



FÉDÉRATION INFORMATIQUE CENTRE VAL
DE LOIRE

ICVL

Dossier scientifique

Dossier d'évaluation pour le HCERES

Laboratoires de recherche :

LI : Laboratoire d'Informatique (EA 6300)

LIFO : Laboratoire d'Informatique Fondamentale d'Orléans (EA 4022)

Porteurs :

Agata Savary (LI)

Mirian Halfeld Ferrari Alves (LIFO)

9 janvier 2017



Présentation synthétique de l'entité

La fédération **Informatique Centre Val de Loire (ICVL)** est une nouvelle structure fédérative en cours de création administrative au moment du dépôt de ce dossier. Elle comprend deux laboratoires académiques de la Région Centre Val de Loire : le **LIFO** (Laboratoire d'Informatique Fondamentale d'Orléans) et le **LI** (Laboratoire d'Informatique de l'Université François Rabelais de Tours). Elle jouit de la tutelle de trois institutions : l'Université d'Orléans, l'Université François Rabelais Tours et l'INSA Centre Val de Loire. Conjointement avec le LI et le LIFO, ICLV demande l'association CNRS pour 2018-2002. Elle est dirigée par un directeur et son adjoint, et pilotée par un conseil et un comité scientifique.

La fédération vise la construction d'un pôle d'excellence dans l'*acquisition et le traitement efficaces, fiables et sécurisés de l'information*. Le projet scientifique de l'ICVL, construit selon une approche consultative, est structuré en 3 axes et 9 sous-axes :

- A** Données massives, complexes ou hétérogènes (47 membres de 4 équipes) :
 - A1** Extraction des connaissances et reconnaissance des formes
 - A2** Données du Web
 - A3** Traitement automatique des langues
 - A4** Analyse et traitement d'images
- B** Algorithmique et optimisation (18 membres de 3 équipes) :
 - B1** Algorithmique des graphes et complexité exponentielle
 - B2** Ordonnancement et transport
- C** Sécurité, fiabilité et performances (27 membres de 4 équipes) :
 - C1** Sécurité des systèmes
 - C2** Programmes corrects et vérification de systèmes
 - C3** Parallélisme et systèmes repartis

L'ICVL rentrera dans le contrat 2018-2022 avec 8 actions inter-laboratoires, une dizaine de projets financés inter-axes ou inter-laboratoires, ainsi que quelques dizaines de facteurs de forte visibilité internationale, nationale et régionale.

Les plus-values de la fédération devraient provenir du fait d'atteindre une masse critique pour l'ouverture vers des horizons ambitieux (projets européens par exemple) et d'offrir une source d'innovation grâce à des interactions entre domaines relativement distincts (représentés par les axes).

Executive Summary

ICVL (Informatique Centre Val de Loire) is a new research federation, to be administratively created shortly after the submission of this report. It comprises two academic research laboratories of the Centre Val de Loire Region in France : **LIFO** (Laboratoire d'Informatique Fondamentale d'Orléans) and **LI** (Laboratoire d'Informatique de l'Université François Rabelais de Tours). ICVL is affiliated with three institutions : the University of Orleans, the François Rabelais University of Tours and the INSA Centre Val de Loire institute. Together with LI and LIFO, ICVL requires CNRS affiliation for 2018-2022. Its coordination lies in the hands of its director and vice-director, its council and its scientific committee.

The Federation aims at becoming a hub of excellence in *efficient, reliable and secure data acquisition and processing*. Its scientific project, resulting from a consultative approach, is structured around 3 axes and 9 sub-axes :

- A** Voluminous, complex and heterogeneous data (47 members from 4 teams) :
 - A1** Knowledge extraction and pattern recognition
 - A2** Web data
 - A3** Natural language processing
 - A4** Image analysis and processing
- B** Algorithmics and optimization (18 members from 3 teams) :
 - B1** Graph algorithmics and exponential complexity
 - B2** Scheduling and transport
- C** Security, reliability and efficiency (27 members from 4 teams) :
 - C1** Security of operating systems
 - C2** Correct programs and system verification
 - C3** Parallelism and distributed systems

At the beginning of the 2018-2022 evaluation period, ICVL will benefit from 8 transversal research actions, a dozen of funded projects, as well as several dozens of international, national and regional visibility factors.

The added value of the Federation should stem from at least two facts : attaining a critical mass necessary for challenging initiatives (such as European projects), and bringing together researchers from relatively distant domains (represented by axes) so as to achieve strong innovation effects.

