAVAIL ET EXCELLENCE



Olivier Dumas, président du Conseil de fondation de l'I'diap

Côtoyer de près un institut de recherche comme l'I'diap provoque parfois des sensations de vertige. Non pas en raison de la complexité des recherches qui y sont menées, ni même grâce à ce splendide escalier qui s'élance dans un large puits de lumière central, au cœur du bâtiment qui abrite, sur quatre étages, l'institut de Martigny!...

Non, le vertige vient d'ailleurs. De cette formidable énergie qui, depuis vingt ans, anime l'I'diap, multipliant son budget par trente, le nombre de ses collaborateurs par cinq, élargissant l'influence de son réseau bien au-delà des frontières nationales et continentales, et le positionnant comme un acteur incontournable dans ses domaines de prédilection qui, eux-mêmes, se situent au cœur des enjeux mondiaux liés aux techniques de l'information. Une adéquation qui permet aujourd'hui à l'Institut d'assurer quelque 60% de son budget grâce aux projets de recherche compétitifs aux niveaux international et national décrochés par ses chercheurs. Adéquation donc, mais aussi excellence.

10 millions de Berne pour 2013-2016

En février 2013, cette excellence était d'ailleurs récompensée par le renouvellement de la confiance et du soutien de Berne pour 2013-2016, pour un montant de 10 millions de francs, soit une augmentation de 3,5 millions par rapport à la dernière période quadriennale. Et dire qu'à sa création, cette subvention s'élevait à 150 000 francs! Que de chemin parcouru... Aujourd'hui, grâce à cette subvention du Secrétariat d'Etat à la formation, à la recherche et à l'innovation (SEFRI), à laquelle il convient d'ajouter les appuis substantiels du Canton du Valais et de la Ville de Martigny, l'I'diap pourra encore solidifier ses bases, lui qui ne cesse de s'élever, mais aussi pérenniser ses activités de recherche et surtout développer encore le transfert de technologie, l'une de ses trois missions de base.

Ce geste fédéral souligne également l'importance de l'alliance stratégique qui unit depuis 2008 l'I'diap et l'EPFL, dans le but de développer les activités scientifiques et académiques d'intérêt commun, mais aussi d'ouvrir de nouvelles voies vers des synergies avec le Pôle EPFL Valais/Wallis, en particulier dans les domaines de l'énergie, de la santé, de la sécurité et de la gestion des risques, dans lesquels l'I'diap est actif depuis plusieurs années.

Des projets pour le Valais, aussi

Car si ses ambitions l'emmènent sur la scène mondiale, l'I'diap décline également ses savoir-faire à l'échelle régionale. Partenaire du projet de Parlement valaisan sans papier et créateur du moteur d'indexation des interventions parlementaires Mediaparl, l'Institut a séduit au printemps 2013 le jury du Bicentenaire de l'entrée du Valais dans la Confédération avec son projet «Valais*Wallis Digital», ou la numérisation de la mémoire collective de notre canton, en partenariat avec la Médiathèque Valais et Migros Valais. (Voir page 13)

Moteur d'innovation, l'Institut participe également à l'incubation d'entreprises sur le site d'IdeArk, où pas moins de douze start-up sont actuellement installées.

Certaines réalisent une progression fulgurante, à l'image de Koemei, qui a signé en 2013 un contrat avec la chaîne Al Jazeera et a été sélectionnée par le World Economic Forum comme «Pionnier en technologie» (*Technology Pioneer*), un titre décerné avant elle à des petites devenues grandes, à l'image de Google. (Voir page 20) En 2013, trois start-up de l'I'diap pouvaient se targuer de figurer dans la liste des 50 «dans lesquelles investir», dressée par le magazine économique *Bilan*.

Oui, l'I'diap donne le vertige. Tant d'audace, d'envie, de travail et d'excellence dans un seul «petit institut», comme le désignent certains médias, peut déstabiliser. Mais perdre pied permet parfois de mieux voir les étoiles.



MESSAGE DU DIRECTEUR

HIER PÔLE NATIONAL, DEMAIN TOUJOURS EN QUÊTE D'EXCELLENCE



Prof. Hervé Bourlard, directeur de l'Idiap

Cette année 2013 a marqué, pour l'Idiap, le terme d'une formidable aventure longue de douze ans: la direction du Pôle national de recherche IM2 (Gestion interactive et multimodale de systèmes d'information). Lorsqu'un tel programme arrive à son terme, on pourrait tomber dans le regret et la nostalgie. Mais on peut également puiser dans cette expérience une force accrue, susceptible de nourrir le futur. Et ce sera sans aucun doute notre cas! Car en douze ans l'Idiap a précisément gagné en force, en réputation nationale et internationale, et en potentiel scientifique et technologique maintenant convoité par de nombreux partenaires.

IM2 a été le berceau de l'Idiap

En 2001, lorsque le Fonds national suisse de la recherche scientifique (FNS), dans sa volonté de maintenir la Suisse parmi les places les plus performantes du monde, choisit parmi plus de 300 soumissions notre institut, celui-ci avait à peine dix ans d'existence! En le retenant, le FNS faisait d'une certaine manière le pari de la jeunesse.

L'Idiap saisit alors l'occasion avec enthousiasme, conscient qu'on ne saurait imaginer conditions plus favorables pour franchir une étape supplémentaire dans sa quête d'excellence. Nous y sommes parvenus puisque, année après année, le comité d'experts mandatés pour analyser nos travaux a renouvelé notre statut de Pôle de recherche national (PRN). Davantage qu'une source de financement supplémentaire de nos projets, nous y avons aussi vu une fantastique reconnaissance ainsi qu'une inépuisable source de motivation. Bien au-delà du financement, il s'agit avant tout de confiance et de motivation, créant un climat unique de créativité et un environnement où nous avons l'impression que tout devient possible!

Oui, IM2 a été le berceau de l'Idiap, une rampe de lancement idéale. L'institut aurait-il sans cela atteint le niveau qui est le sien aujourd'hui? Probablement, au vu de nos domaines de recherche particulièrement porteurs et transversaux, ainsi que de la qualité de nos chercheurs. Mais l'évolution se serait certainement faite à un rythme moins rapide.

10% de financement lié au transfert de technologies

Le chapitre IM2 qui se clôt laissera des traces durables et fécondes, qui nous permettront, demain, de continuer à écrire l'histoire. Car lorsque l'on mène une recherche de qualité, il n'y a pas vraiment de début ni de fin, mais une métamorphose constante et une diversification contrôlée. Dans un monde scientifique qui cède parfois à la tentation de la superficialité et des thèmes à la mode, l'Idiap s'efforce ainsi de maintenir son cap dans ses domaines de recherche privilégiés. Cela a aussi pour effet de maximiser l'impact potentiel au niveau du transfert de technologies qui ne peut réussir que dans un environnement stable et visionnaire sur le long terme, en laissant de côté les «effets de manches» et les «opportunités» parfois bien éphémères. A cet égard, la part de notre budget provenant de projets en lien avec l'industrie, de l'ordre de 10%, s'avère remarquable en comparaison d'institutions souvent prises en exemple, comme le célèbre MIT (Massachusetts Institute of Technology, Cambridge) qui tourne autour des 11%!

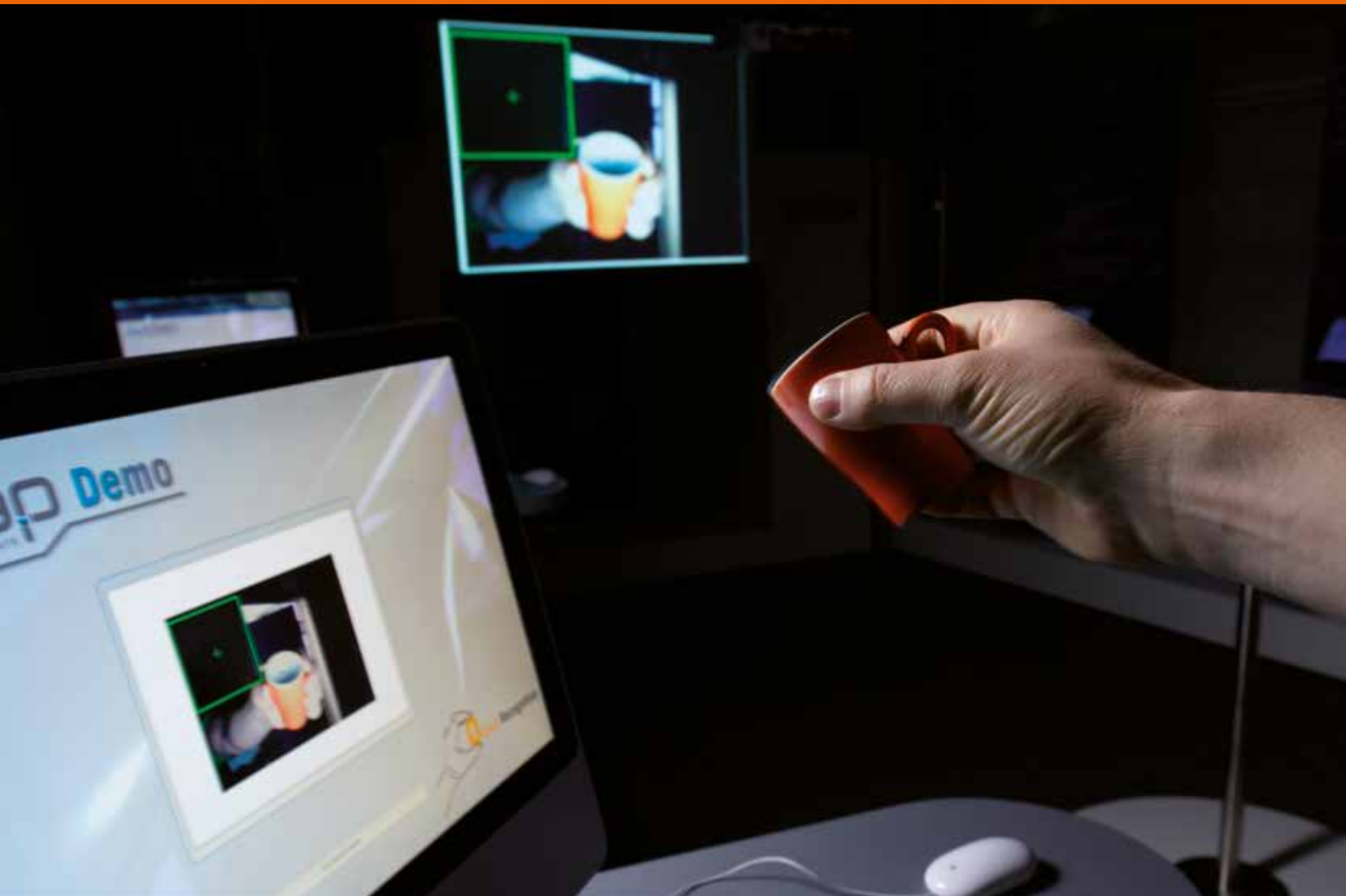
Publications: le défi de la qualité

Notre institut cultive non seulement la qualité de la recherche, mais aussi celle des publications de haut niveau. Dans le monde scientifique, ce sont les revues spécialisées qui en donnent la mesure. En 2013, nous nous étions fixé un objectif particulièrement ambitieux: qu'à trois publications de conférences corresponde une publication dans une revue scientifique. Le but a été si bien atteint que je leur lance un nouveau défi pour 2014: une publication dans un journal spécialisé pour deux comptes rendus de conférences! Bien que ce défi semble hors de portée, nos dernières statistiques montrent que nous avons déjà fait un premier pas dans cette direction.

Nos équipes travaillent sans cesse à relever de nouveaux défis. Je tiens encore à féliciter et à remercier tous les collaborateurs de l'Idiap pour leur engagement. A tous, je souhaite de perpétuer l'esprit de convivialité et de créativité qui nous anime au quotidien et qui crée un climat de travail tellement agréable!

A handwritten signature in blue ink, appearing to read 'H. Bourlard'.

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 de productivité 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 32).

Equipe / En 2013 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 2500 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.
Le transfert de technologies représente 10% du budget de l'Idiap.
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 5 thèmes de recherche et valorisés dans 10 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. Un domaine qui se divise en 5 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

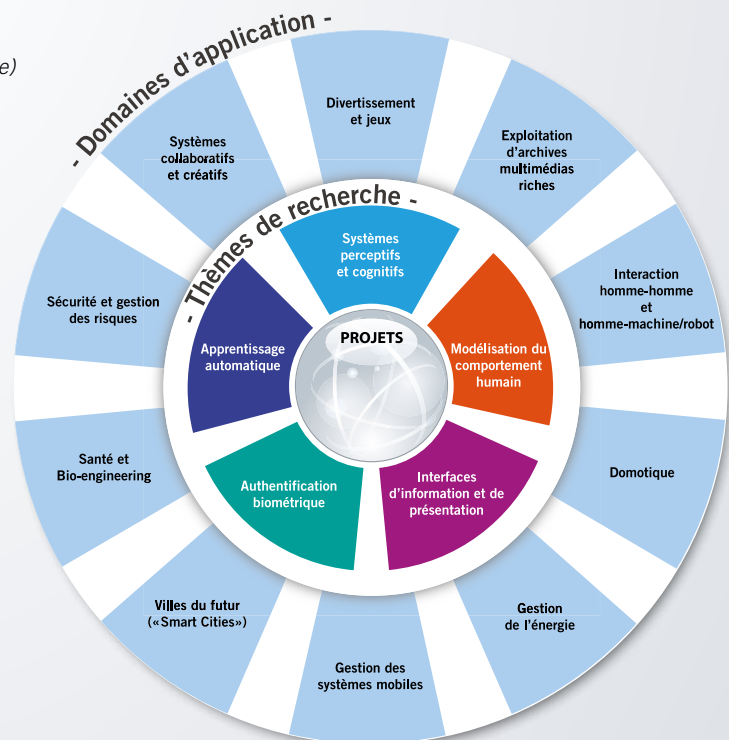
(systèmes d'information multimédias, contextualisation et personnalisation)

4 L'authentification biométrique

(vérification du locuteur, vérification du visage)

5 L'apprentissage automatique

(modélisation statistique, réseaux de neurones, modèles mathématiques)



L'Idiap en chiffres (année 2013)

Ressources humaines

- 1 professeur
- 3 maîtres d'enseignement et de recherche (MER)
- 13 chercheurs permanents et seniors
- 21 postdoctorants
- 40 doctorants
- 17 ingénieurs système et de développement
- 28 stagiaires et visiteurs (moyenne/année)
- 12 collaborateurs administratifs
- 9 titres de docteur décernés
- 40 postes dans les start-up du site IdeArk
- 34 nationalités représentées

Activités scientifiques

- Pôle national de recherche IM2 (Gestion interactive et multimodale de systèmes d'information) de 2001 à 2013
- Participation à 44 programmes de recherche
- Direction de projet dans 12 consortiums
- Participation à la stratégie de développement économique du Canton du Valais à travers le programme The Ark et en particulier la société IdeArk
- 273 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 interlangues, 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

8

**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

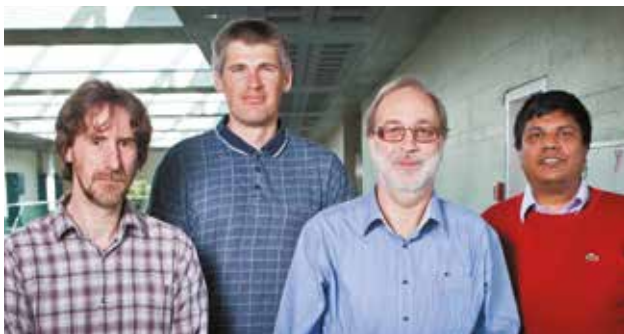
10

DOMOTIQUE



GROUPES DE RECHERCHE

A l'Idiap, les travaux sont menés au sein de huit groupes de recherche. Celui chargé du traitement de la parole et du son est emmené par le directeur de l'institut, Hervé Bourlard, et quatre chercheurs seniors. Les autres groupes sont également pilotés par des chercheurs seniors.



Speech & Audio Processing

Prof. Hervé Bourlard (3^e depuis la gauche), **D^r John Dines** (absent sur la photo), **D^r Phil Garner**, **D^r Petr Motlicek**, **D^r Mathew Magimai Doss**

Le traitement de la parole est depuis de nombreuses années l'un des thèmes de recherche majeurs de l'Idiap. 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'Idiap: 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).



Perception & Activity Understanding

D^r 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 d'objets et 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.



Computer Vision & Learning

D^r François Fleuret

Ce groupe s'intéresse au développement de nouvelles techniques d'apprentissage statistique, essentiellement pour la vision par ordinateur, avec un intérêt particulier pour leurs propriétés computationnelles. Les principales applications sur lesquelles le groupe travaille sont l'analyse automatique d'images, en particulier l'extraction automatique de sémantique, la détection d'objets, et le suivi de personnes et de structures biologiques.



Social Computing

D^r 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, le *crowdsourcing* et le traitement de données urbaines à l'aide de smartphones et de réseaux sociaux mobiles.



Artificial Cognitive Systems D^r 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.



Biometric Person Recognition D^r Sébastien Marcel

En informatique, la biométrie se rapporte à la reconnaissance automatique d'individus par l'utilisation de caractéristiques comportementales et biologiques. Les chercheurs du groupe Biométrie 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 pour certaines modalités biométriques émergentes (électrophysiologie et veines). Le 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.



Applied Machine Learning D^r 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 heuristique-ment. La recherche est motivée par des applications du monde réel, telles

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.



Natural Language Processing D^r Andrei Popescu-Belis

Le groupe Traitement automatique des langues étudie comment l'analyse des textes au niveau sémantique et pragmatique permet d'améliorer les performances dans deux tâches principales, qui sont la traduction automatique et la recherche d'informations. Le groupe s'intéresse en particulier à la façon dont la recom-

mandation d'informations depuis des bases de données multimédias en réseau peut être améliorée en utilisant la structure du réseau et le contenu informationnel de ses nœuds.



PÔLE DE RECHERCHE NATIONAL IM2

BERCEAU ET ACCÉLÉRATEUR DE CROISSANCE

L'Idiap a dirigé le Pôle de recherche national IM2 (Gestion interactive et multimodale de l'information) durant douze ans, de 2001 à 2013. Une expérience dont l'institut tire un bilan extrêmement positif.

A la fin des années 90, le Fonds national suisse de la recherche scientifique (FNS) lance le programme des Pôles de recherche nationaux (PRN). Son but: renforcer la recherche suisse dans des domaines d'importance stratégique tels que la science, l'économie et la société en soutenant des projets de grande envergure.

Quelque 300 propositions sont présentées en réponse à l'appel lancé par le FNS aux instituts de recherche du pays. Une quinzaine sont retenues, dont celle de l'Idiap. L'année de son dixième anniversaire, l'institut donne le coup d'envoi du projet IM2 (Gestion interactive et multimodale de l'information), l'un des 14 PRN financés lors de cette première vague. Un événement exceptionnel pour un institut non universitaire!

Un formidable accélérateur pour l'Idiap

L'aventure va durer douze ans, «davantage si l'on considère que la préparation de notre projet et le processus de sélection se sont déroulés sur près de deux ans, précise Hervé Bourlard, directeur de l'Idiap. Calculée différemment, cette période correspond à trois générations de doctorants.»

A l'heure du bilan, le professeur Bourlard se montre enthousiaste. «IM2 a constitué une expérience géniale, une opportunité formidable! IM2 a été le berceau de l'Idiap. Il a agi comme un accélérateur, permettant à notre institut d'atteindre plus rapidement le niveau qui est le sien aujourd'hui.»

Prestigieuse, intéressante de par le financement qu'elle assure, la direction d'un PNR est aussi fort exigeante. «Décrocher un pôle est difficile, le conserver l'est tout autant!» Chaque année, un comité d'experts extérieurs évalue les travaux réalisés. En outre, le mandat est accordé pour quatre ans, période à l'issue de laquelle il doit être renouvelé, la durée maximale étant fixée à douze ans.

Communauté scientifique élargie

Les retombées positives d'IM2 sont perceptibles dans différents domaines. Sur le plan de la recherche, ce projet a permis d'affiner encore l'expertise de l'institut en matière de gestion interactive et multimodale de l'information (reconnaissance de la parole, indexation de documents, enregistrement et analyse de réunions, etc.). Cette interdisciplinarité, qui représente une occasion unique pour la formation des doctorants, a également contribué à constituer une large communauté scientifique, tissant par exemple des liens entre chercheurs américains et européens. IM2 a largement aidé à accroître la réputation et la visibilité de l'Idiap sur la scène internationale. Sous son égide, le concours International Create Challenge (ICC) a pu être créé en 2012. De plus, IM2 a favorisé l'éclosion de plus de 20 start-up. Nées des technologies de l'Idiap, certaines sont désormais sollicitées par des clients prestigieux et connaissent une belle expansion – c'est le cas notamment de Klewel, Koemei et KeyLemon.

L'influence d'IM2 sur l'Idiap ne s'arrête pas en 2013, Hervé Bourlard en est persuadé. «L'Idiap sort renforcé de cette aventure. Ce que nous avons appris nous servira pour demain. Cela nous encourage à maintenir le cap que nous nous sommes fixé, aussi bien dans nos domaines de recherches de prédilection que dans le transfert de technologies.»



Entre 2001...



... et 2013, l'Idiap a changé d'échelle.



FINAL EVENT, L'HEURE DU BILAN

Une centaine de représentants des mondes scientifique, politique et économique ont participé au «Final Event», manifestation de clôture d'IM2, organisée les 17 et 18 octobre dernier. Deux jours pour tirer le bilan de cette aventure et esquisser des perspectives d'avenir.

Plaisir et fierté habitaient Hervé Bourlard au moment d'accueillir les participants à la manifestation destinée à marquer la clôture d'IM2, cette «aventure très excitante» menée durant douze ans par l'Idiap. «C'est certes la fin d'une histoire, mais aussi le début de nombreuses autres», a notamment déclaré le directeur de l'Idiap aux invités réunis dans les locaux de l'EPFL.

Près de 120 personnes avaient répondu à l'invitation des organisateurs: des chercheurs, des anciens étudiants, des représentants de start-up et des décideurs politiques impliqués dans IM2, mais aussi des invités actifs dans des domaines de recherche proches de ceux de l'Idiap.

