

I-SITE CAP 20-25

Programme Instruments – bilan 2017-21

Dans le dossier initial, le programme « INSTRUMENTS » avait pour vocation d'être un support unissant les compétences techniques du site pour lever les verrous technologiques rencontrés dans la réalisation des enjeux liés aux 4 Challenges. Les développements dans les deux domaines techniques des réseaux de capteurs sans fils et des technologies du « bigdata » (entrepôt de données, méthodes d'apprentissage automatique, ...) ont été mis en priorité.

Le programme « INSTRUMENTS » CAP2025-phase 1 s'est donné comme objectifs principaux de :

1. Répondre aux problématiques technologiques des quatre challenges scientifiques CAP2025, en priorité sur les deux domaines des réseaux de capteurs sans fil et du bigdata.

Dans un premier temps le programme « INSTRUMENTS » a répondu aux besoins déjà répertoriés par les challenges avec l'implication forte dans 2 projets structurants dont la maturation était différente.

- Les besoins de surveillance des systèmes agro-alimentaires du projet CEBA et de l'instrumentation associée des sites pilotes du programme environnemental du CH 1 (du capteur à l'analyse des données) ont drainé l'essentiel des forces du programme « INSTRUMENTS ». Ce dernier a contribué fortement tant sur la partie CLOUD que sur la partie instrumentation et réseau de terrain sur les sites (Développement de nœuds intelligents communicants et de protocole de communication permettant de collecter de manière fiable des données environnementales in situ).

- Le programme « INSTRUMENTS », déjà présent par quelques-uns de ses membres sur des actions d'analyse de données dans le domaine de la santé, a contribué à la définition et la mise en place du projet EMOB du CH3 débuté en 2020, projet phare du challenge sur 'la mobilité personnalisée comme facteur-clé de la santé'. Le programme « INSTRUMENTS » coordonne la partie instrumentale destinée tout d'abord à produire un dispositif innovant de capture des types d'activités et de sédentarité. Ce dispositif est le point clef pour la conception par le programme « INSTRUMENTS » du système intelligent de tri des données, de classement, d'interprétation de la mobilité. Le programme « INSTRUMENTS » met en œuvre également les dispositifs de transmission sans fil, la conception et l'implémentation de l'entrepôt de données. Grâce aux recherches fondamentales qu'il a initié en amont, le programme « INSTRUMENTS » contribuera à toute la partie qualité et interprétation des données par la conception de techniques innovantes d'Intelligence Artificielle.

Les actions entreprises depuis 2017 sur les "capteurs et réseaux de capteurs" ont permis des avancées autour de la technologie de communication sans fil longue distance déployée par le programme « INSTRUMENTS » en articulation avec le CPER ConnecSenS. Ces technologies innovantes intéressent le CH 4 pour l'instrumentation des sites volcaniques en vue de leur surveillance. Un premier réseau de capteurs communicants a été déployé avec succès en 2019 et 2020 permettant de connecter en direct l'Etna à Clermont-Ferrand. L'aspect recherche amont s'est aussi exprimé avec le CH4, dans la contribution par le programme « INSTRUMENTS » à la réalisation et tests d'un détecteur de radon pour la compréhension de l'évolution des réservoirs magmatiques de volcans actifs.

Depuis début 2020, le programme « INSTRUMENTS » a conduit une action volontariste afin de valoriser les résultats de ses recherches et rendre visible son savoir-faire hors du site. A ce titre, deux déclarations d'invention ont été déposées auprès du CNRS et de l'INRAe. Un contrat de collaboration de recherche a été établi avec un partenaire industriel et une commercialisation du nœud communicant avec ce partenaire est en cours de discussion. Enfin, ces travaux innovants ont permis à

l'UCA de se positionner comme un partenaire majeur du projet EQUIPEX+ PIA3 2021 – 2028 TERRA FORMA (Concevoir et tester l'observatoire intelligent des territoires à l'heure de l'Anthropocène).

