

Projet ANR-14-CE24-0002-01

**DYCI2 : Dynamiques créatives de
l'interaction improvisée**

Programme AAPG 2014

AAP Interactions des mondes physiques, de l'humain et du monde numérique

A	Identification.....	2
B	Résumé consolidé public	2
	B.1 Résumé consolidé public en français	2
	B.2 Résumé consolidé public en anglais	4
C	Mémoire scientifique	5
	C.1 Résumé du mémoire	5
	C.2 Enjeux et problématique, état de l'art.....	5
	C.3 Approche scientifique et technique	6
	C.4 Résultats obtenus.....	7
	C.5 Exploitation des résultats.....	10
	C.6 Discussion.....	11
	C.7 Conclusions	12
	C.8 Références	12
D	Liste des livrables.....	13
E	Impact du projet.....	14
	E.1 Indicateurs d'impact	14
	E.2 Liste des publications et communications.....	15
	E.3 Liste des éléments de valorisation	19
	E.4 Bilan et suivi des personnels recrutés en CDD (hors stagiaires)	22

A IDENTIFICATION

Acronyme du projet	DYCI2
Titre du projet	Dynamiques créatives de l'interaction improvisée
Coordinateur du projet (société/organisme)	IRCAM
Période du projet (date de début – date de fin)	01/03/2015 à 30/09/2018
Site web du projet, le cas échéant	dyci2.ircam.fr (voir note en fin de doc.)

Rédacteur de ce rapport	
Civilité, prénom, nom	Gérard Assayag
Téléphone	0667412004
Adresse électronique	gerard.assayag@ircam.fr
Date de rédaction	30/09/2018

Si différent du rédacteur, indiquer un contact pour le projet	
Civilité, prénom, nom	
Téléphone	
Adresse électronique	

Liste des partenaires présents à la fin du projet (société/organisme et responsable scientifique)	IRCAM (Gérard Assayag) Inria Nancy (Emmanuel Vincent) Université de La Rochelle (Sylvain Marchand)
---	--

B RÉSUMÉ CONSOLIDÉ PUBLIC

B.1 RÉSUMÉ CONSOLIDÉ PUBLIC EN FRANÇAIS

L'ordinateur compagnon créatif de nos improvisations musicales

DYCI2 vise les nouveaux besoins de création musicale et les nouveaux scénarios d'interaction multimedia par l'improvisation homme-machine

L'interaction improvisée entre des humains et des artefacts numériques est un domaine récent issu des études sur la créativité artificielle, qui convoque plusieurs problématiques de recherche très actives : l'apprentissage interactif, dont les modèles se construisent dans le temps-même de l'interaction et dont les résultats infléchissent les conditions de cette interaction; la perception artificielle, base de cette interaction; la modélisation de l'interaction sociale et expressive entre agents humains et/ou artificiels, dans ses dimensions à la fois anthropologique, sociale, linguistique, et informatique.

Les défis scientifiques du projet DYCI2 visent les nouveaux besoins en matière de création et de performance musicale, ainsi que les nouveaux scénarios d'interaction créative des utilisateurs avec les objets multimédias en modélisant dans tous ses aspects de perception, d'apprentissage, de générativité et d'interaction la situation de co-improvisation entre des agents humains ou artificiels. DYCI2 met en œuvre, avec la recherche scientifique et le développement technologique d'outils puissants pour la création, un grand nombre d'actions artistiques et pédagogiques au niveau international.

DYCI2 met en œuvre des méthodes d'écoute artificielle, d'apprentissage symbolique, et d'architectures d'interactions pour l'improvisation

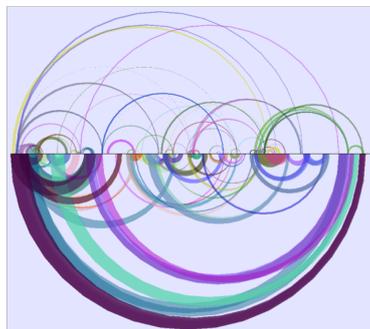
Les agents musicaux, réels (instrumentistes) ou virtuels (ordinateurs) doivent être capable de s'écouter les uns les autres afin de jouer ensemble de manière satisfaisante en conditions d'improvisation. L'ordinateur doit donc être en mesure par l'écoute artificielle d'analyser et de structurer le son musical en fonction du temps, verticalement (à un instant donné) et horizontalement (suivant l'axe temporel), par des méthodes de découverte, de séparation et de prédiction faisant appel à l'apprentissage. À partir des flux de données issues de cette étape d'écoute, il s'agit ensuite d'extraire des modèles symboliques de séquences musicales susceptibles de capturer les structures multidimensionnelles et multi-échelles propres à la musique, en particulier dans un contexte d'improvisation, et constituant des mémoires musicales richement informées. À partir de ces dernières, l'ordinateur détermine les interactions improvisées entre agents artificiels et humains en examinant les questions de dynamiques collectives, d'adaptation aux différentes échelles temporelles, de réactivité et de planification permettant différentes sortes de guidage adaptatif de l'improvisation en fonction de l'écoute.

Résultats majeurs du projet

Des techniques de traitement de signal et d'apprentissage profond, de modélisation statistique de séquences, ainsi que des modèles probabilistes de type bayésien ont été employées avec succès pour la partie modélisation du projet. Des architectures sophistiquées pour l'interaction combinant génération libre, structurée par planification et par réaction instantanée au contexte ont été mises au point et utilisées dans nombre d'applications artistiques. DYCI2 a donné lieu à plusieurs innovations scientifiques dont l'adaptation du Deep Learning à la séparation multicanal live, une première mondiale.

Production scientifique et brevets depuis le début du projet

Le logiciel de séparation de sources par Deep Learning a été déposé à l'APP et valorisé auprès de Samsung. Les résultats du projet ont été diffusés à la communauté scientifique, artistique et au grand public à travers un nombre important de publications, de concerts internationaux (Montreux Jazz Festival, Ensemble Modern Francfort, etc.), et ont ramené deux prix Jeune chercheur. Les prototypes de logiciels musicaux issus du projet ont été choisis pour être présentés par M. Thierry Mandons, secrétaire d'état à la recherche lors de ses vœux à la recherche 2017 au Musée du Quai Branly.



Modélisation en temps-réel d'une improvisation

Le projet DYCI2 est un projet de recherche expérimentale coordonné par l'IRCAM. Il associe aussi Inria Nancy et l'université de La Rochelle, ainsi que l'EHESS et le Metamedia Center de

l'EPFL. Le projet a commencé en mars 2015 et a duré 36 mois. Il a bénéficié d'une aide ANR de 500K € pour un coût global de l'ordre de 1,2M €.

B.2 RÉSUMÉ CONSOLIDÉ PUBLIC EN ANGLAIS

The computer as a creative companion to our musical improvisations

DYCI2 targets new musical creation needs and new multimedia interaction scenarios through human-machine improvisation

Improvised interaction between humans and digital artifacts is a recent field resulting from studies on artificial creativity, which brings together several active research issues: interactive learning, whose models are built in the very time of interaction and whose results influence the conditions of this interaction; artificial perception; modelling of social and expressive interaction between human and/or artificial agents, in its anthropological, social, linguistic, and computer dimensions.

The scientific challenges of the DYCI2 project address new needs in musical creation and performance, as well as new scenarios for creative interaction between users and multimedia objects by modelling in all its aspects of perception, learning, generativity and interaction the situation of co-improvisation between human or artificial agents. DYCI2 implements, along with scientific research and technological development of powerful tools for creation, a large number of artistic and educational actions at the international level.

DYCI2 implements methods of artificial listening, symbolic learning, and interaction architectures for improvisation

Musical agents, real (instrumentalists) or virtual (computers), must be able to listen to each other in order to play together satisfactorily in improvised conditions. The computer must therefore be able, through artificial listening, to analyze and structure musical sound according to time, vertically (at a given moment) and horizontally (along the time axis), by methods of discovery, separation and prediction involving learning. From the data flows resulting from this listening stage, it is then necessary to extract symbolic models of musical sequences likely to capture the multidimensional and multi-scale structures specific to music, in particular in a context of improvisation, and constituting richly informed musical memories. From these, the computer determines the improvised interactions between artificial and human agents by examining questions of collective dynamics, adaptation to different time scales, reactivity and planning allowing different kinds of adaptive guidance of improvisation according to listening.