Impacts structurels positifs sur la région

Outre la présentation des résultats phares du programme IM2, ces deux jours de rencontres et d'ateliers ont aussi permis aux participants de s'interroger sur les grands projets de recherche menés sur le long terme. Leurs impacts structurels positifs sur une région ou un pays ont ainsi été évoqués par Steve Flinter, responsable du programme scientifique de la Science Foundation Ireland, l'agence irlandaise chargée de financer la recherche dans les secteurs stratégiques des biotechnologies, de l'information et de la communication notamment.

Responsable R&D au sein de Klewel SA, basée à Martigny, Sandy Ingram expliquait pour sa part comment IM2 avait aidé la société à naître et à se développer au sein de l'Idiap. Quant à François Foglia, directeur adjoint de l'institut, il évoquait l'International Create Challenge (ICC), le concours lancé en partenariat avec IM2 et IdeArk, qui soutient de jeunes chercheurs dans leurs projets entrepreneuriaux.

Si elles ont permis de se retourner sur l'expérience IM2, ces deux journées ont également été l'occasion d'envisager le futur. Un avenir que le professeur Bourlard envisage avec confiance: «La recherche de qualité trouve toujours son financement, cela va continuer.»



12 ans de recherche dans un livre

La clôture d'IM2 a encore été marquée par la publication d'un livre, «Interactive Multimodal Information Management». Edité sous la direction d'Hervé Bourlard et d'Andrei Popescu-Belis, l'ouvrage résume certains des accomplissements réalisés durant les douze ans d'IM2. Les différents acteurs du programme commentent leur travail, tandis qu'un chapitre présente le regard que portent sur le PRN des personnalités extérieures à l'institut.

Le livre a été remis à tous les participants de la manifestation de clôture d'IM2.

Interactive Multimodal Information Management
EPFL Press, 2013



INSPIRÉ PAR LES RECHERCHES DE L'IDIAP, VENDU PAR APPLE

Inspiré d'une technologie née à l'Idiap et concrétisé dans le cadre du Pôle de recherche national IM2, le microphone multidirectionnel Microcone est désormais commercialisé sur l'Apple Store.

Ancien chercheur de l'Idiap et patron d'une start-up australienne, Iain McCowan a conclu un accord avec Apple pour vendre son microphone multidirectionnel Microcone sur l'Apple Store. L'objet, qui se présente sous la forme d'une petite pyramide, facilite grandement les réunions regroupant un nombre restreint de participants, en simplifiant leur enregistrement, leur marquage et leur analyse.

Dev-Audio, la société qui a mis au point ce microphone de réunion, est installée en Australie, mais c'est à l'Idiap, où Iain McCowan a travaillé en tant que chercheur entre 2001 et 2005, que les premières recherches sur la «smart meeting room» ont été menées. Fort de son expertise en matière de traitement automatique du langage, l'Idiap a en effet lancé de nombreux projets de recherche sur cette «salle de réunion intelligente», dans laquelle un réseau de micros permettait de déterminer qui était l'intervenant lors d'une discussion de groupe.

Désormais, le Microcone enregistre jusqu'à six locuteurs sur des canaux audio distincts, captant et améliorant automatiquement chaque voix tout en réduisant le bruit de fond. Chaque mot est saisi de manière claire, qu'il s'agisse de réunions professionnelles, d'interviews ou de podcasts.

L'Idiap se réjouit de la brillante réussite d'Iain McCowan, et apprécie de voir que ses recherches, au bout de la chaîne, ont une fois de plus abouti à la production d'un objet qui facilite la vie.



L'ancêtre du Microcone, né à l'Idiap...



... et le produit vendu désormais par Apple.

IM2 EN CHIFFRES

1	source de financement, le Fonds national suisse de la recherche (FNS)
6	partenaires en Suisse: Idiap 2 EPF: Lausanne et Zurich 3 universités: Genève, Fribourg et Berne
12	années de recherche, formation et transfert de technologies
88,1	millions de francs suisses de financement sur douze ans (33,8 millions issus du FNS + 54,3 millions de fonds de contrepartie financés par les partenaires)
1,6	franc engagé par les partenaires pour chaque franc investi par le FNS
2168	publications (plus de 180 par année)
57	participations à des projets en Europe et aux Etats-Unis
23	start-up créées, dont 13 toujours actives
72	partenaires industriels

Association AIM2, prolongement d'IM2

Créée en mars 2012, l'AIM2 (Association pour la gestion interactive et multimodale de l'information) vise à assurer la continuité des réalisations technologiques et économiques d'IM2. Présidée par le directeur adjoint de l'Idiap, François Foglia, l'association poursuit son but en se concentrant sur trois domaines principaux: la recherche, le transfert de technologies et la formation. AIM2 rassemble cinq partenaires académiques: Idiap, EPFL, ETHZ, universités de Genève et de Fribourg.

LA MÉMOIRE DU VALAIS, PAR TOUS ET POUR TOUS!

«Valais*Wallis Digital» est un projet fédérateur et ludique, conçu par l'Idiap pour célébrer le 200^e anniversaire de l'entrée du canton dans la Confédération. Archiver et partager ses souvenirs vont devenir un jeu d'enfant.

Une plateforme internet, une application smartphone, un jeu de cartes, voilà quelques-uns des outils que l'Idiap et ses partenaires vont mettre au service de la mémoire du Valais. Cet ambitieux projet a été retenu par l'Etat du Valais dans le cadre des festivités du Bicentenaire. Nul besoin d'être un geek pour se lancer. Une simple connexion internet suffit pour télécharger et partager ses documents via l'application ou la plateforme. Imaginez un enfant face à son grand-père: «Dis, à quoi ça ressemblait l'école dans ton village quand tu étais petit?» Le témoignage filmé sur smartphone peut être téléchargé sur la plateforme Valais*Wallis Digital, qui, grâce aux technologies de l'Idiap, pourra indexer le contenu du fichier, filtrer l'information en fonction de sa pertinence et créer de nouveaux liens. Des outils de reconnaissance vocale et faciale, à développer lors de projets de recherche liés à Valais*Wallis Digital, pourront repérer des contenus et démultiplier les connexions entre eux. Par exemple, une photo du Cervin intitulée «Zermatt» pourrait rejoindre le stock des images consacrées à la mythique montagne valaisanne. Particuliers, communes, écoles et associations seront invités à déposer leurs archives et documents en 2015 et au-delà. Car cette mémoire collective imaginée par l'Idiap, sera en partie reprise en mains par la Médiathèque Valais et les Archives de l'Etat, institutions partenaires du projet.

Des cartes à collectionner

Pour inciter les Valaisannes et les Valaisans de toutes les générations à participer à ce patrimoine en ligne, un jeu de cartes évoquant 200 événements survenus entre 1815 et 2015 dans le Vieux-Pays sera distribué dans les magasins Migros. Ce jeu a été illustré par les jeunes artistes de l'Ecole professionnelle des arts contemporains (EPAC) de Saxon. Ainsi, en rassemblant des partenaires d'horizons différents autour d'un même défi, l'Idiap a trouvé un parfait moyen de mettre ses technologies au service du public et du canton de façon durable. Il va s'appuyer sur son réseau national et européen pour lancer de nouveaux projets de recherche en lien avec les problématiques de Valais*Wallis Digital (développement d'algorithmes d'analyse d'image avec détection de contenu, par exemple).



François Foglia, directeur adjoint de l'Idiap et porteur du projet «Valais*Wallis Digital»



Damian Elsig, directeur de la Médiathèque Valais, partenaire du projet

«Cette collaboration nous ouvre de nouveaux horizons»

*La Médiathèque Valais a pour mission de conserver et de valoriser le patrimoine audiovisuel valaisan. Qu'allez-vous faire des données de Valais*Wallis Digital?*

Nous ne le savons pas encore. Ce projet est un laboratoire passionnant, dont nous allons découvrir les résultats au fur et à mesure de sa création. Il nous invite à plonger dans le monde des réseaux sociaux, il nous interroge sur le patrimoine électronique de ce canton. Dans un premier temps, nous allons définir les critères de sélection qui permettront un pré-tri automatique par l'Idiap. Dès 2015, nous verrons quel matériel la population valaisanne va déposer et si nous voulons le sauvegarder ou pas. Nous pourrions également utiliser cette fenêtre innovante pour solliciter les Valaisans sur certains thèmes de recherche, par exemple.

Ce projet, c'est aussi l'occasion de collaborer très concrètement avec l'Idiap. Qu'attendez-vous de ces nouvelles technologies dans le domaine du traitement de l'information?

C'est une nouvelle dimension qui s'ouvre à nous, avec un potentiel extraordinaire pour nos activités. Actuellement, le traitement d'une image, de la numérisation au catalogage, nous prend beaucoup de temps. La technologie de l'Idiap pourrait nous aider à monter des archives photographiques beaucoup plus rapides, avec un pré-tri et une pré-description. Je suis donc très intéressé par les suites de cette nouvelle collaboration avec l'Idiap. Le rythme des réseaux sociaux va encore s'accélérer, nous aurons besoin d'aide pour capter et trier les contenus intéressants.



LA VOIX, SIGNATURE AU CŒUR DE LA SÉCURITÉ

L'Idiap est le seul institut suisse engagé dans un vaste projet européen visant à repérer des individus suspects grâce à leur voix. Interpol, qui fournira les données aux chercheurs, figurera parmi les premières organisations à tester le système.

À l'automne 2012, la Commission européenne lançait l'un des derniers appels de son programme-cadre FP7. Parmi la multitude de projets soumis par les centres de recherche les plus renommés, celui de l'Idiap a été retenu pour financement par l'Europe. Son nom: SIIP, pour Speaker Identification Integrated Project. Son champ d'action: la sécurité.

«Le but du projet consiste à mettre au point un nouveau système d'identification de suspects, à partir de leur voix», résume Petr Motlicek, qui collabore au projet aux côtés du directeur de l'Idiap Hervé Bourlard et du chercheur senior Sébastien Marcel.

Fusionner différents programmes en un seul

«Le futur système permettra d'identifier la personne qui parle par rapport non seulement à sa voix, mais aussi à la langue dans laquelle elle s'exprime, à son sexe, à son âge approximatif, et aux informations complémentaires extraites de messages écrits, des réseaux sociaux ou de mails.» Il sera en outre capable de repérer les mots qui reviennent fréquemment dans son discours. «Même si le suspect utilise un code en lieu et place de termes comme «bombe» ou «cocaïne», l'appareil détectera qu'il y a quelque chose de louche», précise le spécialiste en traitement de la parole et traitement audio. «A terme, nous espérons adapter le système aux voix enregistrées par Skype, Google Talk, etc. (VoIP, pour Voix sur IP).»

S'il existe actuellement des applications capables de réaliser l'une ou l'autre de ces opérations, le projet SIIP entend les améliorer et, surtout, fusionner ces différents programmes en un seul. Il devra en outre réussir à établir si celui qui parle essaie de déguiser sa voix (en se pinçant le nez, en mettant la main devant sa bouche, en prenant un accent, etc.) ou d'imiter celle d'une autre personne (générée par un ordinateur). «Ces problèmes ne sont pas résolus à ce jour, pas plus que ceux engendrés par des bruits tels que le passage d'un train ou d'une voiture, qui parasitent le système.»

Des milliers d'heures d'enregistrements à «écouter» par la machine

Les chercheurs de l'Idiap travailleront sur des données fournies par Interpol, organisation internationale de police criminelle qui rassemble 190 pays. Ils disposeront d'un corpus comprenant des milliers d'heures d'enregistrements à traiter à des fins d'identification (trouver une même voix dans différents fichiers) et de vérification (ces deux voix proviennent-elles de la même personne?).

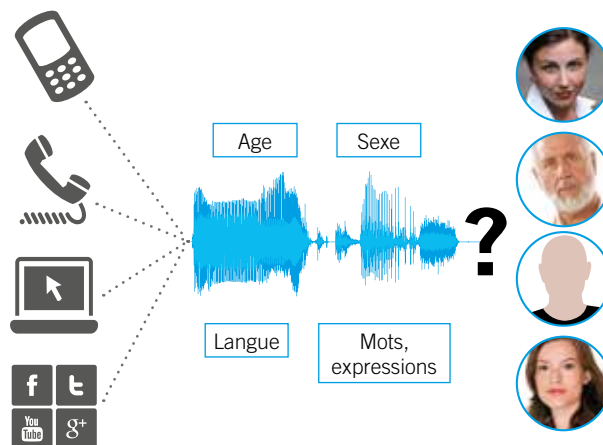
17 partenaires, 10 pays

Le projet SIIP rassemble dix-sept partenaires (universités, entreprises, organisations de police) répartis dans dix pays européens (France, Portugal, Grande-Bretagne, etc.). C'est une entreprise israélienne spécialisée dans le domaine de la sécurité, Verint, qui en assure la coordination.

Le projet devrait débuter au printemps 2014, pour une durée de quatre ans.

«L'analyse de la parole ne débouche pas sur des résultats aussi sûrs que celle des empreintes digitales ou de l'ADN. Mais on atteint grâce à la machine un niveau de confiance qui peut être très élevé.» L'homme n'est pas exclu du processus pour autant. Imaginons que la machine repère trois occurrences d'une même voix, avec des probabilités de 98%, 80% et 20%. A ce stade, ce sont des hommes qui interviendront pour écouter ces trois enregistrements et conclure s'il s'agit du même locuteur ou pas. «L'oreille humaine permet une analyse plus fine. Sauf qu'un homme ne peut pas écouter des milliers d'heures d'enregistrements et que sa mémoire n'est pas infailible: une voix s'oublie vite.»

De la banque à la maison, un système aux applications variées, Petr Motlicek imagine déjà l'utilisation que les banques pourraient faire du futur système, la voix du client devenant le sésame donnant l'accès à ses comptes. La domotique y gagnerait aussi, par exemple avec un téléviseur qui s'allumerait seulement grâce à la voix des parents, et non des enfants.



Le système permettra d'identifier la personne qui parle par rapport à sa voix, à sa langue, à son sexe, à son âge approximatif, et aux informations extraites de messages écrits, des réseaux sociaux ou de mails.

LES TRAVAUX DE L'IDIAP SALUÉS SUR LA SCÈNE EUROPÉENNE

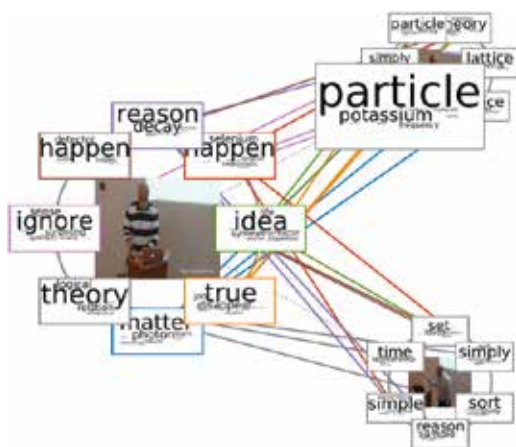
L'Idiap s'est distingué lors de deux campagnes d'évaluation internationales, en proposant de nouvelles approches de l'indexation et de la recommandation de contenus multimédias.

«Le leitmotiv de nos recherches, explique Andrei Popescu-Belis, chercheur senior à l'Idiap, consiste à aller vers une compréhension du contenu multimédia pour améliorer la pertinence des recommandations. Au fond, nous souhaitons permettre à l'utilisateur de «trouver sans chercher». Ainsi, à une personne qui visionne une conférence, nous allons proposer d'autres contenus en lien avec le sujet qui l'a intéressée.»

Andrei Popescu-Belis dirige à l'Idiap le groupe NLP (Natural Language Processing, pour Traitement automatique des langues). «Nous ne sommes pas les seuls à travailler sur ces systèmes de recommandation. Toutes proportions gardées, c'est ce que font les sites de vente en ligne lorsqu'ils suggèrent à leurs clients de nouveaux produits après qu'ils ont effectué un achat. Mais notre approche est nouvelle, d'une part parce qu'elle est appliquée au multimédia, d'autre part, parce qu'elle est basée essentiellement sur l'analyse du contenu de la parole.»

1260 heures d'émissions de la BBC analysées

Deux événements internationaux, organisés en 2013, sont venus confirmer le haut niveau de performance atteint dans ce domaine par les chercheurs de l'Idiap. Lors de la campagne d'évaluation MediaEval 2013, réunissant douze participants, le groupe NLP est arrivé en tête pour la tâche de



Le système de navigation, qui prend en compte à la fois ce que dit l'orateur, les diapositives qu'il projette et ce qu'il écrit au tableau, permet à l'utilisateur de repérer rapidement la thématique qui l'intéresse et de consulter des extraits d'autres conférences traitant de ce même aspect.

création de liens conceptuels, en fournissant plus de 50% de recommandations pertinentes. «Il s'agissait, à partir de 1260 heures d'émissions de la BBC, de repérer des extraits évoquant un sujet donné, par exemple la construction des premiers châteaux forts au Moyen Age. Et, une fois un tel passage trouvé, d'en déterminer d'autres similaires.»

Les chercheurs de l'Idiap, en compagnie de trois autres partenaires du projet européen inEvent, se sont encore distingués à l'ACM Multimedia Grand Challenge 2013, la plus importante conférence mondiale en traitement multimédia. Celle-ci lançait un appel à idées pour segmenter, c'est-à-dire diviser en épisodes, des enregistrements de conférences, ainsi que pour améliorer l'accès à ces segments. «Nous avons proposé de donner une représentation visuelle des segments, et d'utiliser la recommandation de contenus multimédias pour mettre ces extraits en lien avec d'autres contenus similaires.»

Présentation intuitive des résultats

Le graphisme de l'interface, qui allie traits de couleur et nuages de mots clés, permet de repérer très facilement les résultats les plus pertinents. «Une conférence dure généralement entre 30 et 40 minutes. Or, souvent, celui qui la visionne n'est intéressé que par l'un des aspects traités. Ce système de navigation, qui prend en compte à la fois ce que dit l'orateur, les diapositives qu'il projette et ce qu'il écrit au tableau, permet à l'utilisateur de repérer rapidement l'aspect qui l'intéresse et de consulter des extraits d'autres conférences traitant de ce même aspect.»

Le jury de l'ACM s'est montré particulièrement sensible à cette représentation des résultats, dont il a souligné l'intelligence. «Je suis d'autant plus content de la victoire dans ce challenge qu'elle a été obtenue en collaboration avec Klewel, l'une des start-up nées à l'Idiap», se réjouit Andrei Popescu-Belis.

MediaEval 2013 et ACM Multimedia Grand Challenge 2013

Idiap, Groupe NLP	Responsable: Andrei Popescu-Belis, chercheur senior Chidansh Bhatt, postdoctorant, Maryam Habibi, doctorante, Nikolaos Pappas, doctorant
Partenaires	Klewel SA, Martigny (CH) University of Edinburgh (GB) Fraunhofer-Heinrich Hertz Institute (DE)
Soutiens	inEvent, IM2, FNS (via le projet AROLES), Fondation Hasler (via le projet REMUS)



R É S E A U



RECHERCHE ET FORMATION, FRUCTUEUX PARTENARIAT

En février 2013, l'Idiap et l'EPFL renouvelaient leurs vœux en signant pour une nouvelle période de quatre ans un plan de développement commun. Entre le géant académique lausannois et le dynamique institut martignerain, l'alliance est stratégique.

L'Ecole polytechnique fédérale de Lausanne (EPFL), comme son nom l'indique, forme des ingénieurs. L'institut de recherche Idiap, selon la même logique, emploie des chercheurs. Pourtant, l'EPFL mène aussi des activités de recherche, et l'Idiap contribue également à la formation des doctorants: la recherche et la formation ont naturellement des intérêts communs.

Former une relève de qualité

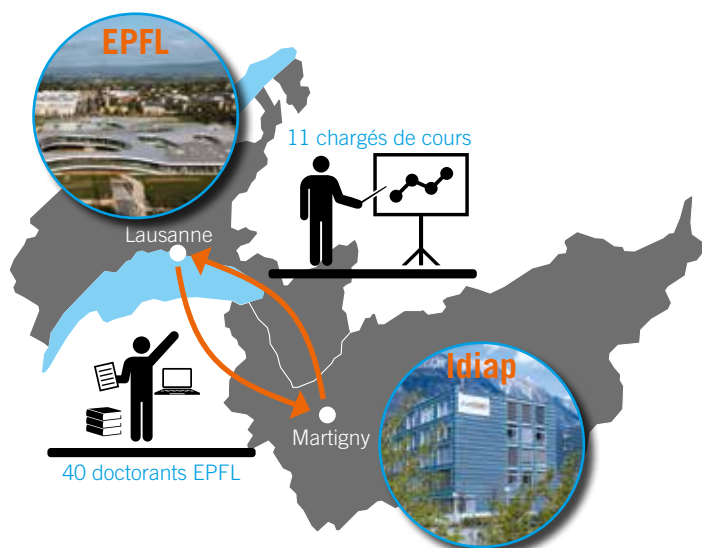
Le savoir n'est pas une matière inerte. Les connaissances de l'homme, en particulier dans les sciences techniques, évoluent au rythme des découvertes réalisées. Quant à la formation, elle gagne à ne pas perdre le contact avec la réalité du terrain, tandis que les chercheurs, en transmettant leur savoir-faire, contribuent à former une relève de qualité.

Forts de ces convictions, l'Idiap et l'EPFL ont signé en 2008 un plan de développement commun assorti, pour l'Idiap, d'une alliance stratégique avec le Domaine des EPF en général, dans une logique d'interdisciplinarité. Le 8 février 2013, les deux institutions signaient pour une nouvelle période de quatre ans, s'engageant ainsi à poursuivre leur collaboration, afin de développer des activités scientifiques et académiques d'intérêt commun.

Reconnaissance académique des chercheurs

Pour l'Idiap, cet accord est important, il lui permet d'offrir à ses chercheurs une reconnaissance académique, ce qui renforce l'attractivité de ses postes de recherche. Aujourd'hui, plusieurs scientifiques de l'institut bénéficient d'un titre EPFL et sont chargés de cours. Ils supervisent également des doctorants qui effectuent leur thèse à l'Idiap. Quant au directeur de l'institut, Hervé Bourlard, professeur EPFL depuis de nom-

breuses années, il dirige le laboratoire LIDAP de la faculté des Sciences et techniques de l'ingénieur (STI) à l'EPFL, antenne de l'Idiap au cœur de l'Ecole polytechnique, où travaillent également les Maîtres d'enseignement et de recherche (MER) François Fleuret, Daniel Gatica-Perez et Jean-Marc Odobez, tous trois chercheurs seniors à l'Idiap.