Projet scientifique pour la période 2018-2022

1 Introduction

La fédération **Informatique Centre Val de Loire (ICVL)** est issue de la rencontre de deux laboratoires académiques de la Région Centre Val de Loire : le **LIFO (Laboratoire d'Informatique Fondamentale d'Orléans)** et le **LI (Laboratoire d'Informatique)** de l'Université François Rabelais de Tours), qui sont de tailles et structures comparables et ont tissé des liens bilatéraux sur une décennie. La table 1 indique la structuration du LI (autour de trois équipes) et du LIFO (autour de cinq équipes).

Équipe LI	Thématiques
BdTIn OC RFAI	Bases de Données et Traitement des Langues Naturelles Ordonnancement et Conduite Reconnaissance des Formes et Analyse d'Images
Équipe LIFO	Thématiques
CA GAMoC LMV Pamda SDS	Contraintes et Apprentissage Graphes, Algorithmes et Modèles de Calcul Logique, Modélisation et Vérification Parallélisme, systèmes distribués et bases de données Sécurité et Distribution des Systèmes

TABLE 1 – Équipes du LI et du LIFO

Via la fédération, les laboratoires ambitionnent l'enrichissement de leurs recherches en informatique, en vue de la construction d'un pôle d'excellence dans *l'acquisition et le traitement efficaces, fiables et sécurisés de l'information*. La fédération demande son association CNRS pour 2018-2022. Elle jouit de la tutelle de trois institutions :

- l'Université d'Orléans,
- l'Université François Rabelais Tours,
- l'INSA Centre Val de Loire

Il s'agit d'une **nouvelle structure fédérative**, en cours de création officielle au moment du dépôt de ce dossier, bien que son activité scientifique soit déjà en route. **Ce document présente une synthèse du projet scientifique de l'ICVL** pour la période 2018-2022 (cf. le projet complet en ligne¹) : les fondements de sa construction, sa structuration en axes et sous-axes scientifiques, les collaborations déjà en cours et devant être poursuivies, ainsi que la plus-value attendue par sa mise en place.

Le projet scientifique est le fruit d'une **démarche** largement **consultative** menées pendant plus de deux ans, incluant :

- des réunions du binôme chef de projet avec les équipes des deux laboratoires,
- une enquête ouverte à tous les membres sur les valeurs communes à défendre au sein de la fédération, ses objectifs, son périmètre d'actions et les craintes vis-à-vis du projet,
- la présentation du projet aux trois tutelles, ainsi qu'à un représentant du CNRS,
- des rencontres avec d'autres fédérations voisines (MIREs, NormaSTIC, Denis Poisson),
- un appel aux actions inter-labo (9 soumissions) et à contributions scientifiques existantes (59 soumissions),
- des ateliers thématiques à la base des contributions (lors des Journées Informatiques en Région Centre à Bourges en mai 2015),
- la rédaction du projet scientifique à la base des ateliers, et son amélioration suite à des retours des membres,
- validation du projet scientifique par les conseils de laboratoire.

L'enquête mentionnée plus haut a donné notamment lieu à la définition des **objectifs** de la fédération :

- valoriser et améliorer la visibilité des activités de recherche informatique en Région Centre ;
- susciter et favoriser des projets de recherche au niveau régional, national et international ;

- dynamiser l'activité de recherche au sein des deux Parties et encourager leur collaboration.

2 Structuration administrative et scientifique

Selon ses **statuts**, la fédération ICVL inclut tous les membres enseignants-chercheurs, chercheurs, personnel administratif et doctorants des laboratoires LIFO et LI. Elle est dirigée par un **directeur** et un directeur **adjoint** (qui échangent leurs rôles à la mi-mandat), élus par un vote direct pour la durée de la convention (signée pour chaque période de 5 ans). Leurs fonctions sont **non cumulables** avec toute autre fonction de direction de recherche dans leurs entités de rattachement respectives.

Le **conseil de fédération** élabore la politique de la fédération. Il est composé du directeur et de son adjoint, des directeurs des deux laboratoires et des représentants des membres de différentes catégories.

Le **comité scientifique**, réuni au moins deux fois pendant la durée de la convention, examine les résultats, la répartition des moyens et les orientations de la fédération. Il est composé du directeur et de son adjoint, des directeurs des deux laboratoires, des représentants des tutelles et de 4 personnalités scientifiques extérieures.