Major project results

Deep signal processing and learning techniques, statistical sequence modelling, and Bayesian probabilistic models were successfully used for the modelling part of the project. Sophisticated architectures for interaction combining free generation, planning and instant reaction to context have been developed and used in many prestigious artistic applications. DYCI2 has given rise to several scientific breakthroughs, including the adaptation of Deep Learning to live multi-channel separation, a world *premiere*.

Scientific production and patents since the beginning of the project

The software for source separation by Deep Learning has been registered with the APP and valorized with Samsung. The results of the project have been disseminated to the scientific, artistic and general public through a large number of publications, workshops and international concerts (College de France, Montreux Jazz Festival, Ensemble Modern Frankfurt, etc.), and have brought back two Young Researcher Awards. The project's musical prototypes have been chosen and shown by Mr. Thierry Mandon, Secretary of State for Research during his 2017 wishes to the research community at the Musée du Quai Branly

The DYCI2 project is an experimental research project coordinated by IRCAM. It also involves Inria Nancy and the University of La Rochelle, as well as CAMS/EHESS and the Metamedia Center at EPFL, sw. The project started in March 2015 and lasted 36 months. It received an ANR grant of €500K for a total cost of around €1.2M.

C MÉMOIRE SCIENTIFIQUE

Mémoire scientifique confidentiel : non

C.1 RÉSUMÉ DU MÉMOIRE

L'interaction improvisée entre des humains et des artefacts numériques est un domaine récent issu des études sur la créativité artificielle, qui convoque plusieurs problématiques de recherche très actives : l'apprentissage interactif, dont les modèles se construisent dans le temps-même de l'interaction et dont les résultats infléchissent les conditions de cette interaction; la perception artificielle, base de cette interaction; la modélisation de l'interaction sociale et expressive entre agents humains et/ou artificiels, dans ses dimensions à la fois anthropologique, sociale, linguistique, et informatique. À partir de ce cadre général de recherche, le projet met en œuvre des situations expérimentales dans lesquelles les agents musicaux, réels (instrumentistes) ou virtuels (ordinateurs) doivent être capables de s'écouter les uns les autres et d'être créatifs afin de jouer ensemble de manière satisfaisante en conditions d'improvisation.

Les défis scientifiques du projet DYCI2 visent les nouveaux besoins en matière de création et de performance musicale, ainsi que les nouveaux scénarios d'interaction créative des utilisateurs avec les objets multimédias en modélisant dans tous ses aspects de perception, d'apprentissage, de générativité et d'interaction la situation de Co-improvisation entre des agents humains ou artificiels. DYCI2 met en œuvre, avec la recherche scientifique et le développement technologique d'outils puissants pour la création, un grand nombre d'actions artistiques et pédagogiques au niveau international.

C.2 ENJEUX ET PROBLÉMATIQUE, ÉTAT DE L'ART

Les défis scientifiques du projet DYCI2 visent les nouveaux besoins en matière de création et de performance musicale, ainsi que les nouveaux scénarios d'interaction créative des utilisateurs avec les objets multimédias.

Dans le cadre du projet, nous souhaitons que les agents musicaux, réels (instrumentistes) ou virtuels (ordinateurs), soient capables de s'écouter les uns les autres et de se comporter de manière créative, afin de jouer ensemble en conditions réalistes d'improvisation. L'ordinateur devait donc être capable **d'analyser et de structurer le son musical en fonction du temps**, verticalement (à un instant donné) et horizontalement (suivant l'axe temporel). Tel était l'enjeu ambitieux de **la tâche WP1**, qui se décomposait en trois sous-tâches : la

séparation informée multicanal, l'écoute structurante et la décomposition / recombinaison par imitation.

La tâche WP2 visait à définir et à **apprendre des modèles symboliques de séquences musicales** susceptibles de capturer les structures multidimensionnelles et multi-échelles émergeant dans un contexte d'improvisation. Au cours du projet, l'état de l'art a évolué avec l'apparition du Deep Learning, que nous avons intégré dans la méthodologie générale.

La tâche **WP3** visait à explorer les conditions dans lesquelles peut se produire **une interaction improvisée riche et créative entre agents artificiels** et humains en examinant les questions de dynamiques collectives, d'adaptation aux différentes échelles temporelles, de structure cognitive interne permettant différents régimes de connaissance, et de guidage adaptatif fonction de l'écoute.

Enfin, **la tâche WP4** avait pour objectif **d'intégrer les nouveaux outils** proposés par les partenaires dans les WP1, 2 et 3, de les mettre en jeu expérimentalement, et de les confronter aux musiciens experts tout au long du projet.

Au début du projet, les travaux existant sur la modélisation du style et l'interaction improvisée s'étaient développés notamment à l'Ircam avec le soutien de l'ANR (projets ImproTech et SOR2). La thèse de F. Maniatakos soutenue en 2012, « Graphs and Automata for the Control of Interaction in Computer Music Improvisation », aborde les aspects théoriques des modèles stylistiques de séquences. La thèse de B. Lévy (soutenance prévue 2014), « Principles and Architectures for an Interactive Music Improvisation System with On the Fly Listening, Learning and Generative capabilities » explorait les architectures d'interaction. Le logiciel OMax issu de ces travaux a servi de plate-forme expérimentale à ces études et constitué une référence reconnue internationalement comme l'a montré le Workshop ImproTech Paris New-York 2012 co-organisé à New-York par l'Ircam, l'EHESS, New York University et Columbia University (<http://repmus.ircam.fr/improtechpny>). Les projets ANR INEDIT sur l'unification des paradigmes de programmation synchrone et asynchrone et EFFICACE (JCJC) sur l'extension réactive des langages pour la composition, sans viser spécifiquement la question de l'improvisation, ont fourni de nouveaux outils conceptuels en contribuant à créer un contexte favorable.

C.3 APPROCHE SCIENTIFIQUE ET TECHNIQUE

La séparation informée multicanal (sous-tâche WP1.1) a été traitée par Inria en début de projet (voir partie résultats), avec deux approches (factorisation matricielle positive et apprentissage profond). Suite à la mutation de Sylvain Marchand, responsable de ce WP1, de l'Université de Bretagne Occidentale à l'Université de La Rochelle, il a fallu trouver un moyen d'explorer les deux autres sous-tâches (WP1.2 et WP1.3) dans un temps réduit (démarrage après la mi-parcours) et avec des moyens adaptés (la thèse initialement prévue n'étant plus possible dans le temps restant). Il se trouve que la méthode Shazam, qui a fait ses preuves en matière d'identification musicale (reconnaissance exacte d'un morceau parmi d'autres), peut aussi servir à retrouver les entités (sources) qui composent un mélange musical (mix). La méthode initiale est très robuste au bruit et à la réverbération. De plus, cette méthode peut être modifiée pour devenir résistante à la transposition ainsi qu'à l'étirement temporel (voir la thèse de Sébastien Fenet). Ainsi, nous avons une méthode existante dont il nous fallait montrer la capacité à faire de l'écoute structurante et dont nous étions convaincus qu'elle avait la capacité d'aller vers l'imitation (résistance à la réverbération, à la transposition, à l'étirement temporel, voire passage au niveau symbolique avec seulement la fondamentale dans le cas des sources harmoniques), répondant ainsi

simultanément aux problématiques des sous-tâches WP1.2 et WP1.3. Par ailleurs, les recherches en Deep Learning de l'équipe RepMus à l'Ircam (master et début de thèse de Tristan Carsaut) ont abouti au développement d'un prototype intégré de module d'écoute et d'un prototype de module de prédiction prenant en entrée un flux audio segmenté par pulsation et retournant respectivement un label d'accord correspondant à la dernière pulsation obtenue, et une progression d'accords inférée pour les mesures à venir. Ces méthodes ont permis d'enrichir notablement les résultats de la tâche WP1.2 portant sur l'écoute structurante et prédictive du WP1.