«Aujourd'hui, explique Hervé Bourlard, environ un tiers de nos collaborateurs sont des doctorants EPFL. Ceux-ci ont la chance de pouvoir œuvrer au sein d'équipes de recherche de niveau mondial, et de notre côté nous bénéficions d'un lien privilégié avec la relève scientifique nationale. D'ailleurs, notre réputation est telle que de nombreux étudiants s'inscrivent désormais en STI pour pouvoir faire leur doctorat chez nous!»

10 millions de subvention fédérale pour 2013-2016

Dans la foulée de son plan de développement commun avec l'EPFL, l'Idiap a vu le Secrétariat d'Etat à la formation, à la recherche et à l'innovation (SEFRI) lui renouveler sa confiance en lui accordant une subvention de 10 millions de francs pour la période 2013-2016, soit 3,5 millions de plus que lors de la précédente période quadriennale.

Cette subvention va permettre de solidifier encore davantage les bases financières de l'institut et de pérenniser ses activités de recherche, de développement et de transfert de technologies.

«Tous les signaux sont au vert pour les quatre années à venir, explique Hervé Bourlard. Sous la condition d'une augmentation de la subvention fédérale, l'Etat du Valais a accepté fin 2013 de relever son soutien de 500 000 francs par an. Une excellente nouvelle pour nous, d'autant que nous avons actuellement deux projets dans le canton, et que nous nous engageons de plus en plus dans le transfert de technologies au niveau régional.»



INTERNATIONAL CREATE CHALLENGE

ACCÉLÉRATION EXPRESS POUR CRÉATION D'ENTREPRISES



La 2^e édition de l'ICC, un programme d'accélération lancé par l'Idiap en 2012, a rencontré un franc succès: 80 dossiers reçus, 3 projets récompensés. Luma7, POWZY et ReMeeting pourraient, demain, changer nos habitudes.

L'appel à projets était publié au printemps 2013: «Turn your idea into a startup!» Conditions de participation: proposer un pro-

jet novateur et de haute qualité scientifique, de préférence dans un domaine de recherche proche de l'Idiap. A la mi-juin, l'institut avait reçu pas moins de 80 dossiers du monde entier, portés par de jeunes gens créatifs, désireux de profiter de cette proposition unique. Dix équipes, regroupant 33 personnes de 15 nationalités, étaient au final retenues et se retrouvèrent le 21 septembre à Martigny. «Vous avez déjà tous gagné, leur lançait alors François Foglia, directeur adjoint de l'Idiap, puisque vous allez enfin pouvoir vous concentrer exclusivement sur l'avancement de vos projets!»

Importants soutiens régionaux

Le 11 octobre, au terme de l'expérience, Luma7, POWZY et ReMeeting obtenaient chacun 10 000 francs, tandis que Luma7 décrochait également un accompagnement de trois ans au sein de l'incubateur IdeArk, d'une valeur approximative de 100 000 francs, sous la condition de créer une entreprise en Valais. Car si le transfert de technologies constitue l'une des missions de l'Idiap, la volonté de l'institut est également de créer de l'emploi dans la région en y faisant naître de nouvelles entreprises. La démarche de l'Idiap, dans cette aventure, bénéficie d'ailleurs du soutien de la Ville de Martigny, de l'Antenne Régions Valais romand, de la fondation The Ark via l'incubateur IdeArk, mais aussi d'Eurofin Ventures, de l'IMD de Lausanne et de l'Hôtel Vatel de Martigny.

L'Idiap félicité par IBM, Nokia, Thalès

Lancé en 2012 par l'Idiap, l'International Create Challenge (ICC) propose en effet aux personnes sélectionnées de vivre trois semaines d'incubation totale, 7 jours sur 7, 24 heures sur 24, tous frais payés, sans avoir à se préoccuper d'autre chose que de faire passer leur idée du projet à la start-up. Durant cette période, les jeunes entrepreneurs bénéficient de l'infrastructure de l'Idiap, de l'accompagnement de ses chercheurs et de son personnel administratif, mais aussi de celui

A noter que, parmi les participants à l'édition 2012, cinq projets sont encore vivants et en cours de développement.

www.createchallenge.org



Luma7, générateur d'idées

Le graphiste Nelio Barros a imaginé Luma7, un générateur automatique de *Mind maps* (cartes d'idées). En tapant par exemple Napoléon, le système proposera une galaxie d'idées allant de l'île d'Elbe à sous-marin en passant par Jeanne d'Arc.

luma7.com



POWZY, publicité mobile non intrusive

Derrière POWZY, une idée: appliquer au business les principes ludiques du jeu, et fidéliser les utilisateurs de smartphone en leur faisant des offres commerciales qui sont en lien avec ce qu'ils sont en train de faire sur leur écran (jouer, chercher leur chemin, etc.).

www.powzy.ch



ReMeeting prend les PV de séances

ReMeeting est une application mobile assortie d'un service de *cloud* personnalisé qui permet d'enregistrer, d'annoter, d'analyser, de retranscrire et d'archiver les séances de travail de manière intelligente. Deux Valaisans sont impliqués dans ce projet.

www.remeeting.com

LES TROIS «K» ONT LE VENT EN POUPE!

On les savait en pleine croissance. Les voilà prêtes à éclore et à donner des fruits. Les jeunes pousses issues des technologies de l'Idiap font partie des «50 start-up dans lesquelles investir», selon le magazine «Bilan». Retour sur une année riche en contrats et en projets pour Klewel, KeyLemon et Koemei.

Les start-up de l'Idiap ont été sélectionnées dans le top 50 des start-up helvétiques du magazine économique *Bilan*. Dans son édition de février 2013, le journal encourage les investisseurs à miser sur ces jeunes entreprises innovantes qui ont besoin de fonds pour se développer. «A condition d'être patient, on peut gagner de 5 à 10 fois sa mise en trois à cinq ans», précise l'article. Des clubs de «business angels» se sont constitués en Suisse romande pour conseiller les investisseurs intéressés à placer intelligemment leurs billes. Durant cette année 2013, les trois spin-off de l'Idiap ont franchi des caps décisifs et emploient désormais, ensemble, une vingtaine de collaborateurs.

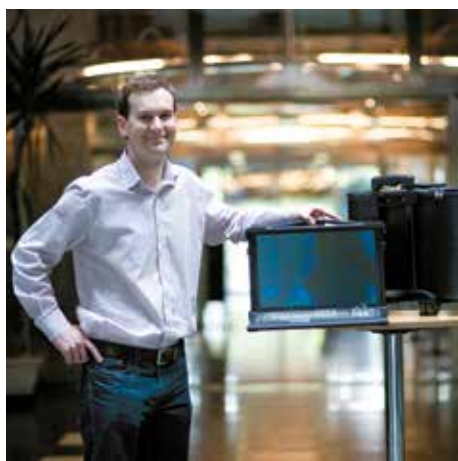


Klewel

Klewel: prêt à servir les pros de l'audiovisuel

Triskel est désormais un produit fini, prêt à conquérir le marché. La version de lancement de cet outil de webcasting a démontré ses atouts auprès de grands clients internationaux ces dernières années. La start-up a d'ailleurs réussi à s'autofinancer grâce à ses prestations de service auprès des organisateurs de congrès. Elle a continué à développer son produit en participant à des projets de recherches national et européen, qui concernent notamment l'intégration de la reconnaissance vocale et de la recommandation automatique de contenu ainsi que la création d'une plateforme web de conférences.

En 2013, Klewel a pu finaliser son Triskel, pour qu'il devienne ce produit «clé en main» facile et rapide d'utilisation depuis l'enregistrement durant l'événement à la publication en ligne. Son fondateur Maël Guillemot se réjouit de cette nouvelle étape: «Nous sommes très satisfaits du résultat. Nous devons aujourd'hui convaincre les agences audiovisuelles d'utiliser notre produit, car ce sont elles qui gèrent la partie technique des congrès. Cela prendra du temps, nous avançons étape par étape. Au lieu de perdre des jours à éditer les conférences, elles pourront synchroniser l'enregistrement audio, la vidéo et les diapositives de la présentation instantanément et publier les contenus en quelques minutes.»



La start-up en bref: 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 régulièrement les enregistrements des conférences d'institutions reconnues telles que Nestlé, la Banque Cantonale Vaudoise ou Shire Pharmaceuticals International et est partenaire du Centre International de Congrès de Genève (CICG).

www.klewel.com





KeyLemon: la confiance des investisseurs

Grâce à une levée de fonds spectaculaire en 2013, KeyLemon décolle sur le marché mondial. La start-up a pu compter sur l'aide de Debiopharm Investment et de Swisscom Venture pour obtenir un financement de 1,5 million de francs. Cet argent lui permet non seulement de développer de nouvelles technologies, mais aussi d'asseoir sa crédibilité. «Cela prouve que l'on croit en nous et que notre produit a un grand potentiel. C'est un signal de confiance très fort pour de futurs investisseurs», se réjouit Gilles Florey, le CEO de KeyLemon, qui prépare une deuxième levée de fonds pour 2014. La start-up a engagé cinq nouvelles personnes pour renforcer le développement et la vente de son produit. Elle est désormais représentée aux Etats-Unis, en Asie et



en Europe. KeyLemon s'investit actuellement dans l'e-learning, l'industrie automobile et les algorithmes pour processeurs. Sa technique de reconnaissances faciale et vocale permettra par exemple de vérifier que c'est le bon étudiant qui passe le bon examen, ou d'alerter un conducteur qui serait en train de s'endormir au volant.

La start-up en bref: 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. 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. www.keylemon.com

KOEMEI

Koemei: la distinction internationale des pionniers

2013 fut une grande année pour Koemei. La start-up active dans la retranscription vocale a été sélectionnée par le World Economic Forum comme Technology Pioneer. Avec cette distinction prestigieuse, elle entre dans «la communauté des pionniers», au même titre que Google (en 2001), Mozilla (2007) ou Dropbox (2012). Pour Temitope Ola, le fondateur de Koemei, figurer sur la liste 2014 des 36 entreprises les plus innovantes de la planète représente un tremplin extraordinaire. En 2013, la start-up s'est d'ailleurs ouverte au marché américain en s'installant à San Francisco. Elle a réussi à lever 0,5 million de francs supplémentaires, portant ainsi ses fonds à 1,3 million. De nouveaux apports financiers sont attendus pour 2014. Koemei vient de lancer une solution destinée à la formation vidéo en ligne, offrant aux apprenants et enseignants des outils de recherche d'information et de collaboration. On connaissait l'énorme potentiel de la retranscription vocale. Investisseurs et clients sont désormais prêts à lui faire confiance.



La start-up en bref: 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. www.koemei.com

AU SERVICE DU PARTAGE MONDIAL DE DONNÉES

Chaque année, 13 terabytes de données quittent l'Idiap à destination de la communauté scientifique mondiale. Si le partage est très courant dans le milieu de la recherche, cela ne se fait pas sans cadre, comme l'explique Frank Formaz, responsable du service informatique de l'institut.

Quelles sont les données échangées par l'Idiap?

Notre institut travaille notamment sur la mise au point de systèmes capables de reconnaître votre visage quand vous allumez votre ordinateur, de retrouver un mot dans une vidéo, etc. L'homme, lui, apprend à reconnaître une image ou un son en emmagasinant, grâce à ses sens, des millions de données. La machine a elle en revanche besoin qu'on lui fournisse ces informations. Les ordinateurs sont donc programmés pour traiter de très grandes quantités de données – son, image, vidéo, texte, etc. Celles-ci intéressent d'autres chercheurs dans le monde, et l'Idiap en obtient aussi d'autres de ses partenaires. De plus, nos scientifiques utilisent et développent des logiciels créés sur mesure pour leurs travaux. Ces logiciels sont aussi échangés dans la communauté scientifique, souvent en version *open source*.

Données et logiciels sont donc échangés en permanence?

C'est effectivement le principe du partage qui prévaut au sein de la communauté scientifique mondiale, ce qui permet naturellement à la recherche d'avancer plus rapidement. Pour nous, il s'agit de faciliter le plus possible ces échanges, tout en respectant le cadre légal. A noter aussi que le volume échangé ne représente qu'une infime part du volume total de données hébergées à l'Idiap.

Pourquoi un cadre légal?

Les données sur lesquelles nous travaillons sont considérées comme personnelles et sensibles au sens de la loi. Il s'agit par exemple d'enregistrements de voix, de vidéos, de photos, etc. Leur utilisation est soumise à la Loi fédérale sur la protection des données (LPD) et régie par des contrats et des licences d'utilisation spécifiques qui nous imposent un contrôle de l'accès et de l'utilisation de ces données. Nous devons donc garantir cet aspect tout en restant le moins contraignants possible pour les chercheurs.

Qu'est-ce que cet échange permanent implique pour vous?

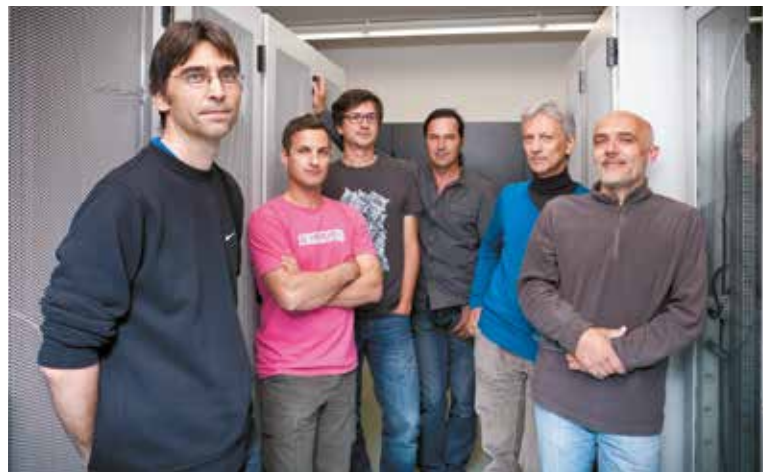
L'Idiap compte une centaine de collaborateurs qui exécutent, ensemble, pas moins de 4000 heures de calcul chaque jour sur ces données. Si 100 personnes veulent faire la même chose au même moment, elles doivent pouvoir le faire en utilisant au mieux les ressources informatiques à disposition. Il s'agit donc pour nous de fournir une infrastructure de stockage et de calcul à la hauteur des besoins, ainsi que des outils pour l'utiliser au mieux.

Concrètement, comment se matérialise cette infrastructure?

A l'Idiap les équipements sont mutualisés afin de permettre à tous de pouvoir accéder au maximum de ressources, tant pour le stockage que pour le calcul. Les serveurs, qui représentent environ 200 ordinateurs et quelque 400 terabytes de stockage, se concentrent dans une dizaine d'armoires informatiques. Cette infrastructure est le cœur des activités de l'institut et son fonctionnement doit être garanti 7 jours sur 7, 24 heures sur 24.

Si une panne d'électricité survient?

Les effets doivent être minimaux. C'est pourquoi tous les systèmes sont redondants. Quoi qu'il arrive, tout est mis en place pour que le système ne s'arrête pas et qu'en cas de problème majeur, les données soient préservées.



Frank Formaz (à gauche) et son équipe offrent aux chercheurs de l'Idiap une infrastructure informatique de haut vol.



VISAGES



PHILIP GARNER, CHERCHEUR SENIOR

«LA VIE N'EST PAS SÉRIEUSE!»

Avec son humour british et ses cheveux en bataille, Phil Garner est une figure incontournable de l'Idiap. A Martigny depuis sept ans, ce chercheur spécialisé dans le traitement de la parole apprécie le côté bon vivant du Valais.

Phil Garner n'a pas d'autre loisir que son travail. Même la guitare qui dort dans son bureau ne réussit pas à le distraire des algorithmes. «Cela fait vingt ans que j'essaie d'en jouer. C'est peine perdue. D'ailleurs, tous les ingénieurs spécialisés dans l'audio/vidéo sont des musiciens ratés!» Car pour ce technicien, l'expression artistique importe moins que le fonctionnement de l'instrument. Dans la banlieue de Manchester où il a grandi, le petit Phil passait son temps à construire et à réparer des objets en compagnie de son père ingénieur en télécommunications. Une passion qui ne l'a jamais quitté. Avec les années et sa formation universitaire, il y a ajouté la dimension scientifique. «Le chercheur essaie de comprendre ce qui existe déjà, alors que l'ingénieur veut résoudre des difficultés techniques. J'ai ces deux facettes en moi. Je passe mon temps à résoudre des problèmes et à en trouver de nouveaux. Mais au fond, cela m'amuse beaucoup, car la vie n'est pas sérieuse, n'est-ce pas?» Le soir, pendant que sa femme pianiste fait ses gammes, il continue de programmer. Il a gardé de sa culture anglaise la cuisine au micro-ondes et le plaisir de sortir dans les bistrot. «Les Valaisans ne sont jamais aussi sympas que lorsqu'ils trinquent et te servent à boire. J'ai découvert le monde des caves et des vigneron. C'est magnifique, même si le vin reste un alcool beaucoup plus dangereux que la bière... Ce qui n'aide pas à améliorer mon français!»

La langue aux chats

Phil avoue volontiers, non sans un brin de culpabilité, son manque d'enthousiasme pour la langue française. Il faut dire qu'il a des circonstances atténuantes. Il a débarqué à Martigny en 2007, après six ans à Tokyo au service de la firme Canon. Il a appris le japonais et continue de le parler au quotidien avec sa femme. A l'Idiap, l'anglais reste la langue commune des chercheurs, y compris lorsqu'il s'agit de partager un verre après le travail. Mais le français l'interpelle dès qu'il met le pied hors de la maison ou du bureau! Bref, pour ce spécialiste de la parole, la communication ne paraît pas toujours simple. «Parfois, je trouve plus reposant de discuter avec mes chats», plaisante-t-il. Grâce à son humour pince-sans-

rire et à son regard aiguisé sur le monde, Phil Garner arrive toujours à ses fins. «Il suffit de dire à un Vaudois: «Vous parlez allemand?» ou à un Zurichois «Vous parlez français?» pour que l'un comme l'autre s'expriment en anglais avec grand plaisir!»

La bonne ambiance

Le Britannique de 44 ans apprécie la vie tranquille et le soleil de Martigny. Ses escapades régulières à l'EPFL de Lausanne pour des heures d'enseignement mises à part, Phil Garner passe le plus clair de son temps dans l'ambiance conviviale de l'institut. «Les moments que nous passons entre collègues à la pause ou en ville sont très importants. La richesse de nos contacts fait la richesse de nos projets.»



Phil Garner: «Je n'ai jamais réussi à jouer convenablement de cette guitare. Comme tous les ingénieurs spécialisés dans l'audio/vidéo, je suis un musicien raté.»



VASIL KHALIDOV, INGÉNIEUR DE DÉVELOPPEMENT SENIOR

«JE SUIS IMPRESSIONNÉ PAR LA QUALITÉ DE L'INFRASTRUCTURE DE L'IDIAP»

L'attractivité d'un institut de recherche de pointe, le charme d'une petite ville. Vasil Khalidov, 30 ans, a trouvé à Martigny les conditions favorables à son épanouissement, tant professionnel que personnel.

Est-ce à Saint-Petersbourg, dont vous êtes originaire, que vous avez appris le français?

Je l'ai appris à Grenoble, où j'ai fait ma thèse en mathématiques appliquées. Cette langue n'est pas indispensable dans le monde de la recherche, mais j'avais envie de m'y initier par intérêt personnel et pour faciliter mon intégration. En arrivant au centre de recherche Inria, j'ai demandé à mes collègues de s'adresser à moi en français. Les gens de l'équipe ont eu suffisamment de patience pour le faire, alors que je n'étais pas capable de dire deux mots dans cette langue!

Leur avez-vous inculqué des rudiments de russe en échange?

Non, mais quand j'ai organisé un voyage à leur intention à Saint-Petersbourg, je leur ai enseigné quelques expressions. Des petites phrases utiles pour éviter de se laisser berner comme certains touristes.

Vous travaillez depuis plus de trois ans à l'Idiap.

Comment l'avez-vous découvert?

A Grenoble, j'ai vu beaucoup de travaux intéressants en provenance de cet institut et j'ai rencontré quelques-uns de ses chercheurs. Grâce à un projet européen auquel participaient l'Inria et l'Idiap, j'ai pu côtoyer Jean-Marc Odobez, qui a été le rapporteur de ma thèse. Quand une possibilité de postdoctorat s'est présentée à l'Idiap, en novembre 2010, je n'ai pas hésité.

Qu'est-ce qui vous a le plus frappé en arrivant à l'institut?

J'ai été particulièrement impressionné par la qualité de l'infrastructure mise à disposition. En plus du niveau des recherches menées à l'Idiap, qui sont à l'avant-garde dans leur domaine, je trouve que l'ambiance y est très conviviale. Sur un plan plus général, j'ai été frappé par la proximité des montagnes, plus grande encore qu'à Grenoble, ce qui est super pour les randonnées. Je n'ai pas choisi l'Idiap en fonction des montagnes... mais je suis très content de me retrouver dans une région montagneuse!

Comment vous sentez-vous à Martigny, où vous avez choisi d'habiter?

Ma femme Olga et moi, nous nous sommes rencontrés à Grenoble; en Russie, il y avait peu de chances que cela arrive puisque nous habitons à 6000 km l'un de l'autre! Aujourd'hui nous nous plaçons beaucoup à Martigny qui, malgré sa taille modeste, est extrêmement bien équipée. On sent l'amour que les habitants portent à leur ville. Je pense par exemple à la manière dont la place Centrale a été remodelée, ou à la Fondation Pierre Gianadda, dont nous visitons régulièrement les expositions.

Vous aviez suivi une école de musique en Russie.

Jouez-vous toujours du piano?

Je continue, oui, la musique me «repose la tête» par rapport à mon travail. J'assiste souvent aux opéras projetés au cinéma de Martigny et je chante dans la chorale TouTanKhanon. Nous sommes douze, trois par registre, et nous interprétons de la variété française, comme Joe Dassin, que je connaissais déjà en Russie, ou Gainsbourg, que j'ai découvert ici. D'une manière générale, je suis toujours passionné par les nouvelles choses, quel que soit le domaine.