Au moment du dépôt de ce dossier aucune **convention** assurant l'application des statuts n'a encore été signée. Sa signature devrait avoir lieu prochainement, et dans tous les cas avant le début du contrat 2018-2022. La fédération n'emploie aucun personnel administratif qui lui soit dédié, elle bénéficie en revanche du support administratif des deux laboratoires membres. Elle n'a pas actuellement de projet d'achat de gros équipements. Ses sources principales de financements devraient provenir essentiellement des subventions des deux laboratoires membres et des trois tutelles. Un éventuel co-financement régional est également envisageable.

Malgré son cadre formel encore naissant, l'**activité scientifique** de l'ICVL a déjà débuté. Elle est structurée autour de trois notions :

- **Axes transversaux** — il s'agit de larges thématiques englobant les compétences et les sensibilités scientifiques représentées par les deux laboratoires. Ils doivent servir à l'affichage externe, mais aussi à la communication et structuration interne. Un axe est souvent transversal à deux laboratoires, mais peut aussi correspondre à l'activité d'une ou plusieurs équipes d'un même laboratoire.

- **Actions** – elles correspondent à des collaborations sur de moyennes et longues durées, existant entre les deux laboratoires.
- **Initiatives** – ce sont des événements plus ponctuels qui visent la création ou le renforcement des actions (par exemple des stages co-encadrés, des journées thématiques, des séminaires, etc.). Les initiatives sont, dans la mesure du possible, financées par le budget de la fédération.

Suite à un travail consultatif important, 3 axes et 9 sous-axes transversaux, décrits dans la section suivante, ont été définis. Au moment du dépôt de ce dossier, **9 actions**, basées sur des collaborations nouvelles ou déjà existantes entre les deux laboratoires, sont en cours.

Parallèlement aux actions, la fédération peut constituer des **groupes de travail** dédiés à des sujets spécifiques. Trois groupes de travail créés jusqu'à présent concernent : (i) l'harmonisation des formations de Master informatique en Région Centre, (ii) le développement des liens avec la COMUE et la fédération MIRES, (iii) la préparation d'une demande d'accréditation auprès du CNRS.

3 Axes transversaux

La définition des axes et sous-axes transversaux a été basée notamment sur les résultats des ateliers thématiques menés par quelques dizaines de membres des deux laboratoires lors des Journées Informatiques en Région Centre en mai 2015. L'un des objectifs majeurs de cette structuration est que chaque membre de l'ICVL puisse s'identifier avec au moins un axe. Ainsi, les axes et les sous-axes peuvent englober des membres :

- des deux laboratoires,
- de plusieurs équipes d'un même laboratoire (axe C, sous-axe B2),
- d'une seule équipe (sous-axe A4).

Le rassemblement en axes a été guidé par la proximité thématique, la possibilité d'interaction plus immédiate et par les collaborations déjà existantes (voir Fig.1, p. 10). La perspective plus lointaine est celle des **interactions inter-axes**, qui induira des rapprochements entre des domaines plus éloignés, ce qui constituera précisément la **plus-value** principale de la fédération, par rapport à une simple co-existence des deux laboratoires en dehors d'une structure fédérative.

Les axes et sous-axes sont présentés ici de manière synthétique. Le texte complet du projet (disponible en ligne¹) inclut leur description détaillée, notamment les noms de leurs membres et les détails des verrous et techniques scientifiques.

3.1 Axe A : Données massives, complexes ou hétérogènes

Cet axe englobe **47 membres** de 4 équipes du LIFO (CA, GAMoC, LMV, Pamda) et 2 du LI (BdTln, RFAI). Il se divise en 4 **sous-axes** :

- A1 Extraction des connaissances et reconnaissance des formes (BdTln, CA, RFAI)
- A2 Données du Web (BdTln, GAMoC, LMV, Pamda)
- A3 Traitement automatique des langues (BdTln, CA)
- A4 Analyse et traitement d'images (RFAI)

Il s'intéresse à de nombreux **défis scientifiques** incluant notamment :

1. Intégration ou adaptation à l'environnement (e.g. : pour permettre l'interaction avec l'utilisateur et la prise en compte de ses préférences ; pour la prise en compte de l'évolution des données ; pour l'intégration et interopérabilité de données du web et du langage naturel)
2. Qualité, fiabilité et sécurité (e.g. pour une modélisation fine des connaissances ou pour l'évaluation de la qualité de données),
3. Complexité et optimisation (e.g. pour les traitements qui passent à l'échelle ou qui respectent les contraintes du temps réel),
4. Modélisation, représentation et génération (e.g. pour la prise en compte des données hétérogènes : structurées, XML, géolocalisées, bioinformatiques, issues des réseaux sociaux, textes, images et vidéos).