Dans le WP2, Nous avons adopté un nouveau paradigme consistant à intégrer les modèles de séquences formels, qui permettent de mémoriser des séquences unidimensionnelles de durée arbitrairement longue, et les modèles probabilistes, qui permettent de représenter des dépendances entre dimensions et de générer des séquences non observées par le passé. Ces travaux principalement portés par la thèse de Ken Deguernel soutenue en mars 2018 et la première moitié de celle de Nathan Libermann ont apporté plusieurs contributions nouvelles dans le domaine de l'interpolation de sous-modèles probabilistes, dans celui des algorithmes de propagation de croyance dans les graphes de clusters Bayésiens, ainsi que dans les applications musicales de l'apprentissage profond à l'aide de réseaux neuronaux.

Les différentes stratégies de génération d'un agent improvisateur embarquant une « mémoire » musicale ont été étudiées et développées en parallèle au sein des différentes sous-tâches du WP3, principalement porté par la fin de thèse (soutenance mars 2016) puis le post-doctorat de Jérôme Nika, puis assemblées dans l'intégration décrite par le rapport livrable L3.3.2. Nous rappelons ici les différents paradigmes pouvant être mobilisés par un agent DYCI2 au sein d'une même performance : de la génération libre à la génération structurée par un scénario temporel, en passant par la réaction aux événements extérieurs analysés par un module d'écoute.

L'approche dans le WP4 était de développer simultanément une librairie de processus génératifs pouvant être interfacés avec d'autres environnements, et une librairie d'agents musicaux pouvant générer des séquences multimédias, de manière autonome ou guidée, en faisant appel à la librairie de processus génératifs. Le travail de recherche et développement porté principalement par J. Nika s'est donc articulé autour de deux librairies : une librairie Python implémentant les processus génératifs (modèles de mémoire et traitement des requêtes) et architectures réactives dans le domaine symbolique, et une librairie d'objets Max pour la performance proposant des interfaces pour le choix des paramètres d'apprentissage, des modalités d'interaction (continuum de paradigme allant de la réactivité à la planification), et des modules de restitution pour jouer les séquences générées dynamiquement. Les versions successives de cette plateforme ont été testées et étudiées sous l'angle des cas d'usage, et validée avec les utilisateurs finaux, c'est-à-dire les improvisateurs humains, à travers un nombre significatif de workshops, de résidences et de productions artistiques.

C.4 RÉSULTATS OBTENUS

Dans la tâche **WP1**, deux méthodes ont été proposées pour la séparation informée multicanal d'enregistrements musicaux (livrable L1.1). La première méthode permet de réduire la « repaïse » dans les enregistrements live, en exploitant les informations a priori concernant la position des microphones par rapport aux instruments. Une première version de cette méthode est basée sur le Kernel Additive Modeling (KAM), puis une seconde version améliorée se base elle sur la factorisation matricielle positive (NMF) fondée sur un modèle

probabiliste, où a été introduite une technique de projection aléatoire des données, qui permet de réduire fortement le coût calculatoire de la NMF et de l'appliquer à des concerts entiers avec un grand nombre de canaux, ou bien en temps réel. La première méthode a permis à l'équipe d'obtenir un **1^{er} prix aequo des campagnes d'évaluation SiSEC 2015 et 2016 pour la séparation d'enregistrements musicaux**. La seconde méthode est basée sur l'apprentissage profond ou Deep Learning. **Nous sommes parmi les premiers au niveau mondial à avoir adapté le Deep Learning à la séparation d'enregistrements musicaux et les premiers à l'avoir fait dans un contexte multicanal** avec un impact très important. Lors de la campagne d'évaluation SiSEC 2018, Sony (dont la méthode reprend certains éléments de celle proposée), a obtenu une qualité de séparation excellente, supérieure de 4 dB par rapport à l'année précédente, en apprenant simplement le réseau de neurones sur un plus grand nombre d'enregistrements. La tâche de SiSEC (séparer les voix, la basse, la batterie et les autres instruments) est maintenant considérée comme résolue sur le plan scientifique.

Pour ce qui est de l'écoute structurante pour la décomposition / recombinaison par imitation (WP1.2 et WP1.3), nous avons dû redévelopper en interne la base de la méthode Shazam, puis construire une base de données afin de tester les performances de la méthode en termes de décomposition musicale (et plus précisément retrouver quelle source joue à quel moment ce qui à notre connaissance, ceci n'avait jamais été fait). De plus, nous ne pouvions pas nous appuyer sur une base existante, ni pour la méthode (globalement protégée, mais seule la base nous était nécessaire), ni pour les données (il nous fallait des morceaux de musique, avec également la structure musicale et les entités sonores les composant, or à notre connaissance rien de ce genre n'était publiquement disponible). Nous sommes parvenus à vérifier la faisabilité d'une décomposition en temps réel par cette méthode, même s'il reste beaucoup de travail pour aboutir à une solution logicielle exploitable par des musiciens (livrable L1.3). Par ailleurs les recherches menées à l'Ircam portant sur l'écoute structurante et prédictive à l'aide de Deep Learning (CNN et RNN) (L1.2) ont abouti au développement d'un prototype intégré de module d'écoute et d'un prototype de module de prédiction prenant en entrée un flux audio segmenté par pulsation et retournant respectivement un label d'accord correspondant à la dernière pulsation obtenue, et à une progression d'accords inférée pour les mesures à venir.

La tâche **WP2** a produit plusieurs résultats nouveaux dans l'apprentissage de séquences musicales symboliques :

- l'introduction du nouveau paradigme d'improvisation combinant un oracle des facteurs unidimensionnel représentant un contexte musical local et un modèle probabiliste représentant le savoir musical acquis sur un corpus plus large capturant les structures musicales multidimensionnelles. (livrable L2.1.1),
- l'extension de ce paradigme à la génération d'improvisations multidimensionnelles, grâce à un système multi-agent modélisant l'interactivité entre plusieurs musiciens ou dimensions, où les communications entre les agents sont réalisées par un algorithme probabiliste de propagation de croyance dans un cadre Bayésien (L2.1.1, L2.1.2),
- une grammaire syntagmatique modélisant la structure multi-échelle d'un morceau et un algorithme permettant d'utiliser une telle grammaire pour diriger le processus génératif (L2.3),
- l'exploration de nouvelles façons de modéliser des séquences musicales par réseaux de neurones récurrents et l'usage des réseaux de neurones profonds et de l'apprentissage par renforcement afin de détecter les « fausses notes » dans l'improvisation en cours et de modifier l'apprentissage pour en tenir compte (L2.2).

La tâche **WP3** a produit plusieurs modèles de génération par navigation dans une « mémoire musicale » qu'un agent DYCI2 peut adopter alternativement au sein d'une même performance : de la génération libre à la génération structurée par un scénario temporel en passant par la réaction aux événements extérieurs analysés par un module d'écoute (voir L3.3.2) :