Le plaisir de vulgariser

Dans son quotidien de développeur de recherches, Vasil Khalidov peut mettre à profit aussi bien ses connaissances en informatique qu'en mathématiques appliquées. «Cela me plaît beaucoup, car nos recherches doivent servir à améliorer la qualité de la vie.» Dans cette optique, Vasil prend beaucoup de plaisir à vulgariser son travail, par exemple en se livrant à des démonstrations publiques du robot humanoïde NAO. «C'est intéressant de voir à quel point les gens, quel que soit leur âge, sont impressionnés par ses capacités. Je me souviens d'une démonstration à Hérisson sous gazon, le festival pour enfants organisé à Charrat: les petits étaient émerveillés!»

«POUR L'AMOUREUX DU VALAIS QUE JE SUIS, TRAVAILLER À L'IDIAP EST EXTRAORDINAIRE!»

C'est en suivant, à l'EPFL, l'enseignement prodigué par Hervé Bourlard, directeur de l'Idiap, que le Haut-Valaisan David Imseng a trouvé sa voie. Rencontre avec un jeune chercheur passionné.

«J'aime ma vie.» Le constat est résolument positif, à l'image de celui qui le pose. Chez David Imseng, 31 ans, postdoctorant à l'Idiap, l'enthousiasme le dispute à la modestie. Homme et scientifique de passions, il conserve le goût des plaisirs simples.

Améliorer le bien-être de la collectivité grâce à la recherche

Ainsi ce chercheur féru de son domaine, la reconnaissance de la parole, envisage-t-il son travail avant tout comme une manière d'améliorer la vie de ses semblables. «Pensez à ce que la reconnaissance de la parole peut apporter à des personnes sourdes, par exemple! Nous devons toujours avoir le souci de réaliser des nouveautés susceptibles de faciliter la vie des gens; c'est dans ce but que nous travaillons.»

Autre exemple, la manière dont le Haut-Valaisan évoque son doctorat en génie électrique, obtenu en mai 2013: «Ce qui me rend heureux, c'est d'avoir été capable de le faire; ce n'est pas de dire à tout le monde que maintenant je suis «docteur.» D'ailleurs, hormis ses collègues de travail, seuls ses amis proches sont au courant de son nouveau titre.

Un docteur qui joue du fifre... jusqu'aux Etats-Unis

Comme David le précise avec humour, «dans le domaine des fifres et tambours, que vous soyez docteur ne change rien!» Ces corps de musique – une tradition particulièrement vivace dans le Haut-Valais – sont une autre passion de David. Joueur de fifre dans la société de Raron-Sankt German, il s'engage comme instructeur auprès des juniors et siège dans les commissions techniques régionale et fédérale. «Socialement, je suis très Valaisan! C'est aussi pour cette raison que je suis si heureux de pouvoir travailler à l'Idiap et de pratiquer ainsi mon métier de chercheur dans mon pays. Quand je suis en répétition de musique, un soir par semaine, je savoure le plaisir de me retrouver avec des amis et de faire le vide dans ma tête.»

David Imseng participera cet été à la Fête fédérale des tambours et fifres de Frauenfeld avant de s'envoler pour Deep River, près de New York, en compagnie d'une soixantaine d'autres instrumentistes suisses – dont son épouse Erika, elle aussi joueuse de fifre –, invités par leurs homologues américains à donner une série de concerts. «J'aime mon canton

mais j'aime aussi en sortir, parce qu'ici les montagnes limitent un peu l'horizon. Je suis toujours ravi d'aller à la découverte d'autres gens et d'autres cultures, que ce soit pour une occasion festive ou dans le cadre de mon travail, pour suivre des conférences ou pour un stage, comme celui que j'ai effectué à l'International Computer Science Institute (ICSI) Berkeley, en Californie, où j'ai passé dix mois.»

Si son travail devait l'amener à s'exiler, le Haut-Valaisan a deux certitudes: «Ce serait une décision difficile à prendre, et, à terme, je suis sûr que je rentrerai en Suisse.» David Imseng retrouverait du même coup ces montagnes où il apprécie tant effectuer des randonnées: «Me trouver en altitude, avec une belle vue, me donne un peu le même sentiment que lorsque je résous un problème, de la joie et de l'apaisement.»



ReMeeting, un outil qui simplifie les réunions

David Imseng figure parmi les lauréats de l'International Create Challenge (ICC 2013 - voir page 18), dont la deuxième édition s'est déroulée à l'automne 2013. Baptisé ReMeeting, son projet permet non seulement d'identifier les personnes qui prennent la parole lors d'une séance et de déterminer à quel moment elles le font, mais aussi de résumer leurs propos, de repérer les mots clés utilisés et d'inscrire automatiquement toute nouvelle échéance dans l'agenda des différents participants. Le prix décerné par l'ICC – 10 000 francs – lui permettra de poursuivre ses recherches et de monter une start-up pour en assurer le transfert technologique.



MICHEL SALAMIN, PROFESSEUR DE FRANÇAIS

«GRÂCE À MES ÉTUDIANTS, JE VOYAGE!»

Michel Salamin dispense depuis douze ans des cours de français aux chercheurs de l'Idiap. Avec un enthousiasme et un engagement qui dépassent largement le cadre du vocabulaire et de la grammaire.

«Ce qui me passionne dans mon travail ici, ce sont les gens!» Tel est le secret de Michel Salamin, la raison pour laquelle, en douze ans de collaboration avec l'Idiap, il n'a jamais ressenti la moindre lassitude. «Mes étudiants viennent de l'autre bout du monde. Quand ils me parlent de leur culture et de leur pays, j'ai l'impression d'y être!»

C'est peu dire que le dynamique Martignerain apprécie les quelques heures qu'il passe chaque semaine à l'institut. D'ailleurs, lorsque l'Ecole de commerce et de culture générale de Martigny, où il enseigne, lui a proposé de compléter son horaire, il a décliné l'offre, afin de pouvoir continuer à accompagner les chercheurs de l'Idiap. «C'est vraiment le contact avec ces gens formidables qui m'a fait rester ici. J'admire l'énergie et l'application qu'ils mettent à apprendre le français, alors qu'ils n'utilisent pas cette langue dans le cadre de leur travail... La situation est cocasse, quand on y pense: ces chercheurs ont des QI extraordinaires et, moi, je leur apprend à compter en français!»

Priorité à l'expression orale

Titulaire d'une licence en lettres et d'un diplôme de maître de gymnase lui permettant d'enseigner la littérature française ainsi que le français et l'italien, le prof adapte ses cours à son public en se concentrant sur l'expression orale. «Je cherche avant tout à transmettre des choses simples, pratiques, qui facilitent leur quotidien.» Tenir une petite conversation, faire ses courses, arriver à communiquer avec une régie immobilière ou avec un coiffeur. «Je me suis aperçu que mes étudiants éprouvent un réel stress s'ils ne savent pas expliquer la coupe de cheveux qu'ils veulent.»

L'attention que porte Michel Salamin à l'intégration de ses protégés ne se limite pas aux cours de français et de conversation, ni même au rôle d'écrivain public qu'il endosse lorsque l'un

d'eux doit rédiger un courrier. Conscient que la langue seule ne suffit pas à garantir l'adaptation, il organise nombre de sorties auxquelles sont conviés non seulement les chercheurs, mais aussi leurs conjoints.



Brisolée, sortie en luge, carnaval, etc.

Au programme des réjouissances: via ferrata Farinetta, brisolée, Foire du Valais, sortie en luge au clair de lune avec fondue, carnaval du Bourg... «Ce n'est rien d'exceptionnel, je fais juste avec eux ce que fait le Valaisan moyen. Je suis content de leur montrer qu'on peut avoir une vie agréable ici, où la nature est si proche.» Il l'avoue, il lui arrive de «gronder» gentiment ses étudiants

s'ils manifestent l'intention de vivre à Montreux ou à Lausanne plutôt qu'en Valais.

Au fil du temps et des activités partagées, de solides liens d'amitié se sont tissés avec des étudiants, dont certains ont fini par mieux connaître que lui la région. Contrat rempli pour celui qui se donne la mission d'être une sorte de «Monsieur Intégration».

Une véritable tour de Babel

Avec plus de 30 nationalités représentées en son sein et presque autant de langues maternelles, l'Idiap fait figure de véritable tour de Babel. Les chercheurs et chercheuses de l'institut proviennent des pays suivants: Allemagne, Australie, Belgique, Brésil, Chine, Costa-Rica, Espagne, Etats-Unis, France, Grande-Bretagne, Grèce, Hongrie, Inde, Indonésie, Iran, Italie, Kenya, Liban, Lituanie, Macédoine, Mexique, Pakistan, Portugal, Slovaquie, République tchèque, Roumanie, Russie, Suisse, Turquie et Vietnam.

ALLÉES ET VENUES

En 2013, on enregistre 19 départs pour 14 arrivées. Parmi les nouveaux arrivants, on compte deux ingénieurs de développement, un professeur invité, quatre postdoctorants, six doctorants et un collaborateur relation industrielle.

ILS SONT ARRIVÉS EN 2013

Prénom, nom, fonction, pays d'origine

Sara Bahaadini Beigy, doctorante, Iran	Renato Martins, ingénieur de développement, Portugal
Gulcan Can, doctorante, Turquie	Srinivas Narayanan, professeur invité, Etats-Unis
Samuel Gaist, ingénieur de développement, Suisse	James Newling, doctorant, Angleterre
Rui Hu, postdoctorante, Chine	Francisco Pinto, postdoctorant, Portugal
Cijo Jose, doctorant, Inde	Blaise Potard, postdoctorant, France
Srikanth Madikeri, postdoctorant, Inde	Marzieh Razavi, doctorante, Iran
Laurent Marti, relation industrielle, Suisse	Matthias Vanoni, doctorant, France

ILS REPARTENT

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

Céline Aymon Fournier, relations publiques, Suisse, 2004, Valais/Wallis Promotion
John Dines, chercheur, Australie, 2003
Charles Dubout, doctorant, Suisse, 2009
Rémi Emonet, postdoctorant, France, 2010, Université Jean Monnet, Saint-Etienne, France
Arjan Gijssberts, postdoctorant, Pays-Bas, 2011
Najeh Hajlaoui, postdoctorant, France, 2011
Dinesh Babu Jayagopi, postdoctorant, Inde, 2007, International Institute of Information Technology (IIIT), Bangalore, Inde
Samuel Kim, postdoctorant, Corée du Sud, 2011
Riwal Lefort, postdoctorant, France, 2011, Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen (GREYC), France
Adolfo Lopez Mendez, postdoctorant, Espagne, 2012, start-up fezoo
Laurent Marti, relation industrielle, Suisse, 2013
Renato Martins, ingénieur de développement, Portugal, 2013, Koemei
Edgar Roman Rangel, doctorant, Mexique, 2008, Université de Genève
Dairazalia Sanchez-Cortes, doctorante, Mexique, 2009
Nicolae Suditu, doctorant, Roumanie, 2008, SwissLitho AG
Romain Tavenard, doctorant, France, 2011, Université de Rennes 2, France
Tatiana Tommasi, doctorante, Italie, 2008
Alessandro Vinciarelli, chercheur senior, Italie, 1999
Majid Yazdani, doctorant, Iran, 2008, Université de Genève



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 2013, le prix de la Recherche a été attribué à **Laurent El Shafey** pour la qualité de ses publications et l'excellence de ses recherches traitant de l'apprentissage automatique reproductible appliqué à la reconnaissance faciale et vocale. Celui de la publication a été remis à **Leonidas Lefakis** pour son excellent article scientifique intitulé «Reservoir Boosting: Between Online and Offline Ensemble Learning».



Laurent El Shafey



Leonidas Lefakis

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.

Vasil Khalidov, Florence Forbes et Radu Horaud

Best Paper Award Nomination at the IEEE Int. Workshop on Multimedia Signal Processing (MMSp) held in Pula (Sardinia, Italy), septembre/octobre 2013

Alignment of Binocular-Binaural Data Using a Moving Audio-Visual Target

Alexander Heili et Jean-Marc Odobez

Best Paper Award at the Workshop on Performance Evaluation of Tracking and Surveillance (PETS), janvier 2013

Parameter Estimation and Contextual Adaptation for a Multi-Object Tracking CRF Model

C. Bhatt, Andrei Popescu-Belis, Maryam Habibi, Sandy Ingram, F. McInnes, S. Masneri, Nikolaos Pappas et O. Schreer

Winner of the Grand Challenge on Lecture Segmentation and Annotation at the ACM Multimedia 2013, Barcelona, octobre 2013

Multi-factor Segmentation for Topic Visualization and Recommendation: the MUST-VIS System

C. Bhatt, Nikolaos Pappas, Maryam Habibi et Andrei Popescu-Belis

Winner of the hyperlinking subtask at the MediaEval 2013, Barcelona, octobre 2013

Idiap at MediaEval 2013: Search and Hyperlinking Task

THÈSES ACHÉVÉES

Neuf étudiants ont achevé leur thèse en 2013: Nicolae Suditu, Gelareh Mohammadi, Dairazalia Sanchez-Cortes, Tatiana Tommasi, Joan-Isaac Biel, Afsaneh Asaei, David Imseng, Charles Dubout et Majid Yazdani

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

Tatiana Tommasi, 2013

Directeurs de thèse: Prof. Hervé Bourlard, D^r Barbara Caputo

Membres du jury: Prof. Dario Floreano, Prof. Jean-Philippe Thiran, Prof. Jim Little, D^r Vittorio Ferrari

- ***Mining Conversational Social Video***

Joan-Isaac Biel, 2013

Directeur de thèse: D^r Daniel Gatica-Perez

Membres du jury: Prof. Sabine Susstrunk, Prof. Pascal Frossard, D^r Bernardo Huberman, D^r Mathias Mehl

- ***Similarity Learning Over Large Collaborative Networks***

Majid Yazdani, 28 février 2013

Directeurs de thèse: Prof. Hervé Bourlard, D^r Andrei Popescu-Belis

Membres du jury: Prof. Karl Aberer, D^r Samy Bengi, D^r Jean-Cédric Chappelier, Prof. Andreas Krause

- ***Computational Methods for Audio-Visual Analysis of Emergent Leadership in Teams***

Dairazalia Sanchez-Cortes, 1^{er} mars 2013

Directeur de thèse: D^r Daniel Gatica-Perez

Membres du jury: Prof. Jean-Philippe Thiran, Prof. Pierre Dillenbourg, Prof. Fabio Pianesi,
Prof. Marianne Schmid Mast

- ***Model-based Sparse Component Analysis for Multiparty Distant Speech Recognition*** Afsaneh Asaei, 21 mars 2013

Directeur de thèse: Prof. Hervé Bourlard

Membres du jury: Prof. Jean-Philippe Thiran, Prof. Christian Jutten, Prof. Dietrich Klakow, Prof. Pierre Vandergheynst

- ***Adaptive Relevance Feedback for Large-scale Image Retrieval***

Nicolae Suditu, 19 avril 2013

Directeur de thèse: D^r François Fleuret

Membres du jury: Prof. M. Grossglauser, D^r M. Ferecatu, D^r S. Marchand-Maillet, Prof. J.-Ph. Thiran

- ***Automatic Personality Perception: Inferring Personality Traits from Nonverbal Vocal Behavior***

Gelareh Mohammadi, 17 juin 2013

Directeurs de thèse: Prof. Hervé Bourlard, D^r Alessandro Vinciarelli

Membres du jury: D^r Jean-Marc Vesin, Prof. Jean-Philippe Thiran, D^r Fabio Pianesi, Prof. Björn Wolfgang Schuller

- ***Multilingual speech recognition: A posterior based approach***

David Imseng, 20 juin 2013

Directeur de thèse: Prof. Hervé Bourlard

Membres du jury: Prof. J.-Ph. Thiran, Prof. N. Morgan, Prof. T. Schultz, D^r J.-M. Vesin

- ***Object Classification and Detection in High Dimensional Feature Space***

Charles Dubout, 17 décembre 2013

Directeur de thèse: D^r François Fleuret

Membres du jury: Prof. Mark Pauly, Prof. Pascal Fua, Prof. Gilles Blanchard, Prof. Frédéric Jurie



FINANCES



COMPTE D'EXPLOITATION

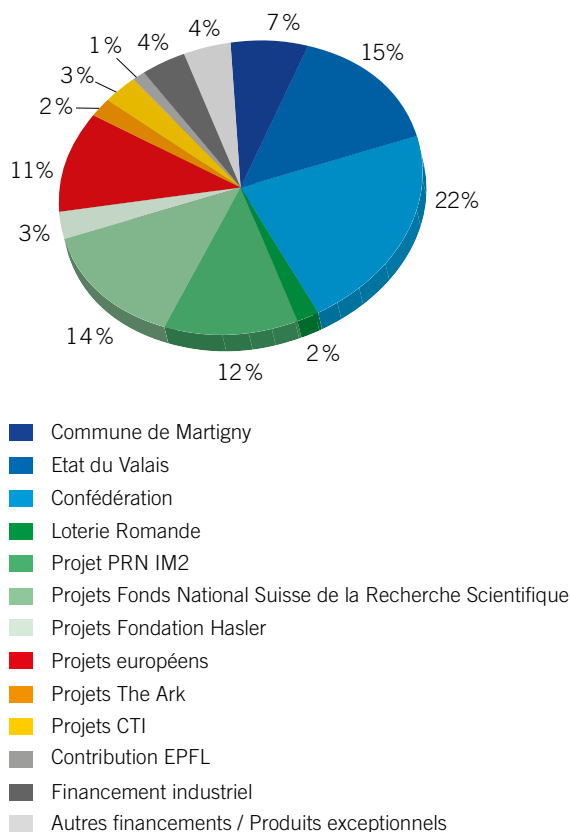
(En francs suisses)

	2012	2013	%
PRODUITS			
Commune de Martigny	650 000	700 000	6,58%
Etat du Valais	1 120 000	1 620 000	15,23%
Confédération	2 393 000	2 400 000	22,56%
Loterie Romande	250 000	250 000	2,35%
Projet PRN IM2	764 071	1 273 628	11,97%
Projet Fonds National Suisse de la Recherche Scientifique	1 203 153	1 489 392	14,00%
Projets Fondation Hasler	347 913	314 717	2,96%
Projets européens	2 219 526	1 132 694	10,65%
Projets The Ark	55 628	189 288	1,78%
Projets CTI	251 168	328 864	3,09%
Contribution EPFL	72 000	72 000	0,68%
Financement industriel	342 766	473 367	4,45%
Autres financements / Produits exceptionnels	669 449	394 213	3,71%
TOTAL DES PRODUITS	10 338 674	10 638 163	100,00%
CHARGES			
Frais de personnel	7 608 490	7 709 161	72,5%
Formation et déplacements	528 042	458 200	4,3%
Partenaires externes	92 469	93 062	0,9%
Informatique: matériel et maintenance	249 146	241 310	2,3%
Frais administratifs	178 872	193 520	1,8%
Promotion et communication	77 007	60 413	0,6%
Loyer et charges	889 306	817 621	7,7%
Amortissement	281 777	184 563	1,7%
Charges exceptionnelles	421 340	97 881	0,9%
Provisions	—	750 000	7,1%
TOTAL DES CHARGES	10 326 449	10 605 731	99,7%
RÉSULTAT D'EXPLOITATION	12 225	32 432	0,3%

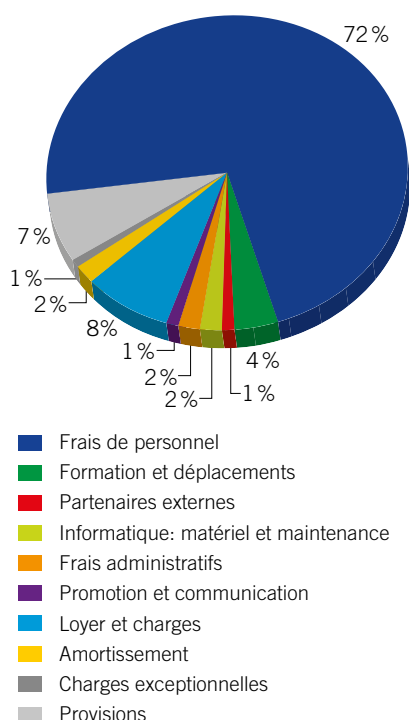


SOURCES DE FINANCEMENT / CHARGES / COMMENTAIRES

Répartition des sources de financement



Répartition des charges



Commentaires sur les comptes 2013

Pour la deuxième année consécutive, la barre mythique des 10 millions est franchie. L'Idiap poursuit sur sa lancée et boucle ses comptes sur un bénéfice de 32432 francs.

La part des subventions publiques se situe toujours dans l'objectif stratégique avec, cette année, un pourcentage de 44,4%. Le canton du Valais et la ville de Martigny renforcent leur contribution financière de 500000 et 50000 francs, respectivement.

Le volume des projets financés par l'UE a diminué de façon importante. C'est une conséquence directe du passage entre le 7^e programme cadre et H2020, transition durant laquelle l'institut a fait face à moins d'appels à propositions, par conséquent, moins d'occasions de financement. Le PRN IM2 et le développement continu des relations industrielles compensent en partie ce manque à gagner.