Il s'efforce de répondre à deux **besoins sociétaux** principaux :

- l'accès à l'information et création de connaissances,
- services à la personne, santé, transport, ville intelligente, production et énergie.

Il emploie des **techniques scientifiques** telles que la classification, la visualisation, l'extraction de données, l'apprentissage statistique et symbolique, ainsi que des **techniques de modélisation** telles que la logique, les grammaires d'arbres, les outils à états finis et les graphes.

Ses **points forts** incluent :

- Grand nombre de participants de diverses équipes des deux laboratoires.
- Nombreuses collaborations en cours ou débutantes.

Le **point faible** principal vient du grand nombre de participants qui pourrait rendre plus difficile le choix des thématiques prioritaires.

3.2 Axe B : Algorithmique et optimisation

Cet axe englobe **18 membres** d'une équipe du LIFO (**GAMoC**) et de 2 équipes du LI (**OC, RFAI**). Il se divise en 2 **sous-axes** :

B1 Algorithmique des graphes et complexité exponentielle (**GAMoC, OC**)

B2 Ordonnancement et transport (**OC, RFAI**)

Il s'intéresse aux **défis scientifiques** suivants :

1. Complexité et l'optimisation (e.g. pour atteindre la complexité optimale, globale et paramétrée, des problèmes NP-difficiles),
2. Modélisation et la structuration de l'information (e.g. pour résoudre des problèmes d'ordonnancement à contraintes multiples et dans de domaines d'application variés).

Il s'efforce de répondre à des **besoins sociétaux** autour des services à la personne dans la vie quotidienne, du transport, de la production et de l'énergie.

Il emploie des **techniques scientifiques** exactes comme approchées sans garantie de performance (décompositions et combinatoire de graphes, heuristiques, métaheuristiques). Les graphes sont ses **techniques de modélisation** principales. Ses **points forts** incluent :

- Le rapprochement naturel entre les deux équipes graphe-complexité, classées A+ à l'évaluation précédente.
- Une complémentarité des approches : théorique et pratique.

Le **point faible** principal est d'offrir un rapprochement naturel, mais qui représente un défis dû aux buts parfois dichotomiques entre théoriciens et praticiens.

3.3 Axe C : Sécurité, fiabilité et performances

Cet axe englobe **27 membres** de 4 équipes du LIFO (**GAMoC, LMV, Pamda, SDS**). Il se divise en 3 **sous-axes** :

C1 Sécurité des systèmes (**SDS, LMV**)

C2 Programmes corrects et vérification de systèmes (**LMV, Pamda**)

C3 Parallélisme et systèmes repartis (**GAMoC, LMV, Pamda**)

Il s'intéresse à de nombreux **défis scientifiques** incluant notamment :

1. Intégration ou adaptation à l'environnement (e.g. composition de programmes parallèles pour la visualisation scientifique),

2. Qualité, fiabilité et sécurité (e.g. pour fournir un système d'exploitation garantissant une sécurité en profondeur, avec un contrôle dynamique de services à degrés de fiabilité différents),
3. Complexité et optimisation (e.g. vérification de compilation pour des langages parallèles),
4. Modélisation, représentation et génération (e.g. pour des modèles non-conventionnels de temps en tant qu'espace continu, ou pour la conception de langages de programmation assurant la vérification et correction automatique des logiciels).

Il s'efforce de répondre à deux **besoins sociétaux** principaux :

- sécurité, sûreté et éthique,
- accès à l'information et création de connaissances.

Il emploie des **techniques scientifiques** telles que la vérification formelle, l'assistant de preuves, la programmation parallèle ou répartie, la programmation fonctionnelle, la cryptographie, la sécurisation et les approches par squelettes, ainsi que des **techniques de modélisation** telles que la logique, les grammaires d'arbres, les réseaux de Petri et les graphes.

Ses **points forts** incluent :

- Synergie entre thématiques relativement différentes,
- Initiation de plusieurs actions/projets avec les autres axes,
- Dynamique transversale au sein du LIFO.

Le **point faible** principal est de contenir uniquement des membres du LIFO.