- Guidage par scénario long-terme : nous avons introduit un modèle « scénario / mémoire » exploité en temps-réel par des algorithmes modélisant des mécanismes anticipatifs, et utilisant un corpus appris en temps réel. Pour atteindre cet objectif, le modèle proposé associe à chaque instant de la génération l'anticipation en assurant la continuité avec le futur du scénario, et la cohérence avec la logique musicale de la mémoire en assurant la continuité avec le passé de la mémoire (voir ImproteK, livrable L3.1.1).
- Guidage par réaction à un flux d'entrée : la génération est guidée par un profil d'activité : une fonction de score continue représentant la pertinence de chaque état de la mémoire musicale déterminée par une écoute multimodale de l'environnement musical. Ce profil d'activité évolue avec le temps et se propage avec un comportement d'oubli introduisant un mécanisme de rémanence cognitive (voir Somax, L3.1.2).
- Guidage par scénarios dynamiques à court terme : la génération est pilotée par des requêtes dynamiques ou « scénarios à court-terme », introduisant ainsi une notion de « meta DJing ». Ce paradigme, conçu dans L3.1.1 et développé dans L3.3.2, voit chaque réaction du système comme une modification de ses « intentions ». Des « anticipations » prenant en compte le nouveau contexte sont ainsi ré-écrites et remplacent les anticipations précédemment générées alors que le système est en train de jouer. Cette architecture d'agent réactif introduit la gestion des requêtes concurrentes, des accès concurrents aux données, et d'une trace d'exécution pour assurer la cohérence entre deux requêtes successives.
- Guidage par scénario de descripteurs audio : les stratégies de guidage « par scénario » ou « par scénarios dynamiques à court terme » peuvent également être utilisées pour naviguer dans une mémoire constituée par l'analyse automatique d'un fichier ou d'un flux audio. Les classes constituant l'alphabet de labels sont obtenues par une analyse du fichier « mémoire » selon une sélection de descripteurs effectuée par l'utilisateur, puis par un clustering discrétisant l'espace de descripteurs en un nombre de classes également choisi par l'utilisateur. La navigation dans une mémoire ainsi constituée par analyse automatique peut également être pilotée par un module d'écoute/analyse temps-réel. Un flux audio capté en temps réel – par exemple un musicien co-improvisant avec l'agent – est analysé selon la même sélection de descripteurs que la mémoire pour créer automatiquement des requêtes de génération (voir L3.3.2).
- Intégration et guidages hybrides : nous avons enfin conçu une architecture d'agent unifiant les paradigmes libre, réactif, et guidé par scénario basée sur les mécanismes de planification modélisant une réaction comme une réécriture d'anticipations préalablement générées, et donnant lieu à de nouveaux paradigmes décrits en discussion (voir L3.3.2).

Ces travaux ont valu à Jérôme Nika deux prix Jeune chercheur : le Prix Jeune Chercheur AFIM (Association Française d'Informatique Musicale), 2016 et le Prix Jeune Chercheur Science/Musique, AFIM, INRIA, IRISA, Univ. Rennes, 2015.

L'intégration logicielle (**WP4**) s'est basée sur la version finale des travaux menés dans le WP3 (L3.3.2) rassemblant les différents paradigmes pouvant être mobilisés par un agent DYCI2 au sein d'une même performance : de la génération libre à la génération structurée par un

scénario temporel, en passant par la réaction aux événements extérieurs analysés par un module d'écoute. La suite logicielle (livrable L4.2) est composée de :

- La librairie Python DYCI2 implémentant les modèles d'apprentissage, les stratégies de navigations dans une mémoire musicale, et les architectures d'agents génératifs. Elle a été conçue de manière modulaire afin de pouvoir être prise en main par des développeurs souhaitant utiliser les outils qu'elle implémente pour créer leurs propres agents improvisateurs, étendre simplement le panel de stratégies de génération, et intégrer ces outils dans un environnement externe, par exemple en utilisant le protocole OSC (ou l'encapsulation en langage C décrite plus loin).
- La librairie Max DYCI2 propose une interface pour chacune des classes de la librairie Python ainsi que différents modules de rendu audio. Elle communique en arrière-plan avec la librairie Python via le protocole OSC et permet de construire aisément son propre dispositif de performance en associant un ou plusieurs agents, modules de rendu, sources de requêtes, et stratégies de génération. Des tutoriels ont été réalisés à partir de patchs de concerts réalisés pour des créations associées au projet DYCI2, et ces exemples sont accompagnés de matériaux musicaux originaux créés par des artistes partenaires du projet.

Les principes de dualité « intuition / connaissances », développés dans le WP2.1, ont été intégrés dans la librairie DYCI2. Ceux-ci permettent une navigation orientée par une intuition unidimensionnelle représentée par un modèle de séquence, à l'aide de connaissances multidimensionnelles représentée par un ensemble de modèles probabilistes interpolés. D'autre part, nous avons intégré les méthodes d'improvisation sur un scénario multi-niveau développées dans WP2.3. Celle-ci permettent de prendre en compte la structure hiérarchique et l'organisation d'un scénario en plusieurs niveaux temporels (par exemple, sections, fonctions, accords...), en mettant en place des principes de navigation et d'anticipation considérant l'importance de chaque niveau lors du déroulé du scénario.

L'environnement unifié résultant du projet permet un degré de réalisme et de flexibilité de l'improvisation homme-machine rarement atteint auparavant.

C.5 EXPLOITATION DES RÉSULTATS

Les résultats ont été publiés dans les revues et conférences de référence du domaine comme en témoigne la riche rubrique publications du site DYCI2. Les logiciels correspondants sont distribués sous licence libre. Le logiciel dnnsep de séparation de sources par réseaux de neurones a été déposé à l'APP. Une version modifiée de ce logiciel a été vendue à Samsung en novembre 2017 pour un usage lié au traitement de la parole.

Notons que les recherches du projet DYCI2 ont été choisies comme domaine d'application pour le développement de nouvelles fonctionnalités du langage Antescofo (variables temporelle) ainsi que pour le développement du nouveau système d'ordonnancement de l'environnement OpenMusic (Antescofo et OpenMusic sont deux composants importants du projet ANR INEDIT).

La suite logicielle intégrée Librairie DYCI2 (WP4), distribuée sous licence libre, est également accompagnée de l'encapsulation et de l'intégration de technologies issues du projet dans des environnements externes dédiés à la performance ou la composition musicale : Castrat (en collaboration avec Diemo Schwarz et l'équipe ISMM de l'Ircam) et OpenMusic (en collaboration avec Jean Bresson, équipe RepMus de l'Ircam).

Recherches et développements ont été menés en dialogue constant avec des musiciens improvisateurs, en intégrant pleinement les interactions et les échanges au processus itératif

de développement des modèles et des systèmes. Ces collaborations ont été mises en place au travers de résidences et des productions décrites dans les rapports L4.1.1 et L4.1.4. Enfin, les résultats ont été largement diffusés à la communauté scientifique, musicale, et au grand public. DYCI2 a été à l'origine de nombreuses productions musicales ambitieuses (voir livrables L4.1.1 et L4.14), notamment avec le compositeur Georges Bloch qui a développé le composant vidéo pour la recréation des archives, et se distingue donc par son réel impact musical : en effet, une vingtaine de musiciens experts a été impliquée sur le long terme et plus de 50 performances artistiques au plus haut niveau ont mis ces outils en jeu depuis 2015 (Annenberg Center, Philadelphia, États-Unis ; Centre Pompidou ; Collège de France ; Montreux Jazz festival, Suisse, etc.). Marc Chemillier (CAMS, EHESS), partenaire extérieur au projet a mis en place un programme ambitieux de validation des outils logiciels dans le cadre des musiques populaires, avec notamment une tournée à Madagascar et Mayotte ("Augmented jazz tour", Bernard Lubat, 2016) et une série de tournées française avec le guitariste malgache virtuose Charles Kely Zana-Rotsy.

C.6 DISCUSSION

En ce qui concerne le WP1.1, l'introduction du Deep Learning a eu un impact remarquable lors de la campagne d'évaluation SiSEC 2018, comme indiqué plus haut. Des progrès importants restent toutefois à faire pour séparer l'ensemble des instruments d'un concert, avec un nombre d'instruments et de micros élevés et des instruments similaires les uns aux autres. Les autres thèmes du WP1 (WP1.2 et WP1.3) sont exploratoires, mais la méthode Shazam a confirmé son potentiel pour l'écoute structurante et la décomposition par imitation. De premiers résultats prometteurs montrent la capacité de la méthode de base à retrouver les entités composant le morceau (structuration), et ce en temps réel, à partir certes de signaux exacts. Toutefois, l'état de l'art montre que cette méthode peut être adaptée pour des cas approchés (imitation). Une publication est en cours de rédaction avec Simon Fargeot. Il est clair que ces premiers résultats ouvrent la voie à des travaux plus conséquents (thèse à envisager). Aussi, dans l'immédiat, il semblait profitable au projet de rediriger des moyens restant du WP1 vers les WP 3 et 4 centraux pour l'intégration finale (compléments de post-docs de Jérôme Nika et Ken Déguernel).