Subventions Confédération, Canton, Commune

(En milliers de francs suisses)

ANNÉE	2011	2012	2013	2014*
Confédération	2357	2393	2400	2550
Canton	900	1120	1620	1720
Commune	650	650	700	700

* Budget

BILAN

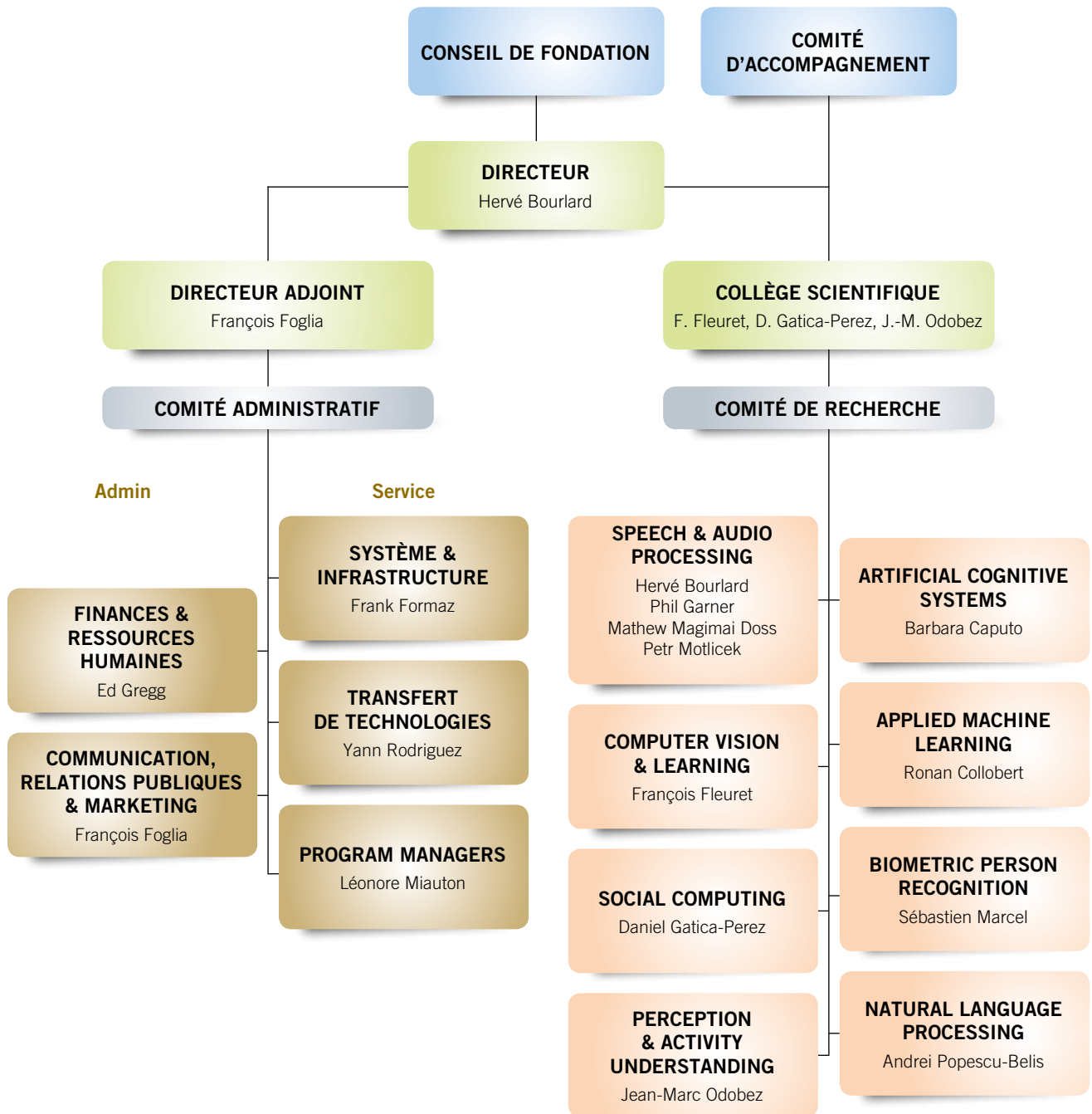
(En francs suisses)

	31.12.2012	31.12.2013
ACTIFS		
Liquidités	2 104 427,07	3 614 085,58
Débiteurs	287 382,35	220 492,98
Actifs de régularisation et divers	2 536 046,69	985 382,87
TOTAL ACTIFS CIRCULANTS	4 927 856,11	4 819 961,43
Mobilier et matériel informatique	504 960,99	340 658,79
Participations financières	10 000,00	10 000,00
TOTAL ACTIFS IMMOBILISÉS	514 960,99	350 658,79
TOTAL ACTIFS	5 442 817,10	5 170 620,22
PASSIFS		
Créanciers	365 838,69	307 725,05
Passifs de régularisation	3 918 238,84	2 921 723,22
Provisions	0,00	750 000,00
TOTAL FONDS ÉTRANGERS	4 284 077,53	3 979 448,27
Capital	40 000,00	40 000,00
Réserve spéciale	1 000 000,00	1 000 000,00
Résultat reporté	106 513,11	118 739,57
Résultat net	12 226,46	32 432,38
TOTAL FONDS PROPRES	1 158 739,57	1 191 171,95
TOTAL PASSIFS	5 442 817,10	5 170 620,22



ORGANISATION







COLLABORATEURS

Personnel scientifique

Prénom, nom, fonction, pays d'origine, (domicile), année d'entrée en fonction

Oya Aran Karakus, chercheuse, Turquie, 2009
Afsaneh Asaei, doctorante, Iran, 2008
Sara Bahaadini Beigy, doctorante, Iran, 2013
Chindansh Bhatt, postdoctorant, Inde, 2012
Joan-Isaac Biel, doctorant, Espagne, 2008
Hervé Bourlard, directeur, Suisse, Saxon, 1996
Gulcan Can, doctorante, Turquie, 2013
Olivier Canévet, doctorant, France, 2012
Barbara Caputo, chercheuse senior, Italie, 2005
Ivana Chingovska, doctorante, Macédoine, 2011
Ronan Collobert, chercheur, France, 2010
Trinh-Minh-Tri Do, postdoctorant, Vietnam, 2009
Elie El Khoury, postdoctorant, Liban, 2011
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
Manuel Günther, postdoctorant, Allemagne, 2012
Maryam Habibi, doctorant, Iran, 2011
Alexandre Heili, doctorant, France, 2010
Pierre-Edouard Honnet, doctorant, France, 2012
Rui Hu, postdoctorante, Chine, 2013
David Imseng, doctorant, Suisse, Rarogne, 2009
Cijo Jose, doctorant, Inde, 2013
Dinesh Babu Jayagopi, postdoctorant, Inde, 2007
Ilja Kuzborskij, doctorant, Lituanie, 2012
Alexandros Lazaridis, postdoctorant, Grèce, 2012
Rémi Lebret, doctorant, France, 2012
Leonidas Lefakis, doctorant, Grèce, 2010
Joël Legrand, doctorant, France, 2012
Srikanth Madikeri, postdoctorant, Inde, 2013
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
Srinivas Narayanan, professeur invité, USA, 2013
James Newling, doctorant, Angleterre, 2013
Laurent Nguyen, doctorant, Suisse, Lausanne, 2011
Jean-Marc Odobez, chercheur senior, France / Suisse, Clarens, 2001
Pedro Oliveira Pinheiro, doctorant, Brésil, 2012
Dimitri Palaz, doctorant, Suisse, Martigny, 2011
Nikolaos Pappas, doctorant, Grèce, 2012
Novi Patricia, doctorante, Indonésie, 2012
Francisco Pinto, postdoctorant, Portugal, 2013
Andrei Popescu-Belis, chercheur senior, France / Roumanie, 2007
Blaise Potard, postdoctorant, France, 2013
André Rabello Dos Anjos, postdoctorant, Brésil, 2010
Ramya Rasipuram, doctorante, Inde, 2010
Marzieh Razavi, doctorante, Iran, 2013
Lakshmi Saheer, doctorante, Inde, 2008
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
Raphaël Ullmann, doctorant, Suisse, Lausanne, 2012
Matthias Vanoni, doctorant, France, 2013
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
Samuel Gaist, ing. de développement, Suisse, Ardon, 2013
Salim Kayal, ing. de développement, Suisse, Vevey, 2011
Vasil Khalidov, ing. de développement, Russie, 2010
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, pays d'origine, (domicile), année d'entrée en fonction

Céline Aymon Fournier, relations publiques, Suisse, Fully, 2004
Antoine Dorsaz, aide-comptable, Suisse, Fully, 2012
Christophe Ecoeur, assistant financier, Suisse, Collombey, 2010
Fellay Martina, junior program manager, Autriche, 2012
François Foglia, adjoint du directeur, Suisse, Saxon, 2006
Edward-Lee Gregg, responsable financier, Etats-Unis, 2004
Léonore Miauton, senior program manager, Suisse, Chexbres, 2012
Sylvie Millius, assistante administrative, Suisse, Vétroz, 1996
Yann Rodriguez, relations industrielles, 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, pays d'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.

Tauseef Ali, Pakistan, University of Twente (NLD)
Krysten Ansermoz, Suisse, EPFL, Lausanne
Afroze Ibrahim Baqapuri, Pakistan, EPFL, Lausanne
Julia Biolley-Franzetti, Suisse, Université de Lausanne
Maxime Bourlard, Suisse, Université de Neuchâtel
Aleksandra Cerekovic, Croatie, University of Zagreb (HRV)
Simon Coquoz, Suisse, EPFL, Lausanne
Pranay Dighe, Inde, Indian Institute of Technology, Kanpu (India)
Matthieu Duval, France, Polytech Nice-Sophia, Sophia Antipolis (FRA)
Alexandre Ferchaud, Suisse, Ecole de commerce, Sion
Manon Fournier, Suisse, Lycée-Collège de l'Abbaye, St-Maurice
Javier Galbally, Espagne, Autonomous University of Madrid (ESP)
Hainan Hu, Chine, EPFL, Lausanne
Nicolas Jacquemin, Suisse, EPFL, Lausanne

César Laurent, Suisse, EPFL, Lausanne
Sharid Loaiciga Sanchez, Costa Rica, Université de Genève
Alvaro Marcos, Espagne, Université de Alcalá (ESP)
Rahesh Mehta, Inde, Tampere University of Technology (FIN)
Braid Meyer, Suisse, ISME, St-Gall
Xingyu Na, Chine, Beijing Institute of Technology, Beijing (CHN)
Billy Okal, Kenya, Jacobs University, Bremen (DEU)
Naoki Peter, Suisse, EPFL, Lausanne
Jérémy Rappaz, Suisse, EPFL, Lausanne
Sandrine Revaz, Suisse, Université de Fribourg
Theresa Schwery, Suisse, EPFL, Lausanne
Abhishek Sharma, Inde, CNRS Paris (FRA)
Kelly Tiraboschi, Italie, Lycée-Collège de l'Abbaye, St-Maurice
Sarah Anna Vomsattel, Suisse, Université de Lausanne

Visiteurs

Prénom, nom, pays d'origine, 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.

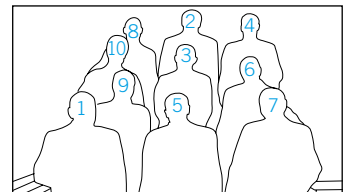
Abhijit Zipru Bendale, Inde, University of Colorado (USA)
Abhishek Dutta, Népal, University of Twente (NLD)
Adrian Penate Sanchez, Espagne, Polytechnic University of Catalonia, Barcelone (ESP)



CONSEIL DE FONDATION 2013



Le Conseil de fondation assume la responsabilité de 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.



1 M. Olivier Dumas, Président
Ancien président de la Ville de Martigny

M. Jean-Daniel Antille, Vice-président (*absent sur la photo*)
Responsable de l'Antenne Régions Valais romand

Prof. Philippe Gillet, Vice-président (*absent sur la photo*)
Vice-président pour les affaires académiques
Ecole polytechnique fédérale de Lausanne (EPFL)

Représenté par:

2 Prof. Karl Aberer
Vice-président pour les systèmes d'information
Ecole polytechnique fédérale de Lausanne (EPFL)

3 M. Marc-André Berclaz
Président du Comité directeur de la HES-SO Valais Wallis

4 M. Stefan Bumann
Chef du service de la formation tertiaire, Etat du Valais

M^{me} Adrienne Corboud Fumagalli (*absente sur la photo*)
Vice-présidente pour l'innovation et la valorisation
Ecole polytechnique fédérale de Lausanne (EPFL)

M. Marc-Henri Favre (*absent sur la photo*)
Président de la Ville de Martigny

5 Prof. Christian Pellegrini
Doyen du Conseil de fondation de l'Idiap
Professeur honoraire à la Faculté des sciences de l'Université de Genève

6 M. Jean-René Germanier
Conseiller national

7 M. Patrick Furrer
Vice-directeur d'Euresearch

8 M. Jordi Montserrat
Directeur Romandie de «venturelab»/«venture kick»

Dominique Perruchoud (*absent sur la photo*)
Président du Conseil d'administration de Cimark SA

M. Walter Steinlin (*absent sur la photo*)
Swisscom relations universitaires
Président de la Commission pour la technologie et l'innovation (CTI)

Hors Conseil de fondation

9 Prof. Hervé Bourlard
Directeur de l'Idiap

10 François Foglia
Directeur adjoint de l'Idiap

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.

D^r Jordan Cohen

Independent Consultant, Spelamode
Half Moon Bay, CA, USA

Prof. Anil K. Jain

Distinguished Professor, Department of Computer Science
& Engineering
Michigan State University, USA

D^r John Makhoul

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

Prof. Kenji Mase

Professor, Graduate School of Information Science
Nagoya University, Japan

Prof. D^r Nelson Morgan

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

Prof. Klaus-Robert Müller

Professor for Computer Science, TU Berlin
Director, Bernstein Focus on Neurotechnology,
Berlin, Germany

D^r David Nahamoo

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

Prof. Gerhard Sagerer

Rector, Bielefeld Universität
Germany

Prof. Bernt Schiele

Max-Planck-Director, MPI Informatics
Professor at Saarland University,
Saarbrücken, Germany

Prof. D^r 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



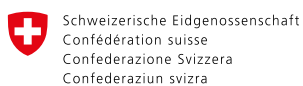
www.theark.ch



www.ideark.ch



www.snf.ch



Agence pour la promotion de l'innovation CTI

www.kti.admin.ch

HASLERSTIFTUNG

www.haslerstiftung.ch

ENCARTÉ SCIENTIFIQUE

IDIAP RESEARCH AREAS: HUMAN AND MEDIA COMPUTING

To face its continuous growth and diversification in different, but complementary, research directions, while building upon its general institutional theme of “Human and Media Computing” (as defining ourselves on our web site www.idiap.ch), Idiap’s expertise mainly spans five research areas, which are summarized in Figure 1 below. In that table, each research area is also described in a few keywords.

Idiap Research Areas (horizontal)	Keywords
Perceptual and cognitive systems	Speech and audio processing, computer vision, handwriting recognition, document processing, computational cognitive sciences, robotics, natural language processing, machine translation
Human and social behavior	Web and mobile social media, social interaction sensing, social signal processing, verbal and nonverbal communication analysis, mobile phone sensing, computational social science
Information and presentation interfaces	Multimedia information systems, user interfaces, contextualization, personalization, system evaluation, mobile HCI using Big Data, data driven services
Biometrics	Speaker recognition, face recognition, multimodal biometric fusion, mobile biometry, spoofing and anti-Spoofing
Machine learning	Statistical and neural network based ML, computational efficiency, online learning, multi-sensor processing, very large datasets

Figure 1: Idiap research areas

- 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.
- Human and social behaviour:** Social media; Mobile media; Social interaction analysis; Social signal processing. This area spans methods for analyzing human and social behaviour 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.
- Information and presentation interfaces:** 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; Counter-measures to spoofing attacks. Biometric person recognition (Biometrics) refers to the process of automatically recognizing a person using distinguishing behavioural patterns (gait, signature, keyboard typing, lip movement, hand-grip) or physiological traits (face, voice, iris, fingerprint, hand geometry, EEG, ECG, ear shape, body odour, body salinity, vascular). Idiap is currently a recognized leader in that field, fully exploiting our multi-disciplinary expertise in image processing, computer vision,



pattern recognition and machine learning. After the end of the MOBIO EU project, this leadership was confirmed in 2011 by the granting of new large EU projects, like the ICT project TABULA RASA and the SECURITY project BEAT. Based on this expertise, Idiap is currently initiating a “Swiss Center for Biometric Security Research and Testing”.

- **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 Deep-STD project), Natural Language Processing with a particular interest in semantic analysis (Sem-Tex 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 speech and audio community; these include the AMI corpus¹ and the Juicer ASR decoder². More recently, the MediaParl database has been released³; other contributions such as ISS and SSP are on the GitHub site⁴. 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. The group is also very active in contributing to KALDI – open-source ASR toolkit⁵.

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 EC Eureka project D-Box⁶), 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 ef-

fectively exploit acoustic and linguistic resources of a different language or multiple languages (funded through SNSF projects like ICS-2010⁷ and FlexASR⁸).

- **Cross-lingual and multi-lingual speech recognition**
In addition to the above, and in the context of an Industry sponsored project (Samsung, Inc. South Korea⁹), we are also investigating new fast acoustic model adaptation techniques (currently referred to as “subspace Gaussian modelling”) 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. In particular, recent advances in speech technology allow training on both languages simultaneously, improving the recognition rate for both.
- **Grapheme-based ASR**
State-of-the-art ASR systems typically use phones as the standard subword units. However, phones are not always well defined, not all languages have well developed phonetic lexicon, 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 AddG2SU¹⁰).
- **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 projects like SNSF funded A-MUSE¹¹ and CTI funded

1 www.idiap.ch/mmm/corpora/ami

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

3 www.idiap.ch/dataset/mediaparl

4 <https://github.com/idiap>

5 <http://kaldi.sourceforge.net>

6 www.idiap.ch/scientific-research/projects/d-box-a-generic-dialog-box-for-multilingual-conversational-applications

7 www.idiap.ch/scientific-research/projects/interactive-cognitive-systems

8 www.idiap.ch/scientific-research/projects/flexible-grapheme-based-automatic-speech-recognition

9 www.idiap.ch/scientific-research/projects/domain-adaptation-using-subspace-models

10 www.idiap.ch/scientific-research/projects/flexible-acoustic-data-driven-grapheme-to-subword-unit-conversion

11 www.idiap.ch/scientific-research/projects/adaptive-multilingual-ASR-TTS



MultiVEO¹², 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 fledged 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-resourced 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, page XII.

Speech Synthesis

- Text-to-speech synthesis (TTS)

Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS is central to two ongoing projects: SIWIS¹³ (Spoken Interaction With Interpretation in Switzerland) and RECOD¹⁴, on low bit rate speech coding (see below).

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 FP7 EMIME¹⁵ 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.

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.

TTS is also pushed towards light-versions to be integrated into open-gaming platform through the EC Eureka project D-BOX¹⁶

12 www.idiap.ch/scientific-research/projects/high-accuracy-speaker-independent-multilingual-automatic-speech-recognition-system

13 www.idiap.ch/project/siwis/

14 www.idiap.ch/scientific-research/projects/low-bit-rate-speech-coding

15 www.emime.org/

16 www.idiap.ch/scientific-research/projects/d-box-a-generic-dialog-box-for-multilingual-conversational-applications

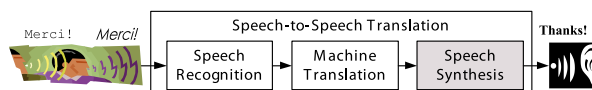


Figure 2: Typical speech-to-speech translation system architecture developed at Idiap.

- 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. Under the RECOD project funded by armasuisse¹⁷, 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.

17 www.ar.admin.ch/internet/armasuisse/en/home.html



- **Microphone arrays**

Idiap further performs studies on audio recorded from microphone arrays, especially those organized in “ad-hoc” geometries. Recently, we mainly focused on the automatic calibration of arbitrary microphone array geometries, fusing Euclidean constraints into generic mathematical concepts. This in turn enables the 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.

- **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 project inEvent¹⁸ and the SNSF project RODI¹⁹.

- **Audio indexing**

In the context of the Hasler Foundation project SESAME²⁰, 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**

Over the last two years, 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 project²¹. 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 PANDA²², 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 applied mostly to the processing of images and video streams.

Our research follows two main axes. The first is the study of novel generic learning methods to cope with feature spaces of large dimension, 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

- **Joint feature selection**

When dealing with very large feature space, it is often desirable to apply a pre-processing step during learning which selects a “good” sub-set of features. Many criteria have been developed for that task, but while it is easy to measure how good a single feature is, they often fail at taking into account the joint quality of a set of features.

We have developed new methods which relies on a Gaussian model of the feature responses, under which we explicitly measure the information content of the subset of features we are selecting. Our main contribution is a “Gaussian compromise”, to approximate the entropy of a mixture of Gaussian densities, for which no analytical or tractable expression exist.

- **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 to identify good “sources” of samples such as web sites, data-bases, or 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.

¹⁸ www.inevent-project.eu/

¹⁹ www.idiap.ch/scientific-research/projects/role-based-speaker-diarization

²⁰ www.idiap.ch/scientific-research/projects/searching-swiss-audio-memories

²¹ www.idiap.ch/scientific-research/projects/low-bit-rate-speech-coding

²² www.idiap.ch/scientific-research/projects/perceptual-background-noise-analysis-for-the-newest-generation-of-telecommunication-systems-1



- **Learning in large dimension feature space**

In the context of the MASH project²³ we have designed new Boosting procedures to learn with hundreds of families of thousands or millions of features, by either 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, or by optimizing on-the-fly the number of samples and features to look at every boosting step.

Our current work addresses the design of artificial neural networks in that context. We investigate novel algorithmic organizations of the computation coupled with sparse learning to reduce the cost for their evaluation.

- **High-dimension density similarity measures**

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

Scene understanding

- **Fast object detection**

In the SNSF project VELASH, we have investigated 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.

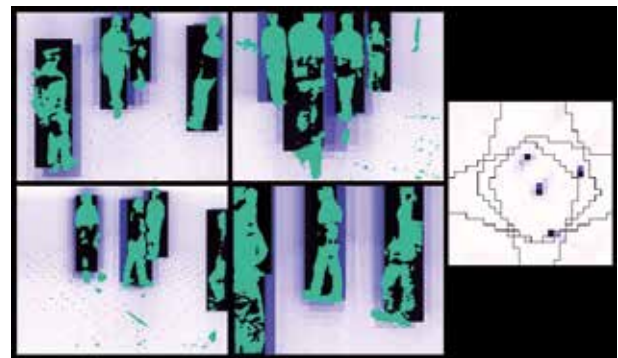


Figure 3: Our fast version of the Deformable Part Model object detector (left) is close to one order of magnitude faster than other existing implementations. The Probabilistic Occupancy Map (right) estimates probabilities of presence on the ground, given the result of a background subtraction procedure in individual views.

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.

²³ <http://mash-project.eu>



- **Multi-camera tracking**

We have a long-lasting collaboration with the CVLab at EPFL around the design of a robust multicamera tracking system, which now takes place in the context of the SNSF WILDTRACK project, in collaboration with ETHZ.

Our system is composed of the Probabilistic Occupancy Map, and a flow-based convex multitarget 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.

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 2013 was composed of 1 head of group, 3 postdoctoral researchers, 1 SNSF Ambizione Research Fellow, 4 PhD students, and 2 visiting researchers (1 postdoc and 1 PhD student). Research lines that were investigated in 2013 included analysis of ubiquitous face-to-face interaction, behavioral analysis of social video, and urban data mining from smartphone and location-based social network data.

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, we have investigated connections between hireability, personality, and performance and nonverbal cues automatically measured from audio and video streams.