En résumé, l'axe A, concentrant un grand nombre de chercheurs, représente des compétences variées en extraction de connaissances, constituant ainsi une grande force nationale dans ce domaine. Les interactions de cet axe avec l'axe C sont nombreuses. Elles sont dues aux besoins manifestes de passage à l'échelle et de solutions efficaces et sûres. L'axe B gagne en rapprochant ses travaux théoriques des recherches applicatives, ce qui ouvre la porte à de nouveaux projets. La fédération se construit dans la perspective de motiver et développer la collaboration entre groupes de recherche complémentaires et, donc, non seulement au sein de chaque axe mais aussi entre les axes. Son objectif à long terme est de devenir un pôle excellence dans l'*acquisition et le traitement efficaces, fiables et sécurisés de l'information*.

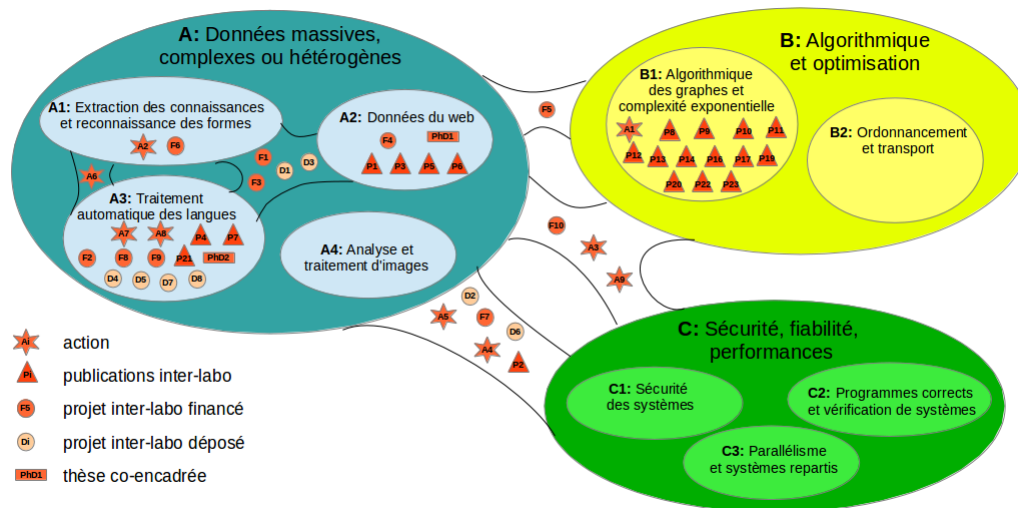


FIGURE 1 – Axe et sous-axes transversaux de la fédération ICVL et leurs interactions

La figure 1 résume la structuration de l'activité scientifique de l'ICVL en axes et sous-axes, ainsi que leurs activités et interactions de divers types existant au moment du dépôt de ce dossier. Celles-ci englobent :

- 8 **actions** inter-laboratoires en cours,
- 10 **projets** financés inter-axes ou inter-laboratoires **passés ou en cours**,
- 10 **projets** financés inter-axes ou inter-laboratoires **soumis**,
- 2 **thèses** de doctorat inter-laboratoires en cours,
- 24 **publications** inter-laboratoires.

Certains de ces éléments sont énumérés dans le chapitre suivant.

4 Actions, projets, ouverture internationale

Huit **actions** (collaborations sur de moyennes et longues durées, existant entre les deux laboratoires) sont en cours et devraient se poursuivre durant le contrat 2018-2022 :

1. Algorithmique exponentielle et paramétré en ordonnancement (axe B).
2. Hybridation des techniques de clustering et d'extraction de motifs (axe A).

3. Extraction de motifs séquentiels à grande échelle dans un environnement Hadoop/MapReduce (axes A et C).
4. Similarité de trajectoires et optimisation (axes A et C).
5. Masses de données dans une approche sémantique : consultation, maintenance, évolution (axe A).
6. Techniques d'apprentissage pour le TALN (axe A).
7. Analyse syntaxique (parsing) pour l'annotation de la temporalité (axe A).
8. Expressions polylexicales et parsing (axe A).
9. Éthique et risques des technologies numériques (axes A et C).

Le LI et le LIFO ont déjà une expérience de collaboration dans des réseaux et projets collaboratifs communs :

1. Projet ANR Codex (2009-2012),
2. Réseau COST PARSEME (2013-2017) - coordination,
3. Projet régional TEMPORAL (2013-2014),
4. Projet régional GIRAFON (2015-2017) - coordination,
5. Projet ANR PARSEME-FR (2016-2020),
6. Projet régional ODIL (2016-2017).