Les objectifs des sous-tâches WP2.1 et WP2.3 ont été largement réalisés. Ils ont donné naissance à de nouvelles approches de l'improvisation basées sur l'utilisation de connaissances multidimensionnelles ou des connaissances a priori d'une structure multi-niveau, ouvrant la voie à des applications inédites en matière d'apprentissage et de générativité dans l'interaction improvisée. Un verrou important restant à franchir concerne l'identification voire l'apprentissage de la structure multi-échelle d'un morceau au cours de l'improvisation et temps réel.

Les objectifs de la sous-tâche WP2.2 ont été adaptés aux nouveaux enjeux du Deep Learning. Notre objectif initial en 2014 était d'identifier automatiquement les dimensions sémantiques utilisées par les autres musiciens afin d'adapter la génération. L'apparition du Deep Learning et son utilisation pour l'improvisation musicale, rendue populaire par le projet Magenta de Google apparu publiquement en juin 2016, ont bouleversé la façon de poser le problème. Comparé aux autres méthodes d'apprentissage automatique qui nécessitent d'extraire et de modéliser séparément les différentes dimensions, le Deep Learning a la capacité à traiter conjointement toutes les dimensions. Plutôt que d'identifier les dimensions sémantiques, le problème devient de détecter les « fausses notes » dans une séquence multidimensionnelle, sans viser à identifier explicitement la ou les dimensions erronées, afin

d'adapter la génération dans une boucle perception/action. Ce nouvel objectif a été atteint sur le plan méthodologique. Toutefois, comme c'est souvent le cas avec le Deep Learning, la meilleure façon de modéliser des séquences musicales par réseaux de neurones reste à déterminer. Des premiers pas dans cette direction ont été effectués et le travail se poursuit dans la deuxième moitié de la thèse de Nathan Libermann. La combinaison des méthodes par réseaux de neurones et par modèles formels est une perspective importante ouverte par le projet, ainsi que l'usage des logiciels développés par des musiciens professionnels dans un contexte de concert.

Les objectifs du WP3 ont été largement réalisés et ont donné naissance à un cadre unifié combinant les paradigmes de génération « libre », « réactive » et « guidée par un scénario » autrefois disjoints, et permis d'atteindre, combinés aux résultats du WP2, un degré de réalisme et de flexibilité de l'improvisation rarement atteint auparavant.

La fusion des stratégies de génération « réactive » et « guidée par un scénario court-terme » a permis d'introduire de nouvelles modalités d'interaction cyber-humaine dépassant le clivage « autonomie *vs* contrôle »: si l'information déterminée par l'écoute artificielle peut être utilisée de manière autonome, elle peut également être combinée avec une action humaine en permettant un contrôle hybride entre le musicien produisant le stimulus et l'opérateur-musicien « composant » en temps réel les modalités de la réactivité dans une dynamique de compagnonnage humain / humain-machine.

Les objectifs du **WP4** sont atteints et ont donné naissance à une suite logicielle (Librairie DYCI2) combinant notamment les paradigmes de génération « libre », « réactive » et « guidée par un scénario » autrefois disjoints en intégrant des méthodologies d'apprentissage des structures musicales et d'écoute structurante qui n'existaient pas au début du projet. Il est aussi à noter que l'association des stratégies de génération développées dans le WP3 et du module d'écoute prédictive (WP1.2) a introduit un nouvel intermédiaire entre les paradigmes de génération « guidée par un scénario à court-terme » et « guidée par une écoute réactive ». Les progressions d'accords à court terme émises par ce module peuvent constituer des requêtes de « scénario à court-terme » envoyées en temps réel à un agent, lui permettant ainsi de générer en réaction des improvisations anticipatives à partir d'un scénario dynamique inféré et non plus seulement prédéfini. Cette première étape - à l'état de prototype mais aux résultats prometteurs - **ouvre une direction de recherche tout à fait inédite qui n'est pas le moindre résultat de ce projet.**

C.7 CONCLUSIONS

Le projet DYCI2 a créé des conditions nouvelles pour l'interaction créative homme-machine en mettant à profit le paradigme de l'improvisation. C'est la base pour un nouveau type de systèmes, dit *cyber-humain*, dans lesquels une certaine continuité peut s'établir de l'univers numérique à la réalité des pratiques humaines. Un nouveau chemin reste à défricher pour prolonger ce continuum vers une interaction cyber-physique, de manière à fermer la boucle d'une véritable réalité musicale augmentée.

C.8 RÉFÉRENCES

Ken Déguernel, Emmanuel Vincent, Gérard Assayag. Probabilistic Factor Oracles for Multidimensional Machine Improvisation. *Computer Music Journal*, MIT Press, 2018, 42 (2), pp.52-66.

Jérôme Nika, Marc Chemillier, Gérard Assayag. *ImproteK: introducing scenarios into human-computer music improvisation*. *ACM Computers in Entertainment*, 2017.

Aditya Arie Nugraha, Antoine Liutkus, Emmanuel Vincent. Multichannel audio source separation with deep neural networks. IEEE/ACM Transactions on Audio, Speech, and Language Processing, IEEE, 2016, 24 (10), pp.1652-1664

Derry Fitzgerald, Antoine Liutkus, Roland Badeau. Projection-based demixing of spatial audio. IEEE Transactions on Audio, Speech and Language Processing, Institute of Electrical and Electronics Engineers, 2016

G rard Assayag. Improvising in Creative Symbolic Interaction. Jordan B. L. Smith; Elaine Chew; G rard Assayag. Mathematical Conversations : Mathematics and Computation in Music Performance and Composition, World Scientific; Imperial College Press, pp.61 - 74, 2016

Marc Chemillier, J r me Nika. «  trangement musical » : les jugements de go t de Bernard Lubat   propos du logiciel d'improvisation ImproteK. Cahiers d'ethnomusicologie, Ateliers d'ethnomusicologie, 2015, Le go t musical, 28, pp.61-80.

D LISTE DES LIVRABLES

Date livr.	N�	Titre	Nature	Partenaires (responsable)	Commentaires
M03	L0.1	Mise en place du Wiki du projet	administrati on	Ircam	
M12	L0.2	Rapport d'avancement annuel	rapport	Ircam, Inria, ULR	
M24	L0.3	Rapport d'avancement annuel	rapport	Ircam, Inria, ULR	
M43	L0.4	Rapport final	rapport	Ircam, Inria, ULR	
M18	L1.1	S�paration inform�e multi-canal (rapport + maquette)	rapport+pro totype	Inria, ULR	re-planifi�e
M42	L1.2	�coute structurante (rapport + maquette)	rapport prototype	Ircam, ULR	re-planifi�e et red�finie (transfert UBO - ULR)
M42	L1.3	D�composition / recombinaison par imitation (rapport + maquette)	rapport+pro totype	<u>ULR</u>	re-planifi�e (transfert UBO - ULR)
M12	L2.1.1	Apprentissage de structures multidimensionnelles, algorithme v1 (rapport + maquette)	rapport+pro totype	Ircam, <u>Inria</u>	
M24	L2.1.2	Apprentissage de structures multidimensionnelles, vers. finale algorithme (rapport + maquette)	rapport+pro totype	Ircam, <u>Inria</u>	
M24	L2.2.1	S�lection de dimensions et apprentissage par renforcement, algorithme v1 (rapport + maquette)	rapport+pro totype	Ircam, <u>Inria</u>	
M36	L2.2.2	S�lection de dimensions et apprentissage par renforcement, vers. finale algorithme (rapport + maquette)	rapport+pro totype	Ircam, <u>Inria</u>	
M36	L2.3	Apprentissage de structures multi-�chelles, algorithme (rapport + maquette)	rapport+pro totype	Ircam, <u>Inria</u>	