Figure 4: 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 transfer learning approaches. This research was conducted in the framework of two SNSF projects (SONVB²⁴: Sensing and Analyzing Organizational Behavior, Sinergia Interdisciplinary Program; and SOBE²⁵, an Ambizione Research Fellowship) and the NISHA²⁶ project (NTT-Idiap Social Behavior Analysis Initiative) funded by NTT Communication Science Laboratories, Japan. Our research in ubiquitous face-to-face interaction was featured in Swiss National Television²⁷.

Social video

The work in this line developed 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 5). We have examined 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. We are using video crowdsourcing with platforms like Amazon Mechanical Turk to collect interpersonal judgments; this is a powerful technique that scales 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 Be-

²⁴ www.idiap.ch/project/sonvb

²⁵ www.idiap.ch/scientific-research/projects/multimodal-computational-modeling-of-nonverbal-social-behavior-in-face-to-face-interaction-1

²⁶ www.idiap.ch/project/nisha/

²⁷ <http://la1.rsi.ch/home/networks/la1/cultura/ll-giardino-di-Albert/2013/05/28/gesto-rivelatore.html#Video>



havior in Social Media). As a follow-up, a PhD student graduate in the group received an EPFL Innogrant Award to support the creation of a social video start-up during 2014.

Urban data mining

Our work in this domain spans two research lines: large-scale smartphone sensing and mobile social network data analysis (Figure 5). For the first line, after co-leading the successful Nokia Mobile Data Challenge 2012 initiative (which attracted participation from several hundred researchers around the world to develop new ideas in mobile computing using the Lausanne smartphone dataset – 180 volunteers in French-speaking Switzerland contributing everyday-life sensor and log data over one year), we undertook the roles of MDC Chair and exclusive worldwide distributor of the MDC data. During 2013, several dozen institutions requested access to this unique research resource²⁸.

Regarding mobile social networks, supported by the SNSF HAI²⁹ project (Human Activity and Interactivity), we are investigating the use of Twitter and FourSquare data to characterize phenomena at city scale, including characterization of urban regions according to language usage. This work received press coverage at the national level³⁰. We also initiated and are currently co-leading the MIRROR multidisciplinary initiative at EPFL (Interrogating Big Data to Improve Decision Making in African Cities), which involves several EPFL laboratories with the goal of developing approaches to characterize and understand large-scale phenomena in African cities.

4. Perception and Activity Understanding

The Perception and Activity Understanding group conducts research in human activities analysis from multi-modal data. This includes the development of principled and efficient algorithms for the characterization of people states from multiple data sources (mainly vision and audio) and the interpretation of sequential data in different scenarios. During 2013, the group was composed of five post-doctoral members, 4 PhD students, and one master student. The year was marked by ending projects (EU VANAHEIM and HUMAVIPS projects, SNSF Promovar) with increased efforts towards the building of real-time demonstrators, and the finalization of journal publications (6 journals published in 2013). Some of the main research carried out this year are described below.

²⁸ www.idiap.ch/project/mdc/

²⁹ www.idiap.ch/scientific-research/projects/human-activity-and-interactivity-modeling

³⁰ www.computerworld.ch/news/kommunikation/artikel/social-media-schweiz-schreibt-englisch-63967/

Tracking in surveillance scenarios

In the context of the VANAHEIM project, we have continued our investigation towards enhanced tracking and behavior recognition for surveillance applications. In particular, several improvements were made within our Conditional-Random-Field (CRF) tracking-by-detection framework. This include the exploitation of image-based motion cues to improve the association of human detections and keep the optimization at the detection level rather than relying on the error prone generation of tracklets through frame-to-frame detection association to build trajectories; the derivation of multi-cue association reliability factors; and further refinements of the unsupervised learning of association parameter distribution parameters. State-of-the-art results were obtained on several benchmark data of the field. The work received the best paper award at the annual IEEE Performance Evaluation of Tracking and Surveillance (PETS) workshop.

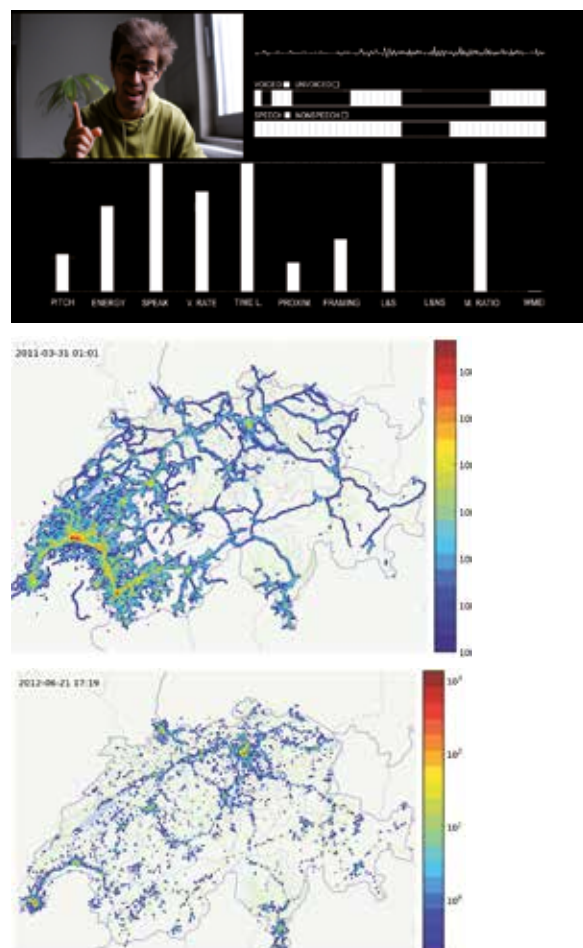


Figure 5: Top: The VlogSense concept: Online vloggers are characterized by a variety of nonverbal behavioral features. Center: Heatmap of the Lausanne smartphone location data in Switzerland. Bottom: Heatmap of check-ins in Switzerland using six months of Foursquare data.



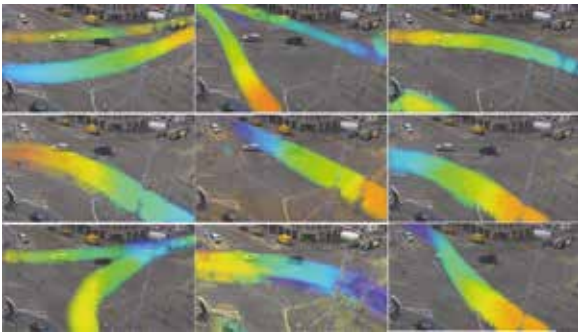


Figure 6: Top: the Idiap human sensing system working on the Nao robot. Bottom : automatic activity discovery from video logs. Each image displays an extracted activity motif that correspond to a sequence of observations that regularly co-occur over time. The color gradient indicates the timing since the start of the motif (from blue to red).



Figure 7: Top: Gaze tracking rendering during a job interview interaction. Bottom: Setup used for the EYEDIAP gaze dataset, the first public database for eyetracker evaluation.

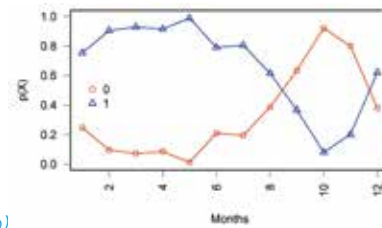
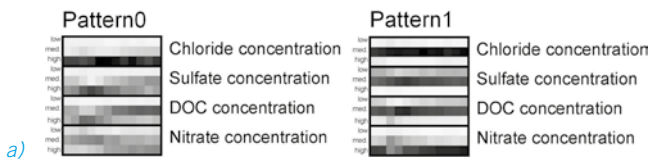


Figure 8: Water quality analysis during and after rainfall events using our topic models. The water quality is characterized by a set of solute-concentration parameters quantified in three different classes: low, medium, high. Water quality time series are decomposed as a mixture of K topic (temporal patterns) represented by the probabilities of observing solute concentration levels a given day (from 2 days before to 1 after the rainfall event). (a) Temporal patterns discovered when $K=2$ (black means high probability). (b) The monthly pattern distribution (averaged over 12 years, month 1 = september). For instance, it resorts that the 0th pattern is typical of the summer season

Non-verbal behaviour analysis in Human-Robot interaction scenarios

Our long-standing work on human interaction modelling has been continued in the context of the European HUMAVIPS project that seeks to endow humanoid robots with basic social skills necessary to deal with small groups of people, and where Idiap developed different sensing components (see left image of Fig. 6). Building upon our multiperson realtime head tracking and pose estimation algorithm, we have improved the recognition of the Visual focus of attention (VFOA) of people by using different contextual models exploiting the robot dialog state (what he is speaking about, who he is addressing) as a strategy for disambiguating amongst looking at different VFOA alternatives in a given situation. Furthermore, we have investigated how context could increase the recognition of addressee (to whom a person is speaking, in particular in HRI, is a person speaking to the robot) whether directly or indirectly through improving VFOA. In particular, we showed that considering only potential addressee targets within the VFOA recognition module rather than considering all VFOA targets (eg. including paintings) was producing much better results.

Multi-modal 3D Gaze tracking using depth sensor data

In the TRACOME project funded by SNSF, we continued our investigation towards the extraction of non-verbal behavior cues and more specifically gaze from cheap RGB-D (depth) cameras like Kinect. Building on our head-pose free gaze estimation framework, we have proposed extensions that accommodate unknown users and can couple the gaze estimation from the two eyes. In a second thread, we have exploited the system and demonstrated its effectiveness at estimating precise mutual gaze events between two persons involved in natural interactions (job interviews) (see Fig. 7). Finally, given the lack of common benchmarks for the evaluation of eye gaze tracking systems, we have recorded a significant large dataset that covers most difficulties of gaze estimation (different persons, sessions and sensing conditions towards 2D -screen- or 3D targets). The dataset is made public and will be presented at the Eye Tracking and Research for Applications conference.



Activity analysis

Motif discovery and stream selection. We have continued our work on 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. The method can be applied to typical situations where one or multiple sensor record the activities of multiple objects/people, like multimodal sensors (proximity, water, light sensors, etc.) in home automation (domotics) applications (see Fig. 8 right for an example in video).

In particular, in the EU Vanaheim project, these models were used this year to perform abnormality detection. More specifically, in collaboration with ethologists from the 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 rankings of video streams according to the unusualness or interestingness of their content made by real human operators were shown to be in good agreement with those resulting from the automatic abnormality rating methods. A real-time demonstrator of the approach applied to video feeds from the Paris metro was live-demonstrated in a public show in Paris in front of more than 60 professionals of the field, along with other Idiap algorithms (tracking, left-luggage detection).

Water chemistry analysis in rainfall events. In collaboration with INRA (Institut National de Recherche en Agronomie), France, we investigated the use of our temporal probabilistic topic framework to analyze hydrology time series, and more precisely the impact of flood-induced variability on water chemistry to verify INRA's assumption that stream chemistry reacts differently to rainfall events according to the season and the weather of the hydrological year. The mixture aspect of our models and robustness to missing data proved to be interesting to explain the observed measures that depend on a mixture of environmental factors. The discovered chemistry motifs and their occurrences provided a good insight about what was going on (chemical wise) at a given water place, and matched the expert view on the data (see Fig. 8).

Multimodal face and person diarization and naming

Identifying people appearing and speaking in multimedia data allows to monitor who said what and when and is thus important for the design of search and fast browsing tools for personal data collection or broadcast programs. This is the challenge addressed in the french REPERE evaluation campaign, in which Idiap is participating as part of the SODA project funded by the french research agency ANR. This year, we have worked on a CRF framework for the audio-visual association and diarization of faces and speakers in broadcast data that accounts to rely on context and cluster analysis to

identify the role of people in the program (guest, anonymous person in the public, etc). Furthermore, leveraging on the text overlaid on videos that often announce speaking or appearing people, we have shown that it is possible to name and tag the most relevant persons of a program.

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 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 NINAPRO³¹ 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 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.

³¹ www.idiap.ch/project/ninapro/



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 America³², 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 investigate new efficient end-to-end systems based on recurrent convolutional neural networks (see Figure 9), alleviating the need of any hand-crafted feature and allowing the modeling of complex spatial dependencies with a low inference cost.



Figure 9: 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.

- **Spoken Term Detection**

The DeepSTD project, funded by the Swiss Hasler Stiftung³³, is concerned about applying new deep learning techniques to audio processing, with a particular interest in fast spoken term detection. The year 2012 and 2013 were devoted to estimating phoneme class conditional probabilities directly from raw speech signal, using convolutional neural networks, as well as investigation ways to automate speech segmentation in these models.

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 SODS project, funded by the Swiss Hasler Stiftung³⁴, we wish to develop a new type of search engine distributed over available web servers (in the same spirit than YaCy³⁵), 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.

³² www.nec-labs.com/research-departments/machine-learning/machine-learning-home

³³ www.haslerstiftung.ch

³⁴ www.haslerstiftung.ch. The SODS project is part of the SmartWorld special call.

³⁵ <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 KeyLemon³⁶ 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 2013, we delivered a state-of-the-art version of our multi-pose face and gender detection system (see Figure 10), which runs in real time with good performance even in adverse conditions.



Figure 10: Our real time face detector detects multi-pose faces and labels gender (pink for female, bluish for male) in adverse conditions (left). It also outputs 22 precise facial features, in multi-pose conditions (right).

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

SENNA³⁷ is a standalone software distributed under a non-commercial license, that we develop through our ongoing collaboration with NEC Laboratories America³⁸. 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 Torch³⁹ 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.

³⁶ <https://www.keylemon.com>

³⁷ <http://ml.nec-labs.com/senna>

³⁸ www.nec-labs.com/research-departments/machine-learning/machine-learning-home

³⁹ 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 localisation 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 11).

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 localisation 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.

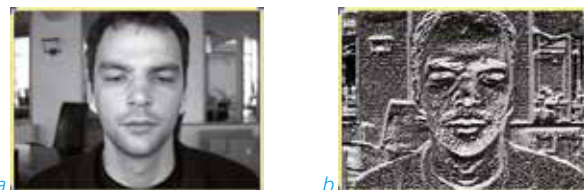


Figure 11: Illustration of the LBP feature map (b) for the original image (a).

Face and Speaker Recognition

We leveraged prior work on distribution modelling for part-based face recognition (Figure 12) using session variability modelling techniques. Session variability modelling aims to explicitly model and suppress detrimental within-class (inter-session) variation. We examined two techniques to do this, inter-session variability modelling (ISV) and joint factor analysis (JFA), which were initially developed for speaker recognition.

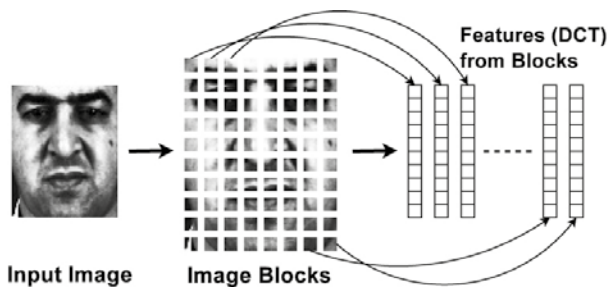


Figure 12: Illustration of the LBP feature map (b) for the original image (a).

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 recognise 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 modelling, motion detection or spatio-temporal pattern recognition. We also organised 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 project⁴⁰, 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 parameterisations 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 normalisation algorithm in a preconfigured face recognition tool-chain. 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 tool-chain and measure its impact on the overall system performance;
3. A scientific challenge is organised in which users contribute with features that will be used in a pre-configured key-point localisation algorithm, for face recognition. The user wishes to upload his/her own feature-extraction algorithm to the contest and observe how that performs.

⁴⁰ www.beat-eu.org/



8. Natural Language Processing

The Idiap Natural Language Processing group carries out research in semantic-level language analysis with two principal applications: machine translation and information retrieval, more specifically multimedia recommendation. The NLP Group reached in 2013 its maximal extension to date, with eight people. Its research is aimed at improving text analysis performance by combining robust semantic and pragmatic analysis methods with state-of-the-art, data-driven techniques for the two applications: respectively, phrase-based statistical machine translation, and content-based or collaborative filtering algorithms for multimedia recommendation over networked data.

Machine Translation

Work on this topic focuses on a problem that is not so often targeted in the current statistical machine translation (SMT) paradigm: the translation of words that depend on long-range relationships between sentences. The work is mainly supported by the SNSF through its Sinergia program, first in the COMTIS project⁴¹ and then in the MODERN project⁴², both coordinated by Idiap. We are collaborating with two teams from the University of Geneva, in linguistics and computational learning, a team from the University of Zurich in machine translation, and one from the University of Utrecht in linguistic modeling and experimentation.

We have analyzed large collections of various types of dependencies between sentences which are problematic for current MT engines, specifically rhetorical relations signaled by discourse connectives, and verb tenses which must be coherently translated across a text. We have proposed computational methods for the automatic disambiguation of discourse connectives, using classifiers such as maximum entropy or decision trees based on surface features, but also using more semantic features such as synonymy, word polarity, or temporal markers. The accuracy of our classifiers for discourse connectives (e.g. distinguishing between temporal and causal uses of the English word *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.

Similarly, we dealt with verb tenses, which also depend on relations across sentences, and raise numerous challenges

when translating from English to French. We collected data and performed classification experiments on the translation of English past tenses into French, in collaboration with the University of Geneva, and through an internship at Idiap. A classifier using surface features was designed to predict a linguistic feature ('narrativity') in unseen texts, and the predictions were combined with SMT. Moreover, using a precise alignment of EN/FR verb phrases, we trained a classifier predicting directly the translated tense (e.g. for EN simple past into either FR *passé composé*, *passé simple*, or *imparfait*). We have demonstrated that the translation of verb phrases is improved by this method.

Similarity and Recommendation over 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 past AMI⁴³ and AMIDA⁴⁴ EU projects, and the IM2 Swiss NCCR⁴⁵ finished at the end of 2013. We also are also supported by the InEvent⁴⁶ EU project and the REMUS Hasler project, and receive SNSF funding for technology transfer in collaboration with Klewel SA, the Idiap spin-off dedicated to conference recording and broadcasting.

In the past year, a PhD thesis was completed on learning 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 (VP) from a graph node to another node, and a distance measure was defined based on it. Moreover, several embeddings of items, offering compact representations, quicker computation and better generalization capacities were learned from VP. These similarity learning models over graphs were applied to NLP tasks such as word similarity or document clustering, and to information retrieval ones, including learning-to-rank.

Our first application framework is the Automatic Content Linking Device (ACLD), which is a multimedia retrieval system with spoken queries (explicit or implicit), running in real-time to enrich a conversation by making suggestions of potentially relevant documents. We have proposed a method, using a submodular optimization criterion, to extract from a conversation fragment, with high precision, a diverse set of keywords that reflect the diversity of topics hypothesized in the fragment. We have shown using human subjects recruited through a crowdsourcing plat-

⁴¹ 2010-2013, see www.idiap.ch/project/comtis

⁴² 2013-2016, see www.idiap.ch/project/modern

⁴³ www.amiproject.org

⁴⁴ www.amidaproject.org

⁴⁵ www.im2.ch

⁴⁶ www.inevent-project.eu



form that our keywords were judged as more relevant than competing methods, and that they allowed the retrieval of more relevant recommendations for the conversation fragment, especially at the top levels. A method for diverse retrieval was also designed, with application to implicit but also explicit queries. We have studied 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 – we have collected its metadata from TED and made it publicly available⁴⁷. We compared a number of approaches on this data set: 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. We have considered aspect-based sentiment analysis and proposed an innovative method based on multiple-instance learning to determine the contribution of each sentence in a review to aspect-related judgments, demonstrating improvement over the state-of-the-art on seven data sets (including TED).

The NLP Group participated in the MediaEval 2013 campaign, specifically in the Search and Hyperlinking task. The search sub-task required finding a determined segment of a show (from 1260 hours of BBC TV material) based on a test query looking for a known item, while the hyperlinking sub-task required finding items related to the known items. We proposed a unified approach to both sub-tasks, using content-based recommendation techniques to compute the most similar segments to a given text query or to another segment, based on words but also on visual concepts. For the search task, we ranked third out of seven institutions, while for the hyperlinking task (the closest to our expertise from inEvent) our team was ranked first, reaching a mean average precision above 50%.



Figure 13: The MUST-VIS navigation graph (bottom) and video/slide player (top). Each lecture is represented either with a key-frame or with keyword clouds (the two modes are switched by clicking). The lecture in focus, center, is surrounded by lectures with related segments.

In collaboration with our partners in the inEvent project, and in particular with Klewel, the NLP Group was the winner of one of the five ACM Multimedia 2013 Grand Challenge Solutions, specifically the one for lecture segmentation and retrieval. Our system, represented in Figure 13, was called MUST-VIS for Multi-factor Segmentation for Topic Visualization and Recommendation, and allowed users to visualize a lecture as a series of segments, with relations to other similar lectures and segments, computed using recommendation techniques.