2. Fédérer les expertises du site et favoriser l'interaction avec les acteurs des axes stratégiques

Pour la thématique "capteurs et réseaux de capteurs", le programme « INSTRUMENTS » s'est structuré en s'appuyant sur la communauté ConneSenS du CPER, qui réunissait des laboratoires du site autour d'actions de recherche et d'innovation sur les capteurs et réseaux de capteurs pour l'environnement. Le développement des travaux dans le domaine du "bigdata" s'est appuyé sur le CPER AUDACE et la communauté associée. Cette initiative a permis la consolidation de la plateforme de mise au point et d'expérimentation d'architectures d'intégration, de gestion de données et d'algorithmes d'apprentissage automatique. L'initiative a également contribué au développement de travaux de recherche autour de l'optimisation des entrepôts de données.

Les travaux de recherche impulsés par le programme « INSTRUMENTS » ont permis une ouverture sur la communauté IA du site Clermontois et également la consolidation des synergies avec des acteurs du monde médical, initiant ainsi un programme de recherche à plus long terme autour de l'exploitation des données en santé.

Sur le plan opérationnel, afin de créer une dynamique nouvelle de recherche et développement autour des capteurs communicants et du cloud environnemental auprès de la communauté académique du site (étudiants, chercheurs, ingénieurs ...), le programme « INSTRUMENTS » a mis en place plusieurs actions incitatives :

- Des appels à proposition de sujets de thèse et de stage de Master ont été lancés dès 2017. Le programme « INSTRUMENTS » a privilégié les sujets qui concernent une problématique scientifique d'un challenge ou permettant de lever un verrou technologique par une recherche amont.
- Organisation à Clermont-Ferrand de la Conférence 'XLDB2017' (100 participants), forum majeur qui permet la rencontre des acteurs académiques et industriels clés dans le domaine du bigdata. Initiée en 2007 à Stanford, 'XLDB2017' est la première édition de cette conférence organisée en dehors des États-Unis.
- Organisation de moments d'échanges sous forme d'ateliers. L'objectif est de partager et de faire connaître les résultats du programme « INSTRUMENTS » au-delà du groupe de l'action, ainsi que ses compétences scientifiques et techniques. Cela a permis l'élargissement de la communauté du programme « INSTRUMENTS » et l'initiation de nouveaux projets (partie instrumentation dans le projet EMOB et actions avec le CH4).
- Installation d'une salle technique de 200 m² sur le campus des Cézeaux, à disposition de tout acteur du site travaillant sur les capteurs / réseaux de capteurs. Aménagement d'une partie de la salle en espace d'exposition.
- Développement des recherches en amont afin de maintenir et développer l'excellence sur nos domaines d'expertise.
- La réalisation de deux projets pédagogiques innovants, soutenus par le programme LEARN'IN Auvergne : laboratoire de formation à la conception et à la mise en œuvre d'IoT; mise en commun de ressources pédagogiques sur l'Internet des objets connectés.

- Présentation des réalisations/publications les plus remarquables du programme

<p>Déclarations de logiciel, N°13473-02 Valorisation avec la société YesItIs</p>	<p>Les réseaux communicants mis en place dans les différents projets de recherche menés dans le cadre de l'axe Instruments sont de type LoRaWAN et s'appuient sur le nœud communicant SoLo développé par les laboratoires de l'UCA, du CNRS et de l'INRAE. Ce nœud communicant est en cours de valorisation auprès de la Société Yesitis, suite à un contrat de collaboration de recherche et à une déclaration de logiciel auprès du CNRS. Une autre déclaration d'invention sur la partie protocole réseau a été déposée auprès de l'INRAE en 2020.</p>
<p>L Terray et al., 'From Sensor to Cloud: An IoT Network of Radon Outdoor Probes to Monitor Active Volcanoes. Sensors 20(10): 2755 (2020). 'https://hal.archives-ouvertes.fr/hal-02572054/</p>	<p>While radon in soil gases has been identified for decades as a potential precursor of volcanic eruptions, there has been a recent interest for monitoring radon in air on active volcanoes. We present here the first network of outdoor air radon sensors that was installed successfully on Mt. Etna volcano, Sicily, Italy in September 2019. Small radon sensors designed for workers and home dosimetry were tropicalized in order to be operated continuously in harsh volcanic conditions with an autonomy of several months. Two stations have been installed on the south flank of the volcano at ~3000 m of elevation. A private network has been deployed in order to transfer the measurements from the stations directly to a server located in France, using a low-power wide-area transmission technology from Internet of Things (IoT) called LoRaWAN. Data finally feed a data lake, allowing flexibility in data management and sharing. A first analysis of the radon datasets confirms previous observations, while adding temporal information never accessed before. The observed performances confirm IoT solutions are very adapted to active volcano monitoring in terms of range, autonomy, and data loss.</p>
<p>Alexandre Guitton, Megumi Kaneko: Improving LoRa Scalability by a Recursive Reuse of Demodulators. GLOBECOM 2020.</p>	<p>Long Range (LoRa) is a protocol that enables low-power wireless communications over long distances for a wide range of IoT applications. Its main drawback is its limited throughput, which is further reduced by the small number of demodulators in the hardware of the gateway. In this paper, we propose to use each demodulator as efficiently as possible. To do so, we reuse them for short frames during the preamble of long frames. By smartly planning the demodulation of multiple frames, the proposed method enables a recursive reuse of each demodulator. Compared to the benchmark packet arbiter policy, our method is shown to offer throughput and fairness enhancements even with a large number of users, thereby improving the scalability of LoRa systems. Our simulation results show that when the number of nodes is large, 6.5% more frames are decoded and the rate fairness among nodes is improved by 11%.</p>