M12	L3.1.1	Guidage de l'interaction improvisée I par scénario (rapport et maquette)	rapport+prototype	Ircam , Inria	
M42	L3.1.2	Guidage de l'interaction improvisée II par flux d'entrée (rapport et maquette)	rapport+prototype	Ircam , Inria	re-planifiée
M18	L3.2.1	Adaptation temporelle de l'interaction I (rapport et maquette)	rapport+prototype	Ircam , Inria	
M30	L3.2.2	Adaptation temporelle de l'interaction II (rapport et maquette)	rapport+prototype	Ircam , Inria	
M24	L3.3.1	Mémoire, connaissance et contrôle des agents créatifs I (rapport et maquette)	rapport+prototype	Ircam , Inria	
M42	L3.3.2	Mémoire, connaissance et contrôle des agents créatifs II (rapport et maquette)	rapport+prototype	Ircam , Inria	re-planifiée
M12	L4.1.1	Expérimentations WP1,2,3 musiciens en studio (rapport)	rapport	Ircam , Inria	
M30	L4.1.2	Colloque International Improvisation et nouvelles technologies (Colloque)	rapport	Ircam , Inria, ULR	re-planifiée Improtech 2017
M40	L4.1.3	Track/Workshop dans une conférence internationale (Workshop)	rapport	Ircam , Inria, ULR	re-planifiée, SMC2018
M24	L4.1.4	Artistes en résidence (Rapport)	rapport	Ircam , Inria	
M30	L4.1.5	Écoute et interaction improvisée Archives Montreux Jazz Festival (rapport et démonstrateur)	rapport + prototype	Ircam , Inria	Devenu projet artistique "Meute"
M42	L4.2	Intégration suite logicielle des outils de WP1,2 et 3 (logiciel)	rapport + logiciel	Ircam , Inria, ULR	Fin de projet

E IMPACT DU PROJET

E.1 INDICATEURS D'IMPACT

Nombre de publications et de communications (à détailler en E.2)

		Publications multipartenaires	Publications monopartenaires
International	Reuves à comité de lecture	4	4
	Ouvrages ou chapitres d'ouvrage		3
	Communications (conférence)	7	8
France	Reuves à comité de lecture		
	Ouvrages ou chapitres d'ouvrage		
	Communications (conférence)	1	1
Actions de diffusion	Articles vulgarisation	1	
	Conférences vulgarisation	25	
	Autres Évènements publics artistiques, scientifiques (recensés sur dyci2.ircam.fr rubrique 'évents')	60	

Autres valorisations scientifiques (à détailler en E.3)

	Nombre, années et commentaires (valorisations avérées ou probables)
Brevets internationaux obtenus	
Brevets internationaux en cours d'obtention	
Brevets nationaux obtenus	
Brevets nationaux en cours d'obtention	
Licences d'exploitation (obtention / cession)	Cession logiciel DNNSEP, SAMSUNG
Créations d'entreprises ou essaimage	Partenariat industriel en construction avec HyVibe, startup issue de l'Ircam.
Nouveaux projets collaboratifs	Soumis : CREST 2018 ; MERCI 2019 ; MIRE 2019.
Colloques scientifiques	ImproTech 2017 Philadelphia, SMC 2018 Cyprus, ImproTech 2019 Athens
Autres (préciser)	

E.2 LISTE DES PUBLICATIONS ET COMMUNICATIONS

Journal articles

- Dimitri Bouche, Jérôme Nika, Alex Chechile, Jean Bresson. Computer-aided Composition of Musical Processes. *Journal of New Music Research*, Taylor & Francis (Routledge), 2017, 46 (1), <hal-01370792>
- Marc Chemillier, Jérôme Nika. « Étrangement musical » : les jugements de goût de Bernard Lubat à propos du logiciel d'improvisation ImproteK. *Cahiers d'ethnomusicologie, Ateliers d'ethnomusicologie*, 2015, Le goût musical, 28, pp.61-80. <hal-01262592>
- Ken Déguernel, Emmanuel Vincent, Jérôme Nika, Gérard Assayag, Kamel Smaïli, Learning of hierarchical temporal structures for guided improvisation, accepté par *Computer Music Journal*, Massachusetts Institute of Technology Press (MIT Press), 2018
- Ken Déguernel, Emmanuel Vincent, Gérard Assayag. Probabilistic Factor Oracles for Multidimensional Machine Improvisation. *Computer Music Journal*, Massachusetts Institute of Technology Press (MIT Press), 2018, 42 (2), pp.52-66. <hal-01693750v2>
- Derry Fitzgerald, Antoine Liutkus, Roland Badeau. Projection-based demixing of spatial audio. *IEEE Transactions on Audio, Speech and Language Processing*, Institute of Electrical and Electronics Engineers, 2016. <hal-01260588v2>
- Antoine Liutkus, Emmanuel Vincent. Démixer la musique. *Interstices*, INRIA, 2016, <https://interstices.info/jcms/p_84088/demixer-la-musique. <hal-01350450>
- Jérôme Nika, Marc Chemillier, Gérard Assayag. ImproteK: introducing scenarios into human-computer music improvisation. *ACM Computers in Entertainment*, 2017. <hal-01380163>
- Aditya Arie Nugraha, Antoine Liutkus, Emmanuel Vincent. Multichannel audio source separation with deep neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, IEEE, 2016, 24 (10), pp.1652-1664. <hal-01163369v5>

Conference papers

- Gérard Assayag. Creative Symbolic Interaction: Key Note Talk.  A. Georgaki and G. Kouroupetroglou (Eds). 40th Intl. Comp. Mus. Conf. and 11th Sound and Music Comp. Conf.

- (ICMC / SMC joint conf.), Sep 2014, Athenes, Greece. ICMA, SMC, National and Kapodistrian University of Athens, IRMA, pp 1-6, 2014. [〈hal-01146734〉](#)
- Diego Carlo, Antoine Liutkus, Ken Déguernel. Interference reduction on full-length live recordings. ICASSP 2018 - IEEE International Conference on Acoustics, Speech, and Signal Processing, Apr 2018, Calgary, Canada. IEEE, pp.736-740. [〈hal-01713889〉](#)
- Tristan Carsault, Jérôme Nika, Philippe Esling. Using musical relationships between chord labels in automatic chord extraction tasks. International Society for Music Information Retrieval Conference (ISMIR 2018), Sep 2018, Paris, France. 2018, [〈hal-01875784〉](#)
- Ken Déguernel, Jérôme Nika, Emmanuel Vincent, Gérard Assayag. Generating Equivalent Chord Progressions to Enrich Guided Improvisation: Application to Rhythm Changes. SMC 2017 - 14th Sound and Music Computing Conference, Jul 2017, Espoo, Finland. pp.8. [〈hal-01528559〉](#)
- Ken Déguernel, Emmanuel Vincent, Gérard Assayag. Using Multidimensional Sequences for Improvisation In The OMax Paradigm. 13th Sound and Music Computing Conference, Aug 2016, Hamburg, Germany. [〈hal-01346797〉](#)
- Diego Di Carlo, Ken Déguernel, Antoine Liutkus. Gaussian framework for interference reduction in live recordings. AES International Conference on Semantic Audio, Jun 2017, Erlangen, Germany. 2017. [〈hal-01515971〉](#)
- Nathan Libermann, Frédéric Bimbot, Emmanuel Vincent. Exploration de dépendances structurelles mélodiques par réseaux de neurones récurrents. JIM 2018 - Journées d'Informatique Musicale, May 2018, Amiens, France. pp.81-86, 2018, [〈http://www.algomus.fr/jim2018〉](http://www.algomus.fr/jim2018) . [〈hal-01791381〉](#)
- Sylvain Marchand. Spatial Manipulation of Musical Sound: Informed Source Separation and Respatialization. Vienna Talk (VITA) on Music Acoustics, Sep 2015, Vienne, Austria. [〈hal-01373021〉](#)
- Sylvain Marchand. Audio Scene Transformation Using Informed Source Separation. 5th Joint Meeting of the Acoustical Society of America / Acoustical Society of Japan, Nov 2016, Honolulu, United States. [〈hal-01731107〉](#)
- Jérôme Nika, Ken Déguernel, Axel Chemla--Romeu-Santos, Emmanuel Vincent, Gérard Assayag. DYCI2 agents: merging the "free", "reactive", and "scenario-based" music generation paradigms. International Computer Music Conference, Oct 2017, Shangai, China. 2017. [〈hal-01583089〉](#)
- Jérôme Nika, Marc Chemillier, Gérard Assayag. Guider l'improvisation musicale homme-machine : une synthèse sur le système ImproteK. Journées d'Informatiques Musicales (JIM) 2016, Mar 2016, Albi, France. [〈hal-01361233〉](#)
- Jérôme Nika, Dimitri Bouche, Jean Bresson, Marc Chemillier, Gérard Assayag. Guided improvisation as dynamic calls to an offline model. Sound and Music Computing (SMC), Jul 2015, Maynooth, Ireland. [〈hal-01184642〉](#)
- Aditya Arie Nugraha, Antoine Liutkus, Emmanuel Vincent. Multichannel Music Separation with Deep Neural Networks. European Signal Processing Conference (EUSIPCO), Aug 2016, Budapest, Hungary. pp.1748-1752, Proceedings of the 24th European Signal Processing Conference (EUSIPCO). [〈hal-01334614v2〉](#)
- Thomas Prätzlich, Rachel Bittner, Antoine Liutkus, Meinard Müller. Kernel additive modeling for interference reduction in multi-channel music recordings. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr 2015, Brisbane, Australia. 2015. [〈hal-01116686v2〉](#)
- Kevin Sanlaville, Gérard Assayag, Frédéric Bevilacqua, Catherine Pelachaud. Modèles Probabilistes pour l'Interaction entre agents. Workshop Affect • Compagnon Artificiel • Interaction (WACAI 2016), Jun 2016, Brest, France. 2016. [〈hal-01379649〉](#)