⁴⁷ www.idiap.ch/dataset/ted



MAIN PROJECTS IN PROGRESS

Projects funded by European agencies

- Name** **BBFOR2** (Bayesian Biometrics for Forensics)
Funding FP7 - Marie Curie Actions (ITN)
Coordinator Radboud University Nijmegen - Stichting Katholieke Universiteit
Duration 01.01.2010-31.12.2013
Partner(s) Katholieke Universiteit Leuven, Idiap Research Institute, Högskolan i Halmstad, University of York, Universiteit Twente, Universidad Autónoma de Madrid, Politecnico di Torino, Netherlands Forensic Institute
- Name** **BEAT** (Biometrics Evaluation and Testing)
Funding FP7 - STREP
Coordinator Idiap Research Institute
Duration 01.03.2012-28.02.2016
Partner(s) University of Surrey, Universidad Autónoma de Madrid, Tüvit, Tubutak Informatics and Information Security, Advanced Technologies Research Center, Sagem Sécurité S.A., Katholieke Universiteit Leuven, École Polytechnique Fédérale de Lausanne, Commissariat à l'Énergie Atomique
- Name** **EUMSSI** (Event Understanding through Multimodal Social Stream Interpretation)
Funding FP7 - IP
Coordinator Fundacio Barcelona Media
Duration 01.12.2013-30.11.2016
Partner(s) Video Stream Networks S.L., Université du Maine, Gottfried Wilhelm Leibniz Universität Hannover, Idiap Research Institute, Gesellschaft zur Förderung der angewandten Informationsforschung EV, Deutsche Welle
- Name** **DBOX** (A generic dialog box for multilingual conversational applications)
Funding EUREKA's Eurostars Programme
Coordinator Mi'pu'mi Games GmbH
Duration 01.11.2012-31.10.2015
Partner(s) Universität des Saarlandes, SIKOM Software GmbH, Koemei SA, Idiap Research Institute, Acapela
- Name** **HUMAVIPS** (Humanoids with auditory and visual abilities in populated spaces)
Funding FP7 - STREP
Coordinator Institut de Recherche en Informatique et en Automatique
Duration 01.02.2010-31.01.2013
Partner(s) Jean-Marc Odobez, Hervé Boulard, François Fleuret, François Fleuret
- Name** **INEVENT** (Accessing Dynamic Networked Multimedia Events)
Funding FP7 - STREP
Coordinator Idiap Research Institute
Duration 01.11.2011-31.10.2014
Partner(s) University of Edinburgh, Radvision LTD, Klewel, IBM Israel LTD, Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
- Name** **MASH** (Massive Sets of Heuristics for Machine Learning)
Funding FP7 - STREP
Coordinator Idiap Research Institute
Duration 01.01.2010-30.06.2013
Partner(s) Centre national de la recherche scientifique, Universitaet Potsdam, Institut de Recherche en Informatique et en Automatique, Czech Technical University in Prague
- Name** **PASCAL2** (Pattern Analysis, Statistical Modelling and Computational Learning 2)
Funding FP7 - NoE
Coordinator University of Southampton
Duration 01.03.2008-28.02.2013
Partner(s) University College London, Helsinki University of Technology, Max Planck Society for the Advancement of Sciences, Jozef Stefan Institute, Idiap Research Institute, Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Stichting Centrum voor Wiskunde en Informatica, Centre national de la recherche scientifique, Xerox SAS, Université Pierre et Marie Curie-Paris 6, University of Bristol, Università degli Studi di Milano, University of Manchester, Helsingin yliopisto, University of Edinburgh



9. Name **ROCKIT** (Roadmap for Conversational Interaction Technologies)
 Funding FP7- CSA
 Coordinator University of Edinburgh
 Duration 01.12.2013-30.11.2015
 Partner(s) Vodera Limited, Universität des Saarlandes, LT-Innovate, Idiap Research Institute, Europe Limited SA
10. Name **SSPNET** (Social Signal Processing Network)
 Funding FP7 - NoE
 Coordinator Idiap Research Institute
 Duration 01.02.2009-31.01.2014
 Partner(s) Universiteit Twente, Università di Roma Tre, University of Geneva, University of Edinburgh, Delft University of Technology, Queen's University Belfast, Institut de Recherche en Informatique et en Automatique, Imperial College London, German Research Centre for Artificial Intelligence
11. Name **TABULARASA** (Trusted Biometrics under Spoofing Attacks)
 Funding FP7 - STREP
 Coordinator Idiap Research Institute
 Partner(s) University of Southampton, University of Oulu, Università Degli Studi Di Cagliari, Universidad Autónoma de Madrid, Startlab Barcelona SL, Sagem Sécurité S.A., KeyLemon SA, Eurocom, Biometry.com AG, Centre for Science, Society and Citizenship, Institute of Automation Chinese Academy of Sciences
12. Name **VANAHEIM** (Video/Audio Networked surveillance system enhancement through Human-centered adaptive Monitoring)
 Funding FP7 - IP
 Coordinator Multitel ASBL
 Duration 01.02.2010-30.09.2013
 Partner(s) Thales Italia spa, Thales Communications SA, Régie Autonome des Transports Parisiens, Ludwig Boltzmann Institute for Urban Ethology, Institut de Recherche en Informatique et en Automatique, Gruppo Torinese Trasporti
13. Name **SODA** (Person Recognition in debate and broadcast news)
 Funding ANR - Agence Nationale de la Recherche
 Coordinator Université du Maine
 Duration 01.06.2011-31.05.2014
 Partner(s) Idiap Research Institute



Projects funded by Swiss agencies

- Name **A-MUSE** (Adaptive Multilingual Speech Processing)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.10.2012-30.09.2015
Partner(s)
- Name **AROLIS** (Automatic Recommendation of Lectures and Snippets)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.06.2012-31.05.2014
Partner(s) Klewel
- Name **COMTIS** (Improving the coherence of machine translation output by modeling intersentential relations)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.03.2010-31.07.2013
Partner(s) University of Geneva
- Name **DASH** (Object Detection with Active Sample Harvesting)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.09.2012-31.08.2015
Partner(s)
- Name **DIMHA** (Diarizing Massive Amounts of Heterogeneous Audio)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.06.2012-31.05.2014
Partner(s) Koemei
- Name **FLEXASR** (Flexible Grapheme-Based Automatic Speech Recognition)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.03.2010-30.04.2013
Partner(s)
- Name **FLEXASR** (Flexible Grapheme-Based Automatic Speech Recognition)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.05.2013-30.04.2014
Partner(s)
- Name **HAI-2010** (Human activity and interactivity modeling)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.10.2010-30.09.2013
Partner(s)
- Name **ICS-2** (Interactive Cognitive Systems, Indoor Scene Recognition for Intelligent Systems)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.04.2013-31.12.2013
Partner(s)
- Name **IM2-3** (Interactive Multimodal Information Management Phase 3)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.01.2010-31.12.2013
Partner(s)
- Name **LOBI** (Low Complexity Binary Features for Robust-to-Noise Speaker Recognition)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.06.2012-31.05.2014
Partner(s) KeyLemon SA



12. Name **MAAYA** (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.06.2013-31.05.2016
 Partner(s) University of Geneva, University of Bonn
13. Name **MCM-FF** (Multimodal Computational Modeling of Nonverbal Social Behavior in Face to Face Interaction)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.11.2011-31.10.2014
 Partner(s)
14. Name **MODERN** (Modeling discourse entities and relations for coherent machine translation)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.08.2013-31.07.2016
 Partner(s) Universiteit Utrecht, Universität Zürich, University of Geneva
15. Name **NINAPRO** (Non-Invasive Adaptive Hand Prosthetics)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.01.2011-30.06.2014
 Partner(s) HES-SO Valais, German Aerospace Research Center
16. Name **PROMOVAR** (Probabilistic Motifs for Video Action Recognition)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.02.2012-31.07.2013
 Partner(s)
17. Name **RODI** (Role based speaker diarization)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.11.2011-31.10.2014
 Partner(s)
18. Name **SIVI** (Situating Vision to Perceive Object Shape and Affordances)
 Funding SNSF - Swiss National Science Foundation
 Coordinator University of Bonn
 Duration 25.10.2011-30.09.2014
 Partner(s) Technische Universität Wien, RWTH Aachen, Idiap Research Institute
19. Name **SIWIS** (Spoken Interaction with Interpretation in Switzerland)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.12.2012-30.11.2015
 Partner(s) University of Geneva, University of Edinburgh, Eidgenössische Technische Hochschule Zürich
20. Name **SONVB** (Sensing and Analysing Organizational Nonverbal Behavior)
 Funding SNSF - Swiss National Science Foundation
 Coordinator Idiap Research Institute
 Duration 01.06.2010-30.06.2014
 Partner(s) Université de Neuchâtel, The Trustees of Dartmouth College Corp.
21. Name **TRACOME** (Robust face tracking, feature extraction and multimodal fusion for audio-visual speech recognition)
 Funding SNSF - Swiss National Science Foundation
 Coordinator École Polytechnique Fédérale de Lausanne
 Duration 01.01.2011-31.03.2014
 Partner(s) Idiap Research Institute
22. Name **UBM** (Understanding Brain Morphogenesis)
 Funding SNSF - Swiss National Science Foundation



- Coordinator University of Basel
Duration 01.01.2011-31.12.2013
Partner(s) Idiap Research Institute, École Polytechnique Fédérale de Lausanne
23. Name **VELASH** (Very Large Sets of Heuristics for Scene Interpretation)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.09.2012-31.08.2013
Partner(s)
24. Name **ADDG2SU** (Flexible Acoustic Data-Driven Grapheme to Subword Unit Conversion)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.03.2013-01.03.2016
Partner(s)
25. Name **DEEPSTD** (Universal Spoken Term Detection with Deep Learning) Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.12.2011-30.11.2014
Partner(s)
26. Name **IMAGECLEF** (The Robot Vision Task @ ImageCLEF: Towards Web-Robotics)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.01.2013-28.02.2014
Partner(s)
27. Name **MASH-2** (Massive Sets of Heuristics for Machine Learning II)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.07.2013-30.06.2016
Partner(s)
28. Name **SESAME** (SEarching Swiss Audio MEMories)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.01.2011-28.02.2015
Partner(s)
29. Name **SODS** (Semantically Self-Organized Distributed Web Search)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.01.2012-31.12.2014
Partner(s)
30. Name **UBSL** (User-Based Similarity Learning for Interactive Image Retrieval)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.08.2012-31.03.2013
Partner(s)
31. Name **MEDIAPARL** (MediaParl) Funding Etat du Valais
Coordinator Idiap Research Institute
Duration 01.01.2012-31.12.2013
Partner(s)
32. Name **VALAIS-2015** (Valais*Wallis Digital)
Funding Etat du Valais
Coordinator Idiap Research Institute
Duration 01.04.2013-31.12.2015
Partner(s) Société Coopérative Migros Valais, Médiathèque Valais et Archives de l'Etat, Ecole professionnelle des arts contemporains



Projects funded by industrial partners

- Name** **ARIS** (Automatic French and German Speech Recognition System for Efficient Indexing and Search)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 01.04.2013-31.01.2014
Partner(s) Koemei, Klewel
- Name** **COMMIC2** (Système cognitif pour une commande intelligente de chauffage, phase2)
Funding Fondation The Ark
Coordinator Centre de recherches energetiques et municipales
Duration 01.09.2012-30.06.2013
Partner(s) Idiap Research Institute, HES-SO Valais, ARLO Communications
- Name** **LAMI** (Large scale multi-energy information platform)
Funding Fondation The Ark
Coordinator Centre de recherches energetiques et municipales
Duration 01.08.2013-28.02.2014
Partner(s) ESR - Energie Sion Region, Sinergy, Idiap Research Institute, HES-SO Valais
- Name** **SMARTCAP** (Smart Caps Orientation)
Funding Fondation The Ark
Coordinator AISA Automation Industrielle SA
Duration 01.05.2013-21.02.2014
Partner(s) Idiap Research Institute
- Name** **ASLEEP** (Adapting the Static Luggage Detection module for the Protect Rail demonstrator)
Funding IND - Industrial Project
Coordinator Idiap Research Institute
Duration 15.11.2012-31.05.2013
Partner(s)
- Name** **DAUM2012** (Domain Adaptation Using Sub-Space Models)
Funding IND - Industrial Project
Coordinator Idiap Research Institute
Duration 15.10.2012-14.10.2013
Partner(s)
- Name** **NEC** (NEC collaboration)
Funding IND - Industrial Project
Coordinator Idiap Research Institute
Duration 01.05.2012-30.04.2014
Partner(s)
- Name** **NTT** (NISHA - NTT Idiap Social Behaviour Analysis Initiative)
Funding IND - Industrial Project
Coordinator Idiap Research Institute
Duration 15.11.2012-30.09.2014
Partner(s)
- Name** **CLEAR** (Online Cloud-based Platform for Efficient and Robust Face Recognition Servi)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.06.2012-30.11.2013
Partner(s) KeyLemon
- Name** **DOMOCARE** (DomoCare - A new Home Care Preventive Protocol)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.10.2013-30.09.2014
Partner(s) DomoSafety
- Name** **FEDARS** (Feature Extraction from Deep learning Architectures for face Recognition Systems)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.08.2013-31.01.2014
Partner(s) KeyLemon SA



12. Name **MULTIVEO** (High Accuracy Speaker-Independent Multilingual Automatic Speech Recognition System)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.11.2013-31.05.2015
Partner(s) Veovox
13. Name **PANDA** (Perceptual Background Noise Analysis for the Newest Generation of Telecommunication Systems)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.07.2012-31.12.2013
Partner(s) SwissQual AG
14. Name **RECOD2013** (low bit-rate speech coding)
Funding ArmaSuisse
Coordinator Idiap Research Institute
Duration 01.03.2013-31.12.2013
Partner(s)



Projects awarded in 2013 and starting in 2014

- Name** **GENEEMO** (Geneemo: An Expressive Audio Content Generation Tool)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.04.2014-30.09.2015
Partner(s)
- Name** **REMUS** (REMUS: Re-ranking Multiple Search Results for Just-in-Time Document Recommendation)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.01.2014-31.10.2014
Partner(s)
- Name** **SCOREL2** (Automatic scoring and adaptive pedagogy for oral language learning)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.04.2014-30.11.2015
Partner(s) SpeedLingua
- Name** **SIIP** (Speaker Identification Integrated Project)
Funding FP7 - IP
Coordinator Verint System Ltd
Duration 01.05.2014-30.04.2018
Partner(s) Loquendo SpA, The International Criminal Police Organization (Interpol), Inov Inesc inovacao - Instituto De Novas Tecnologias, Idiap Research Institute, International Biometric Group (UK) Limited, green Fusion Limited (data Fusion International), Cassidan SAS; University of Warwick, Snthema S.R.L, Singularlogic Anonymos Etairia Pliroforiakon Systematon & Efarmogon Pliroforkis, Sail Labs Technology AG, Rijksuniversiteit Gronigen, Police Service of Northern Ireland, Ministério da Justiça, Lisboa, ok2go, Laboratorio Di Scienze Della Cittadinanza
- Name** **SUVA** (Intégration de la transcription vocale dans le dossier patient informatisé CRR)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 01.03.2014-31.03.2014
Partner(s) Clinique romande de réadaptation, Cimark
- Name** **UBIMPRESSED** (UBImpressed: Ubiquitous First Impressions and Ubiquitous Awareness)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.01.2014-31.12.2016
Partner(s) Université de Neuchâtel, Cornell University
- Name** **WILDTRACK** (Tracking in the Wild)
Funding SNSF - Swiss National Science Foundation
Coordinator Idiap Research Institute
Duration 01.01.2014-31.12.2016
Partner(s) Eidgenössische Technische Hochschule Zürich, École Polytechnique Fédérale de Lausanne
- Name** **YOUTH@NIGHT** (A multi-disciplinary multi-method study of young people's outgoing and drinking behaviors)
Funding SNSF - Swiss National Science Foundation
Coordinator Sucht Schweiz - Research Institute
Duration 01.01.2014-31.12.2016
Partner(s) Universität Zürich, Idiap Research Institute



LIST OF PUBLICATIONS – 2013

1. Book

1. Hervé Bourlard and Andrei Popescu-Belis. *Interactive Multimodal Information Management*. EPFL Press, Lausanne, 2013.
2. Bonnie Webber, Andrei Popescu-Belis, Katja Markert, and Jorg Tiedemann. *Proceedings of the ACL Workshop on Discourse in Machine Translation (DiscoMT 2013)*. Association for Computational Linguistics, August 2013.

2. Book Chapters

1. Barbara Caputo. Medical image annotation. In Hervé Bourlard and Andrei Popescu-Belis, editors, *Interactive Multimodal Information Management*. EPFL Press, 2013.
2. Remi Emonet and Jean-Marc Odobez. Unsupervised methods for activity analysis and detection of abnormal events. In J. Y. Dufour, editor, *Intelligent Video Surveillance Systems (ISTE)*. Wiley, February 2013.
3. Francois Fleuret, Horesh Ben Shitrit, and Pascal Fua. Re-identification for improved people tracking. In *Person Re-Identification*. Springer, 2013.
4. Mathew Magimai-Doss. Speech processing. In Hervé Bourlard and Andrei Popescu-Belis, editors, *Interactive Multimodal Information Management*, chapter 15, pages 221–245. EPFL Press, 2013.
5. Fabian Nater, Tatiana Tommasi, Luc Van Gool, and Barbara Caputo. Learning to learn new models of human activities in indoor settings. In Hervé Bourlard and Andrei Popescu-Belis, editors, *Interactive Multimodal Information Management*. EPFL Press, 2013.
6. Andrei Popescu-Belis and Hervé Bourlard. Interactive multimodal information management: Shaping the vision. In Hervé Bourlard and Andrei Popescu-Belis, editors, *Interactive Multimodal Information Management*, pages 1–17. EPFL Press, Lausanne, 2013.

3. Articles in Scientific Journals

1. Marcel Alcoverro, Xavier Suau, Adolfo Lopez-Mendez, Josep R. Morros, Javier Ruiz-Hidalgo, Albert Gil, and Josep R. Casas. Gesture control interface for immersive panoramic displays. *Multimedia Tools and Applications*, 1380-7501:1–27, July 2013.
2. André Anjos, Murali Mohan Chakka, and Sébastien Marcel. Motion-based counter-measures to photo attacks in face recognition. *Institution of Engineering and Technology Journal on Biometrics*, July 2013.
3. Afsaneh Asaei, Hervé Bourlard, Bhiksha Raj, Mohammad J. Taghizadeh, and Volkan Cevher. Binary sparse coding for optimally speech reconstruction. *IEEE Transactions on Audio, Speech and Language Processing*, 2013.
4. Afsaneh Asaei, Hervé Bourlard, Mohammad J. Taghizadeh, and Volkan Cevher. Computational methods for spatio-spectral speech recovery via sparse modeling. *Journal of Signal Processing Systems*, 2013.
5. Afsaneh Asaei, Mohammad Golbabaee, Hervé Bourlard, and Volkan Cevher. Structured sparsity models for reverberant speech separation. *IEEE Transaction on Audio, Speech and Language Processing*, 2013.
6. Alice Aubert, Romain Tavenard, Remi Emonet, A. de Lavenne, Simon Malinowski, Thomas Guyet, René Quiniou, Jean-Marc Odobez, Philippe Merot, and Chantal Gascuel. Clustering flood events from water quality time-series using latent dirichlet allocation model. *Water Resources Research*, 2013. Online published version before inclusion in an issue.
7. Horesh Ben Shitrit, Jerome Berclaz, Francois Fleuret, and Pascal Fua. Multi-commodity network flow for tracking multiple people. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013.
8. Joan-Isaac Biel, Daniel Gatica-Perez, John Dines, and Vagia Tsminiaki. Hi youtube! personality impressions and verbal content in social video. *15th ACM International Conference on Multimodal Interaction, Sydney, Australia, ACM, 2013*, 2013.
9. Antoine Bordes, Léon Bottou, Ronan Collobert, Dan Roth, Jason Weston, and Luke Zettlemoyer. Introduction to the special issue on learning semantics. *Machine Learning*, June 2013.



10. Bruno Cartoni, Sandrine Zufferey, and Thomas Meyer. Annotating the meaning of discourse connectives by looking at their translation: The translation-spotting technique. *Dialogue and Discourse*, 4(2):65–86, April 2013.
11. Bruno Cartoni, Sandrine Zufferey, and Thomas Meyer. Using the europarl corpus for cross-linguistic research. *Belgian Journal of Linguistics*, (27):23–42, December 2013.
12. Trinh-Minh-Tri Do and Daniel Gatica-Perez. The places of our lives: Visiting patterns and automatic labeling from longitudinal smartphone data. *IEEE Transactions on Mobile Computing*, 2013.
13. Trinh-Minh-Tri Do and Daniel Gatica-Perez. Where and what: Using smartphones to predict next locations and applications in daily life. *Pervasive and Mobile Computing*, May 2013.
14. Stefan Duffner and Jean-Marc Odobez. A track creation and deletion framework for long-term online multi-face tracking. *IEEE Transactions on Image Processing*, March 2013.
15. Laurent El Shafey, Chris McCool, Roy Wallace, and Sébastien Marcel. A scalable formulation of probabilistic linear discriminant analysis: Applied to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(7):1788–1794, July 2013.
16. Remi Emonet, Jagannadan Varadarajan, and Jean-Marc Odobez. Temporal analysis of motif mixtures using dirichlet processes. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, May 2013.
17. Philip N. Garner, Milos Cernak, and Petr Motlicek. A simple continuous pitch estimation algorithm. *IEEE Signal Processing Letters*, 20(1):102–105, January 2013.
18. David Imseng, Hervé Bourlard, John Dines, Philip N. Garner, and Mathew Magimai-Doss. Applying multi- and cross-lingual stochastic phone space transformations to non-native speech recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, 2013.
19. David Imseng, Petr Motlicek, Hervé Bourlard, and Philip N. Garner. Using out-of-language data to improve an under-resourced speech recognizer. *Speech Communication*, 2013.
20. Elie Khoury, Laurent El Shafey, Chris McCool, Manuel Günther, and Sébastien Marcel. Bi-modal biometric authentication on mobile phones in challenging conditions. *Image and Vision Computing*, October 2013.
21. S. R. Krishnan, Mathew Magimai-Doss, and C. S. Seelamantula. A savitzky-golay filtering perspective of dynamic feature computation. *IEEE Signal Processing Letters*, 20(3):281 – 284, March 2013.
22. J. K. Laurila, Daniel Gatica-Perez, Jan Blom, Olivier Bornet, Trinh-Minh-Tri Do, O. Dousse, Julien Eberle, and Markus Mietinen. From big smartphone data to worldwide research: The mobile data challenge. *Pervasive and Mobile Computing*, 2013.
23. Riwal Lefort and Francois Fleuret. treekl: A distance between high dimension empirical distributions. *Pattern Recognition Letters*, 34(2):140–145, 2013.
24. Patrick Marmoroli, M. Carmona, Xavier Falourd, Hervé Lissek, and Jean-Marc Odobez. Observation of vehicle axles through pass-by noise: A strategy of microphone array design. *IEEE Trans. on Intelligent Transportation Systems*, March 2013.
25. Chris McCool, Roy Wallace, Mitchell McLaren, Laurent El Shafey, and Sébastien Marcel. Session variability modelling for face authentication. *IET Biometrics*, 2(3):117–129, September 2013.
26. Michael McCoy, Volkan Cevher, Quoc Tran Dinh, Afsaneh Asaei, and Luca Baldassarre. Convexity in source separation: Models, geometry, and algorithms. *IEEE Signal Processing Magazine, Special Issue on Source Separation and Applications*, 2013.
27. Petr Motlicek, Stefan Duffner, Danil Korchagin, Hervé Bourlard, Carl Scheffler, Jean-Marc Odobez, Giovanni Del Galdo, Markus Kallinger, and Oliver Thiergart. Real-time audio-visual analysis for multiperson videoconferencing. *Advances in Multimedia*, 2013:21, August 2013. Hindawi Publishing Corporation, Article ID 175745.
28. Lakshmi Saheer, Junichi Yamagishi, Philip N. Garner, and John Dines. Combining vocal tract length normalization with hierarchical linear transformations. *IEEE Journal of Selected Topics in Signal Processing - Special Issue on Statistical Parametric Speech Synthesis*, 2013.
29. Björn Schuller, Stefan Steidl, Anton Batliner, Elmar Nöth, Alessandro Vinciarelli, Felix Burkhardt, Felix Weninger, Florian Eyben, Tobias Bocklet, Gelareh Mohammadi, and Benjamin Weiss. A survey on perceived speaker traits: Personality, likability, pathology and the first challenge. *Computer Speech and Language*, 2013.
30. Jagannadan Varadarajan, Remi Emonet, and Jean-Marc Odobez. A sequential topic model for mining recurrent activities from long term video logs. *International Journal of Computer Vision*, 103(1):100–126, May 2013.