Plusieurs autres projets, notamment 2 propositions de projet européens, ont été récemment déposés et non retenus, mais évolueront vers de nouvelles propositions à re-soumettre au cours du prochain contrat. L'incitation et la formation au montage de projets européens est notamment au programme des activités à mener au sein de l'ICVL.

La fédération n'a pas vocation à rester fermée entre le LI et le LIFO, la collaboration avec d'autres laboratoires, d'autres universités françaises et étrangères est un des ses objectifs majeurs. Nous avons identifié 24, 14 et 2 facteurs de forte visibilité internationale, nationale et régionale, respectivement (ces facteurs correspondent à des projets ou collaborations listés dans l'annexe C du document en ligne¹ contenant le projet scientifique complet). Pour n'en mentionner que quelques-uns dont la portée devrait s'étendre au **contrat 2018-2022** :

- participation au Master Erasmus Mundus IT4BI avec Bruxelles, Barcelone, Berlin et Paris ; ce programme (se terminant en 2018) évoluera vers un programme Erasmus+ "Big Data Management and Analytics" (BDMA) avec le consortium élargi aux Pays-Bas (proposition retenue pour financement en juillet 2016),

- coordination du projet ANR bilatéral international "jeune chercheur" e-VRO avec les USA et la Colombie (2015-2019),
- collaboration de longue durée avec le Brésil, le Chili et le Sénégal (plusieurs projets bilatéraux ou co-tutelles de thèse),
- participation au Centre d'Excellence Africain en Mathématiques, Informatique et Technologies de l'Information et de la Communication (3 co-tutelles de thèse démarrant en 2016),
- participation aux projets ANR GraphEn (2015-2019) et PARSEME-FR (2016-2020),
- coordination d'un groupe de travail au sein d'un GDR : GT SDA2 du GDR IM (depuis 2015),
- coordination académique d'un projet national ATELYN (2016-2019),
- coordination du Réseau Thématique Régional sur le Risque Médical (2015-2018).

5 Conclusion

La fédération ICVL est une structure nouvelle, en cours de création administrative au moment du dépôt de ce dossier, mais dont l'activité scientifique montre déjà une bonne dynamique. En effet, nous avons à notre actif 8 actions inter-laboratoires en cours, 10 projets financés inter-axes ou inter-laboratoires passés ou en cours, 10 projets financés inter-axes ou inter-laboratoires soumis, 2 thèses de doctorat inter-laboratoires, 24 publications inter-laboratoires, ainsi que 24, 14 et 2 facteurs de forte visibilité internationale, nationale et régionale, respectivement.

La fédération a pour ambition la qualité et le dynamisme dans la recherche. Plus précisément elle vise la construction d'un pôle d'excellence dans l'*acquisition et le traitement efficaces, fiables et sécurisés de l'information*. Son projet scientifique, construit depuis 2014 selon une approche consultative, est structuré en 3 axes et 9 sous-axes transversaux.

Nos forces principales consistent en le rapprochement des deux laboratoires ayant des thématiques proches ou complémentaires et dans de nouvelles collaborations déjà entamées. Les plus-values de la fédération, par opposition à une co-habitation du LIFO et du LI hors structure fédérative, devraient provenir du fait d'atteindre une masse critique pour l'ouverture vers des horizons plus ambitieux (projets européens par exemple) et d'offrir une source d'innovation grâce à

des interactions entre domaines relativement distincts (représentés par les axes). La puissance interne des axes consiste, majoritairement, à faire émerger des recherches, possibles uniquement grâce à la mise en commun des compétences des deux laboratoires. Leur puissance externe repose sur l'offre d'un savoir faire qui peut s'adapter ou proposer des solutions à d'autres axes. L'ensemble de toutes ces collaborations permettra de donner des réponses pluridisciplinaires aux besoins sociétaux entrelacés entre les différentes thématiques des axes.

La fédération est d'ores et déjà un terrain de discussion pour de nouvelles perspectives, comme par exemple une fusion éventuelle des deux laboratoires ou bien une association de l'ICVL avec d'autres laboratoires ou fédérations.

Notes

¹Projet scientifique ICVL, publié comme rapport de recherche numéro RR-2016-01 ; disponible sur les sites <http://www.univ-orleans.fr/lifo/prodsci/rapports/RR/RR2016/RR-2016-01.pdf> ou <http://www.info.univ-tours.fr/~savary/ICVL/ICVL-projet-scientifique.pdf>