Kevin Sanlaville, Gérard Assayag, Frédéric Bevilacqua, Catherine Pelachaud. Emergence of synchrony in an Adaptive Interaction Model. Intelligent Virtual Agents 2015 Doctoral Consortium, Aug 2015, Delft, Netherlands. IVA Doctoral Consortium, 2015. <hal-01164604>

Book sections

Gérard Assayag. Improvising in Creative Symbolic Interaction. Jordan B. L. Smith; Elaine Chew; Gérard Assayag. Mathematical Conversations: Mathematics and Computation in Music Performance and Composition, World Scientific; Imperial College Press, pp.61 - 74, 2016, Lecture Notes Series, 978-981-3140-09-7. <hal-01378904>

Georges Bloch, Jérôme Nika. Edith Piaf, Billie Holiday, and Elisabeth Schwarzkopf Making Music Together. Kapoula Z., Volle E., Renoult J., Andreatta M. Exploring Transdisciplinarity in Art and Sciences, Springer, pp.275-299, 2018, <hal-01875772>

Directions of work or proceedings

Jordan Smith, Elaine Chew, Gérard Assayag. Mathemusal Conversations: Mathematics and Computation in Music Performance and Composition. Jordan B. L. Smith; Elaine Chew; Gérard Assayag. Mathemusal Conversation, Feb 2015, Singapour, Singapore. 32, World Scientific; Imperial College Press, 2016, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore. <hal-01378897>

Master thesis

Rémi Decelle. Apprentissage par renforcement pour l'improvisation musicale automatique. Intelligence artificielle [cs.AI]. 2017. <hal-01591521>

Ken Deguernel. Apprentissage de structures multidimensionnelles pour l'improvisation musicale. Informatique et langage [cs.CL]. 2015. <hal-01267408>

Diego Di Carlo. Gaussian Framework for Interference Reduction in Live Recordings. Signal and Image processing. 2017. <hal-01870918>

This Bazin. Deep Learning for musical scenario inference and prediction: Application to structured co-improvisation. [Internship report] Ircam UMR STMS 9912. 2016. <hal-01388645>

Tristan Carsault. Automatic chord extraction and musical structure prediction through semi-supervised learning, application to human-computer improvisation. [Research Report] Ircam UMR STMS 9912. 2017. <hal-01877337>

Axel Chmela--Romeu-Santos. Guidages de l'improvisation. [Stage] Ircam UMR STMS 9912. 2015. <hal-01388641>

Other publications

Ken Déguernel, Nathan Libermann, Emmanuel Vincent. La musique comme une langue. Commission française pour l'enseignement des mathématiques livret "Mathématiques et langage. 2017. <hal-01485209>

Poster communications

Ken Déguernel, Emmanuel Vincent, Gérard Assayag. Improvisation musicale multidimensionnelle dans le paradigme OMax. Journées Jeunes Chercheurs en Acoustique, Audition et Signal, Nov. 2016, Paris, France. 2016. <hal-01437490>

Thèses

Ken Déguernel. Apprentissage de structures musicales en contexte d'improvisation. Intelligence artificielle [cs.AI]. Université de Lorraine, 2018. Français. 〈NNT : 2018LORR0011〉 . 〈tel-01735308〉

Jérôme Nika. Guiding human-computer music improvisation: introducing authoring and control with temporal scenarios. Sound [cs.SD]. UPMC - Université Paris 6 Pierre et Marie Curie, 2016. English. 〈tel-01361835〉

Invited Talks

Artefacts sonores, cycle « le son des autres », Centre Georges Pompidou, Invited Talk, Gérard Assayag ; Marc Chemillier ; Bernard Lubat, 18/06/2018

Forum Vertigo, Centre Pompidou, Débat « Code et IA : de nouveaux maîtres pour la création ? », invited talk by Gérard Assayag, 13 juin 2018.

Invited Workshop, Athens University, Music Dot, 28 – 30 June, 2018. « DYC2 tools for Jazz students », Gérard Assayag and Jérôme Nika.

E.coli des Mines, Paris. « Improvisation with Creative Agents », invited talk by Gerard Assayag, September 20th, 2018

Geidai University of the arts, Tokyo, « Symbolic Interaction with Creative Agents », invited talk by Gerard Assayag, March 29th, 2018

Today Tokyo University « Symbolic Interaction with Creative Agents », invited talk by Gerard Assayag, April 25th 2018

Interview "AI and music" in "L'esprit sorcier" : Jérôme Nika and Philippe Esling were interviewed by L'esprit sorcier ("C'est pas sorcier 2.0"), a French programme aimed at the general public, for their issue about artificial intelligence and music. Juin 2018.

2018 edition of the Darmstadt Summer Course. Jérôme Nika and Rémi Fox at Darmstadt Summer Course (Germany) on July 27 2018.

Science Festival of Université de Lorraine 2018. October 12 2018 in Nancy, Emmanuel Vincent.

Music Composition and Creative Interaction with Machine Learning Workshop, SMC'18, Cyprus University of Technology, Limassol, Jul. 2018. Jerome Nika (Université de La Rochelle / IRCAM) : Symbolic human-machine musical interaction. Jean Bresson (IRCAM): Machine learning in the compositional workflow (applications in OpenMusic)

Invited Talk at University of Edinburgh. Seminar at the Reid School of Music (30/04/2018), University of Edinburgh, College of Art, United Kingdom. Ken Déguernel (Inria Nancy - Ircam / STMS): "Learning of Musical Structures for Machine Improvisation Systems".

Welcome event new CNRS-INS2I researchers 2018. Presentation by Jérôme Nika "Generative processes for human-computer music co-improvisation" at the welcome event for the new CNRS-INS2I researchers 2018, Ircam, March 19 2018.

Seminar at Pompeu Fabra University (08/02/2018), Barcelona, Spain. Invited Talk by Emmanuel Vincent (Inria Nancy): "Multidimensional and multi-level learning of music structure for machine improvisation in the DYCI2 project".

Workshop Improtech Paris-Philly 2017. University of Pennsylvania. Invited Talks : "The Djazz project: Jazz machines and Anthropology" - Marc Chemillier (EHESS). "Merging free, reactive, and scenario-based features in human-computer co-improvisation" - Jérôme Nika (Ircam / STMS). "Multi-dimensional and multi-scale learning of music structure for machine improvisation" - Ken Déguernel (Inria Nancy - Ircam / STMS)

MOOC Digital Media.Paris, Mairie de Paris, EPSAA, MCD, « Recherche musicale et intelligence digitale », cycle de conférences invitées en ligne de Gérard Assayag.