31. Alessandro Vinciarelli and Gelareh Mohammadi. A survey of personality computing. *IEEE Transaction on Affective Computing*, 2013.
32. Majid Yazdani and Andrei Popescu-Belis. Computing text semantic relatedness using the contents and links of a hyper-text encyclopedia. *Artificial Intelligence Journal*, 194:176202, January 2013.

4. Articles in Conference Proceedings

1. Aniruddha Adiga, Mathew Magimai-Doss, and Chandra Sekhar Seelamantula. Gammatone wavelet cepstral coefficients for robust speech recognition. In *Proceedings of IEEE TENCON*, October 2013.
2. Oya Aran and Daniel Gatica-Perez. Cross-domain personality prediction: From video blogs to small group meetings. In *15th ACM International Conference on Multimodal Interaction*, 2013.
3. Oya Aran and Daniel Gatica-Perez. One of a kind: Inferring personality impressions in meetings. In *15th ACM International Conference on Multimodal Interaction*, 2013.
4. Afsaneh Asaei, Mohammad Golbabaee, Hervé Bourlard, and Volkan Cevher. Room acoustic modeling exploiting joint sparsity and lowrank structures. In *Signal Processing with Adaptive Sparse Structured Representations SPARS*, 2013.
5. Afsaneh Asaei, Bhiksha Raj, Hervé Bourlard, and Volkan Cevher. A multipath sparse beamforming method. In *Signal Processing with Adaptive Sparse Structured Representations SPARS*, 2013.
6. Alice Aubert, Romain Tavenard, Simon Malinowski, Thomas Guyet, René Quiniou, Jean-Marc Odobez, Remi Emonet, and Chantal Gascuel. Discovering temporal patterns in water quality time series, focusing on floods with the Ida method. In *European Geosciences Union*, 2013.
7. Chidansh A. Bhatt, Nikolaos Pappas, Maryam Habibi, and Andrei Popescu-Belis. I diap at mediaeval 2013: Search and hyperlinking task. In *MediaEval 2013 Workshop*, CEUR Workshop Proceedings. CEURWS.org, October 2013.
8. Chidansh A. Bhatt, Andrei Popescu-Belis, Maryam Habibi, Sandy Ingram, Stefano Masneri, Fergus McInnes, Nikolaos Pappas, and Oliver Schreer. Multi-factor segmentation for topic visualization and recommendation: the must-vis system. In *Proceedings of the 21st ACM International Conference on Multimedia*. ACM, 2013.
9. Hervé Bourlard, Marc Ferras, Nikolaos Pappas, Andrei Popescu-Belis, Steve Renals, Fergus McInnes, Peter Bell, Sandy Ingram, and Maël Guillemot. Processing and linking audio events in large multimedia archives: The eu inevent project. In *Workshop on Speech, Language and Audio in Multimedia*, July 2013.
10. Milos Cernak, Petr Motlicek, and Philip N. Garner. On the (un) importance of the contextual factors in hmm-based speech synthesis. In *Proceedings of the IEEE Intl. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8140 – 8143, May 2013.
11. Milos Cernak, Xingyu Na, and Philip N. Garner. Syllable-based pitch encoding for low bit rate speech coding with recognition/synthesis architecture. In *Proc. of Interspeech 2013*, August 2013.
12. Ivana Chingovska, André Anjos, and Sébastien Marcel. Anti-spoofing in action: joint operation with a verification system. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Biometrics*, June 2013.
13. Ivana Chingovska, Jinwei Yang, Zhen Lei, Dong Yi, Stan Z.Li, Olga Kahm, Naser Damer, Christian Glaser, Arjan Kuijper, Alexander Nouak, Jukka Komulainen, Tiago de Freitas Pereira, Shubham Gupta, Shubham Bansal, Shubham Khandelwal, Ayush Rai, Tarun Krishna, Dushyant Goyal, Muhammad-Adeel Waris, Honglei Zhang, Iftikhar Ahmad, Serkan Kiranyaz, Moncef Gabbouj, Roberto Tronci, Maurizio Pili, Nicola Sirena, Fabio Roli, Javier Galbally, Julian Fierrez, Allan Pinto, Helio Pedrini, William Robson Schwartz, Anderson Rocha, André Anjos, and Sébastien Marcel. The 2nd competition on counter measures to 2d face spoofing attacks. In *International Conference of Biometrics 2013*, June 2013.
14. Tiago de Freitas Pereira, André Anjos, José Mario De Martino, and Sébastien Marcel. Can face anti-spoofing countermeasures work in a real world scenario? In *International Conference on Biometrics*, June 2013.
15. Trinh-Minh-Tri Do, Kyriaki Kalimeri, Bruno Lepri, Fabio Pianesi, and Daniel Gatica-Perez. Inferring social activities with mobile sensor networks. In *15th ACM International Conference on Multimodal Interaction*, December 2013.
16. Charles Dubout and Francois Fleuret. Accelerated training of linear object detectors. In *CVPR 2013 Workshop on Structured Prediction*, 2013.
17. Charles Dubout and Francois Fleuret. Deformable part models with individual part scaling. In *British Machine Vision Conference*, 2013.



18. Nesli Erdogmus and Sébastien Marcel. Spoofing attacks to 2d face recognition systems with 3d masks. In *International Conference of the Biometrics Special Interest Group*, September 2013.
19. Nesli Erdogmus and Sébastien Marcel. Spoofing in 2d face recognition with 3d masks and antispoofing with kinect. In *Biometrics: Theory, Applications and Systems*, September 2013.
20. Marco Fornoni, Barbara Caputo, and Francesco Orabona. Multiclass latent locally linear support vector machines. In Cheng Soon Ong and Tu-Bao Ho, editors, *JMLR W&CP, Volume 29: Asian Conference on Machine Learning*, pages 229–244, 2013.
21. Kenneth Alberto Funes Mora. 3d head pose and gaze tracking and their application to diverse multimodal tasks. In *Doctoral consortium of the 15th ACM International Conference on Multimodal Interaction*, December 2013.
22. Kenneth Alberto Funes Mora, Laurent Son Nguyen, Daniel Gatica-Perez, and Jean-Marc Odobez. A semi-automated system for accurate gaze coding in natural dyadic interactions. In *15th ACM International Conference on Multimodal Interaction*. ACM, December 2013.
23. Kenneth Alberto Funes Mora and Jean-Marc Odobez. Person independent 3d gaze estimation from remote rgb-d cameras. In *International Conference on Image Processing*. IEEE, September 2013.
24. Arjan Gijsberts and Barbara Caputo. Exploiting accelerometers to improve movement classification for prosthetics. In *International Conference on Rehabilitation Robotics*, 2013.
25. Baran ozcu, Afsaneh Asaei, and Volkan Cevher. Manifold sparse beamforming. In *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, 2013.
26. Manuel Günther, Artur Costa-Pazo, Changxing Ding, Elhocine Boutellaa, Giovanni Chiacchia, Honglei Zhang, Marcus de Assis Angeloni, Vitomir Struc, Elie Khoury, Esteban Vazquez-Fernandez, Dacheng Tao, Messaoud Bengherabi, David Cox, Serkan Kiranyaz, Tiago de Freitas Pereira, Jerneja Zganec-Gros, Enrique Argones-Rua, Nicolas Pinto, Moncef Gabbouj, Flavio Simões, Simon Dobrisek, Daniel Gonzalez-Jiménez, Anderson Rocha, Mario Uliani Neto, Nikola Pavesic, Alexandre Falcao, Ricardo Violato, and Sébastien Marcel. The 2013 face recognition evaluation in mobile environment. In *The 6th IAPR International Conference on Biometrics*, June 2013.
27. Maryam Habibi and Andrei Popescu-Belis. Diverse keyword extraction from conversations. In *Proceedings of the ACL 2013 (51th Annual Meeting of the Association for Computational Linguistics), Short Papers*, 2013.
28. Najeh Hajlaoui. Are act's scores increasing with better translation quality? In *Are ACT's scores increasing with better translation quality?*, page 6, July 2013.
29. Najeh Hajlaoui and Andrei Popescu-Belis. Assessing the accuracy of discourse connective translations: Validation of an automatic metric. In *14th International Conference on Intelligent Text Processing and Computational Linguistics*, page 12. University of the Aegean, Springer, March 2013.
30. Alexandre Heili and Jean-Marc Odobez. Parameter estimation and contextual adaptation for a multi-object tracking crf model. In *IEEE Workshop on Performance Evaluation of Tracking and Surveillance*, 2013.
31. David Imseng and Hervé Bourlard. Speaker adaptive kullback-leibler divergence based hidden markov models. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.
32. David Imseng, Petr Motlicek, Philip N. Garner, and Hervé Bourlard. Impact of deep mlp architecture on different acoustic modeling techniques for under-resourced speech recognition. In *Proceedings of the IEEE workshop on Automatic Speech Recognition and Understanding*, December 2013.
33. Dinesh Babu Jayagopi and Jean-Marc Odobez. Given that, should I respond? contextual addressee estimation in multi-party human-robot interactions. In *Proceedings of Human Robot Interaction (HRI) Conference*, 2013.
34. Dinesh Babu Jayagopi, Samira Sheikhi, David Klotz, Johannes Wienke, Jean-Marc Odobez, Sebastian Wrede, Vasil Khalidov, Laurent Son Nguyen, Britta Wrede, and Daniel Gatica-Perez. The vernissage corpus: a conversational human-robot-interaction dataset. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, 2013.
35. Vasil Khalidov, Florence Forbes, and Radu Horaud. Alignment of binocular-binaural data using a moving audio-visual target. In *Proc. of IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, 2013.
36. Elie Khoury, Paul Gay, and Jean-Marc Odobez. Fusing matching and biometric similarity measures for face diarization in video. In *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*, pages 97–104. ACM, April 2013.
37. Elie Khoury, Manuel Günther, Laurent El Shafey, and Sébastien Marcel. On the improvements of uni-modal and bi-modal fusions of speaker and face recognition for mobile biometrics. In *Biometric Technologies in Forensic Science*, October 2013.



38. Elie Khoury, Bostjan Vesnicher, Javier Franco-Pedroso, Ricardo Violato, Zenelabidine Boulkenafet, Luis-Miguel Mazaira Fernandez, Mireia Diez, Justina Kosmala, Houssemeddine Khemiri, Tomas Cipr, Rahim Saedi, Manuel Günther, Jerneja Zganec-Gros, Ruben Zazo Candil, Flavio Simões, Messaoud Bengherabi, Augustin Alvarez Marquina, Mikel Penagarikano, Alberto Abad, Mehdi Boulayemen, Petr Schwarz, David Van Leeuwen, Javier Gonzalez-DomiVilda, Amparo Varona, Dijana Petrovska-Delacretaz, Pavel Matejka, Joaquin Gonzalez-Rodriguez, Tiago de Freitas Pereira, Farid Harizi, Luis Javier Rodriguez-Fuentes, Laurent El Shafey, Marcus Angeloni, German Bordel, Gérard Chollet, and Sébastien Marcel. The 2013 speaker recognition evaluation in mobile environment. In *The 6th IAPR International Conference on Biometrics*, June 2013.
39. Jukka Komulainen, Abdenour Hadid, Matti Pietikainen, André Anjos, and Sébastien Marcel. Complementary countermeasures for detecting scenic face spoofing attacks. In *International Conference on Biometrics*, June 2013.
40. Ilya Kuzborskij and Francesco Orabona. Stability and hypothesis transfer learning. In *International Conference on Machine Learning*, June 2013.
41. Ilya Kuzborskij, Francesco Orabona, and Barbara Caputo. From n to $n+1$: Multiclass transfer incremental learning. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, June 2013.
42. Leonidas Lefakis and Francois Fleuret. Reservoir boosting : Between online and offline ensemble learning. In *Proceedings of the international conference on Neural Information Processing Systems*, 2013.
43. Eric Malmi, Trinh-Minh-Tri Do, and Daniel Gatica-Perez. From foursquare to my square: Learning check-in behavior from multiple sources. In *The 7th International AAAI Conference on Weblogs and Social Media*, July 2013.
44. Alvaro Marcos-Ramiro, Daniel Pizarro-Perez, Marta Marron-Romera, Laurent Son Nguyen, and Daniel Gatica-Perez. Body communicative cue extraction for conversational analysis. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, April 2013.
45. Jesus Martinez-Gomez, Ismael Garcia-Varea, Miguel Cazorla, and Barbara Caputo. Overview of the imageclef 2013 robot vision task. In *Working Notes, CLEF 2013*, 2013.
46. Thomas Meyer, Cristina Grisot, and Andrei Popescu-Belis. Detecting narrativity to improve english to french translation of simple past verbs. In *Proceedings of the 1st DiscoMT Workshop at ACL 2013 (51st Annual Meeting of the Association for Computational Linguistics)*, pages 33–42, June 2013.
47. Thomas Meyer and Lucie Polakova. Machine translation with many manually labeled discourse connectives. In *Proceedings of the 1st DiscoMT Workshop at ACL 2013 (51st Annual Meeting of the Association for Computational Linguistics)*, pages 43–50, June 2013.
48. Thomas Meyer and Bonnie Webber. Implication of discourse connectives in (machine) translation. In *Proceedings of the 1st DiscoMT Workshop at ACL 2013 (51st Annual Meeting of the Association for Computational Linguistics)*, pages 19–26, June 2013.
49. Gelareh Mohammadi, Sunghyun Park, Kenji Sagae, Alessandro Vinciarelli, and Louis-Philippe Morency. Who is persuasive? the role of perceived personality and communication modality in social multimedia. In *International Conference on Multimodal Interaction*, 2013.
50. Petr Motlicek, Philip N. Garner, Namhoo Kim, and Jeongmi Cho. Accent adaptation using subspace gaussian mixture models. In *The 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 7170–7174. IEEE, May 2013.
51. Petr Motlicek, David Imseng, and Philip N. Garner. Crosslingual tandem-sgmm: Exploiting out-of-language data for acoustic model and feature level adaptation. In *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013)*, pages 510–514. ISCA - International Speech Communication Association, ISCA, August 2013.
52. Petr Motlicek, Daniel Povey, and Martin Karafiat. Feature and score level combination of subspace gaussians in lvcsr task. In *The 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 7604–7608. IEEE, May 2013.
53. Laurent Son Nguyen, Alvaro Marcos-Ramiro, Marta Marron-Romera, and Daniel Gatica-Perez. Multimodal analysis of body communication cues in employment interviews. In *15th ACM International Conference on Multimodal Interaction Proceedings*, 2013.
54. Youssef Oualil, Mathew Magimai-Doss, Friedrich Faubel, and Dietrich Klakow. A probabilistic framework for multiple speaker localization. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.
55. Dimitri Palaz, Ronan Collobert, and Mathew Magimai-Doss. Estimating phoneme class conditional probabilities from raw speech signal using convolutional neural networks. In *Proceedings of Interspeech*, August 2013.



56. Nikolaos Pappas, Georgios Katsimpras, and Efstathios Stamatatos. Distinguishing the popularity between topics: A system for up-to-date opinion retrieval and mining in the web. In *14th International Conference on Intelligent Text Processing and Computational Linguistics*. LNCS, ACM, 2013.
57. Nikolaos Pappas and Andrei Popescu-Belis. Combining content with user preferences for ted lecture recommendation. In *11th International Workshop on Content Based Multimedia Indexing*. IEEE, 2013.
58. Nikolaos Pappas and Andrei Popescu-Belis. Sentiment analysis of user comments for one-class collaborative filtering over ted talks. In *36th ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2013.
59. Ramya Rasipuram, Peter Bell, and Mathew Magimai-Doss. Grapheme and multilingual posterior features for under-resourced speech recognition: A study on scottish gaelic. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.
60. Ramya Rasipuram and Mathew Magimai-Doss. Improving grapheme-based asr by probabilistic lexical modeling approach. In *Proceedings of Interspeech*, 2013.
61. Ramya Rasipuram, Marzieh Razavi, and Mathew Magimai-Doss. Probabilistic lexical modeling and unsupervised training for zero-resourced asr. In *Proceedings of the IEEE workshop on Automatic Speech Recognition and Understanding*, December 2013.
62. Edgar Roman-Rangel, Jean-Marc Odobez, and Daniel Gatica-Perez. Evaluating shape descriptors for detection of maya hieroglyphs. In *Proc. Mexican Conf. on Pattern Recognition*, June 2013.
63. Mickael Rouvier, Gregor Dupuy, Paul Gay, Elie Khoury, Teva Merlin, and Sylvain Meignier. An open-source state-of-the-art toolbox for broadcast news diarization. In *INTERSPEECH*, August 2013.
64. Rahim Saedi, Kong Aik Lee, Tomi Kinnunen, Tawfik Hasan, Benoit Fauve, Pierre-Michel Bousquet, Elie Khoury, Pablo Luis Sordo Martinez, Jia Min Karen Kua, Changhuai You, Hanwu Sun, Anthony Larcher, Padmanabhan Rajan, Ville Hautamaki, Cemal Hanilci, Billy Braithwaite, Gonzalez-Hautamaki Rosa, Seyed Omid Sadjadi, Gang Liu, Hynek Boril, Navid Shokouhi, Driss Matrouf, Laurent El Shafey, Pejman Mowlaei, Julien Epps, Tharmarajah Thiruvaran, David Van Leeuwen, Bin Ma, Haizhou Li, John Hansen, Jean-François Bonastre, Sébastien Marcel, John Mason, and Eliathamby Ambikairajah. I4u submission to nist sre 2012: a large-scale collaborative effort for noise-robust speaker verification. In *INTERSPEECH*, August 2013.
65. Lakshmi Saheer and Milos Cernak. Automatic staging of audio with emotions. In *International Conference on Affective Computing and Intelligent Interaction*, 2013.
66. Lakshmi Saheer and Blaise Potard. Understanding factors in emotion perception. In *ISCA Speech Synthesis Workshop*, 2013.
67. Albert Ali Salah, Hayley Hung, Oya Aran, and Hatice Gunes. Creative applications of human behavior understanding. hbu 2013: 1-14. In *Human Behavior Understanding*, pages 1–14, 2013.
68. Dairazalia Sanchez-Cortes, Joan-Isaac Biel, Shiro Kumano, Junji Yamato, Kazuhiro Otsuka, and Daniel Gatica-Perez. Inferring mood in ubiquitous conversational video. In *12th International Conference on Mobile and Ubiquitous Multimedia*. ACM Press, December 2013.
69. Darshan Santani and Daniel Gatica-Perez. Revisiting the generality of the rank-based human mobility model. In *Proceedings of 3rd International Workshop on Pervasive Urban Applications (PURBA)*, 2013.
70. Darshan Santani and Daniel Gatica-Perez. Speaking swiss: Languages and venues in foursquare. In *Proceedings of the 21st ACM International Conference on Multimedia*, 2013.
71. A. Sapru and Hervé Bourlard. Automatic social role recognition in professional meetings using conditional random fields. In *Proceedings of Interspeech*, 2013.
72. A. Sapru and Hervé Bourlard. Investigating the impact of language style and vocal expression on social roles of participants in professional meetings. In *Affective Computing and Intelligent Interaction*, page 6, September 2013.
73. Samira Sheikhi, Dinesh Babu Jayagopi, Vasil Khalidov, and Jean-Marc Odobez. Context aware addressee estimation for human robot interaction. In *Proceedings of the 6th Workshop on Eye Gaze in Intelligent Human Machine Interaction: Gaze in Multimodal Interaction*, 2013.
74. Samira Sheikhi, Vasil Khalidov, David Klotz, Britta Wrede, and Jean-Marc Odobez. Leveraging the robot dialog state for visual focus of attention recognition. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, 2013.
75. Gyorgy Szaszak and Andras Beke. Using phonological phrase segmentation to improve automatic keyword spotting for the highly agglutinating hungarian language. In *Proc. of Interspeech*, 2013.



76. Gyorgy Szaszak and Philip N. Garner. Evaluating intra- and crosslingual adaptation for non-native speech recognition in a bilingual environment. In *Proceedings of the 4th IEEE International Conference on Cognitive Infocommunications*, pages 357–361. IEEE, December 2013.
77. Raphael Sznitman, Carlos Becker, Francois Fleuret, and Pascal Fua. Fast object detection with entropy-driven evaluation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2013.
78. Mohammad J. Taghizadeh, Reza Parhizkar, Philip N. Garner, and Hervé Bourlard. Euclidean distance matrix completion for ad-hoc microphone array calibration. In *Proceedings IEEE International Conference On Digital Signal Processing*, July 2013.
79. Romain Tavenard, Remi Emonet, and Jean-Marc Odobez. Time-sensitive topic models for action recognition in videos. In *IEEE International Conference on Image Processing*, 2013.
80. Raphael Ullmann, Hervé Bourlard, Jens Berger, and Anna Llagostera Casanovas. Noise intrusiveness factors in speech telecommunications. In *Proceedings of the AIA-DAGA 2013 International Conference on Acoustics*, pages 436–439, March 2013.
81. Majid Yazdani, Ronan Collobert, and Andrei Popescu-Belis. Learning to rank on network data. In *Mining and Learning with Graphs*, 2013.
82. Majid Yazdani and Andrei Popescu-Belis. Computing text semantic relatedness using the contents and links of a hyper-text encyclopedia. In *International Joint Conference on artificial intelligence*, 2013.
83. Sree Harsha Yella and Hervé Bourlard. Improved overlap speech diarization of meeting recordings using long-term conversational features. In *ICASSP*, 2013.
84. Sree Harsha Yella and Hervé Bourlard. Improved overlap speech diarization of meeting recordings using long-term conversational features. In *ICASSP*, 2013.





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