<p>EN-NEJJARY, D., PINET, F., KANG, M. - 2019. A Method to Improve the Performance of Raster Selection Based on a User-defined Condition: an Example of Application for Agri-environmental Data. <i>Advances in Intelligent Systems and Computing</i>, 190-201., Springer.</p>	<p>More and more environmental and agricultural data are now acquired with a high precision and temporal frequency. These data are often represented in the form of rasters and are useful for agricultural activities or climate change analyses. In this paper, we propose a new method to process very large raster. We present a new technique to improve the execution time of the selection and calculation of data summaries (e.g., the average temperature for a region) on a temporal sequence of rasters. We illustrate the use of our approach on the case of temperature data, which is important information both for agriculture and for climate change analyses. We have generated several data sets in order to analyze the influence of the different value properties on the process performance. One of our final goals is to provide information about the value conditions in which the proposed processing should be used.</p>
<p>Abdoul Jalil Djiberou Mahamadou, Violaine Antoine, Engelbert Mephu Nguifo, Sylvain Moreno: Categorical fuzzy entropy c-means. <i>FUZZ-IEEE 2020</i>: 1-6</p>	<p>Hard and fuzzy clustering algorithms are part of the partition-based clustering family. They are widely used in real-world applications to cluster numerical and categorical data. While in hard clustering an object is assigned to a cluster with certainty, in fuzzy clustering an object can be assigned to different clusters given a membership degree. For both types of method an entropy can be incorporated into the objective function, mostly to avoid solutions raising too much uncertainties. In this paper, we present an extension of a fuzzy clustering method for categorical data using fuzzy centroids. The new algorithm, referred to as Categorical Fuzzy Entropy (CFE), integrates an entropy term in the objective function. This allows a better fuzzification of the cluster prototypes. Experiments on ten real-world data sets and statistical comparisons show that the new method can efficiently handle categorical data.</p>
<p>Ousmane Issa, Angela Bonifati, Farouk Toumani: Evaluating Top-k Queries with Inconsistency Degrees. <i>Proc. VLDB Endow.</i> 13(11): 2146-2158 (2020)</p>	<p>We study the problem of augmenting relational tuples with inconsistency awareness and tackling top-k queries under a set of denial constraints (DCs). We define a notion of inconsistent tuples with respect to a set of DCs and define two measures of inconsistency degrees, which consider single and multiple violations of constraints. In order to compute these measures, we leverage two models of provenance, namely why-provenance and provenance polynomials. We investigate top-k queries that allow to rank the answer tuples by their inconsistency degrees. Since one of our measure is monotonic and the other non-monotonic, we design an integrated top-k algorithm to compute the top-k results of a query w.r.t. both inconsistency measures. By means of an extensive experimental study, we gauge the effectiveness of inconsistency-aware query answering and the efficiency of our algorithm with respect to a baseline, where query results are fully computed and ranked afterwards.</p>