Colloque Abraham Moles et l'École de Strasbourg, Université de Strasbourg, Septembre 2017, Invited talk by Gérard Assayag « Si proche, si lointain ». Performance DYCI2, compositeur Georges Bloch, Pianiste Hervé Sellin (« La meute kitsch »).

Seminar at Queen Mary University of London, Center for Digital Music (28/06/2017). Invited Talk by Ken Déguernel (Inria Nancy - Ircam / STMS): "Learning of Musical Structures in the Context of Improvisation"

Collège de France "Créativité musicale artificielle », 27 mai 2016, dans le colloque "Art et sciences, de nouveaux domaines pour l'informatique » de la chaire d'informatique de Gérard Berry. Conférence invitée de Gérard Assayag et de Jérôme Nika.

Workshop Ircam @ Boston, Boston University, 25-29 avril 2016, présentations de G. Assayag, M. Noisternig, J. Bresson (et des orateurs invités de MIT et de Rensselaer Institute)

France Inter : Radio program "Chercheurs d'avenir", 2016, presented by Mathieu Vidard, first broadcasted on France Inter, July 10, 2016 ; Invited speakers : Gérard Berry, Stéphanie Delaune, and Jérôme Nika (Ircam / STMS).

Creative Dynamics of Improvised Musical Interaction, Mathemusical Conversations. Mathematics and Computation in Music Performance and Composition, International Workshop, NUS, Singapour, 13-15 février 2015

Annual seminar of the National Network of Complex Systems (RNCS). Invited talk by Jean Bresson (Ircam / STMS) and Jérôme Nika (Ircam / STMS), Le Havre, November 2-5, 2015.

Young Researcher Prize in Science and Music 2015. Science / Music day - Young Researcher Prize in Science and Music 2015, Rennes, Maison des associations, September 25, 2015. Invited talk by Jérôme Nika (Ircam / STMS).

MIR and artificial creativity, Lorentz Workshop on Music Similarity, Concepts, Cognition and Computation, 19-23 janvier 2015, Lorentz Center, Leiden. Invited Talk by Gérard Assayag « Improvised Interaction, a Creative Perspective on Similarity »

Montreux Jazz Festival International Workshop : Le musicien et la machine. Invited talks by Gérard Assayag, Jérôme Nika, Rémi Fox (Ircam – STMS). Vended 17 Juillet 2015, Petit Palais.

Science & You 2015, 3 to 6 juin 2015 in Nancy, Talks by K. Déguernel and E. Vincent (Inria Nancy).

Seminar "Complex Systems in Social Sciences", Invited talk by Jérôme Nika (Ircam / EHESS), EHESS, Paris, April 3, 2015. "Modèle pour l'improvisation musicale homme-machine guidée par un scénario"

Brunch Ethnomusika, Péniche Anako, quai de Seine, Paris, February 8, 2015, invited talk "Marovany zither, rhythm, variations, and computer modeling using the software ImproteK" Concert-conference, presentation by M. Chemillier (EHESS) and demonstration / performance by M. Chemillier (EHESS) and J. Nika (Ircam / STMS).

"Unexpected workshops": traveling school, workshops, discussions, performances, Lille (France), Mons / Tournai (Belgium), 2015. "Mons capitale de la culture 2015". Talks by M. Chemillier (EHESS) and J. Nika (Ircam / STMS)

NYU @ Abu Dhabi Rhythm International Workshop 2014. Cross-disciplinary and Multi-Cultural Perspectives on Musical Rhythm II, New York University @ Abu Dhabi, October 12-15, 2014. Invited talks by Gérard Assayag (Ircam / STMS) , Jérôme Nika (Ircam / STMS), Marc Chemillier (EHESS)

E.3 LISTE DES ÉLÉMENTS DE VALORISATION

Le logiciel dnnsep de séparation de sources par réseaux de neurones a été déposé à l'APP (dépôt IDDN FR 001 480003 000 S P 2016 000 10000 du 18/11/2016). Une version modifiée de ce logiciel a été vendue à Samsung en novembre 2017 pour un usage lié au traitement de la parole. Il est à noter que lors de la campagne d'évaluation SiSEC 2018, Sony (dont la méthode reprend certains éléments de celle proposée), a obtenu une qualité de séparation excellente, contribuant ainsi à sa popularisation.

Des partenariats ont été noués avec l'EPFL à travers le MetaMedia Center, le Center for Research in Entertainment and Learning (CREL) de l'université de San Diego, de grandes

universités américaines (U. of Pennsylvania New York University, Princeton, California Institute for the Arts CalArts), des acteurs institutionnels (Région Nouvelle Aquitaine, Collegium Musicae Sorbonne Université, Compagnie Lubat) et des entreprises (HyVibe, Samsung, Yamaha).

Les résultats du projet ont été généralement diffusés à la communauté scientifique, musicale, et au grand public. En effet, DYCI2 a été à l'origine de nombreuses productions musicales ambitieuses (voir livrables L4.1.1 et L4.14), et se distingue donc par son réel impact musical : en effet, une vingtaine de musiciens experts a été impliquée sur le long terme et plus de 50 performances artistiques au plus haut niveau ont mis ces outils en jeu depuis 2015 (Annenberg Center, Philadelphia, États-Unis ; Centre Pompidou ; Collège de France ; Montreux Jazz festival, Suisse, etc.).

Notons que la suite logicielle DYCI2 a été choisie par le compositeur Pascal Dusapin pour son projet évolutif « Lullaby experience » (<https://www.lullaby-experience.eu>) qui sera créé avec l'Ensemble Modern à Francfort en février 2019 et dont de nombreuses reprises sont d'ores et déjà en préparation. De même, le festival-workshop ImproTech, émanation du projet, a été sélectionné pour financement par la fondation Onassis pour se tenir à Athènes de nouveau en 2019 après l'édition américaine à Philadelphie en 2017.

Le projet a également donné lieu à l'organisation d'un workshop « Music composition and Créative Interaction with Machine learning » dans la conférence internationale Sound and Music Computing (L4.1.3) ainsi que d'un festival-colloque « Improtech Paris-Philadelphia 2017 » rassemblant sur 3 jours à Philadelphie plus d'une soixantaine de musiciens et de chercheurs (L4.1.2).

Enfin, l'impact et la dissémination se traduisent également par des prix et distinctions, et des actes de dissémination parmi lesquels on peut citer :

- Travaux choisis et présentés par M. Thierry Mandons, **Secrétaire d'état à la Recherche** lors de ses vœux à la Recherche 2017 au Musée du Quai Branly
- **Prix Jeune Chercheur** AFIM (Association Française d'Informatique Musicale), 2016
- **Prix Jeune Chercheur** Science/Musique, AFIM, INRIA, IRISA, Univ. Rennes, 2015
- A.A. Nugraha, A. Liutkus & E. Vincent, **1er ex aequo des campagnes d'évaluation SiSEC** 2015 et 2016 pour la séparation d'enregistrements musicaux (travaux autour du WP1).
- Émission "Intelligence artificielle", L'esprit sorcier, avril 2018
- Festival-colloque "Improtech", Philadelphie, États-Unis, décembre 2017, avec Ircam, Université de Pennsylvanie, Collegium Musicae SU, région Aquitaine.
- Émission "Chercheurs d'avenir", France Inter, avec Gérard Berry, 10 juillet 2016
- "Créativité artificielle", présentation invitée et concert, **Collège de France**, 27 mai 2016
- "Musician and machine", Workshop au **Montreux Jazz festival**, Montreux, Suisse, 17 juillet 2015
- Articles *Libération* hors-série Voyage au cœur de l'IA, « La créativité devient un objet d'étude en soi », interview de Gérard Assayag par Mathieu Vidard / France Inter, 2 janvier 2019.
- Article *Le Monde*, Pixels / chronique des révolutions numériques, 2 février 2018, « Effet de mode ou révolution ? Comment l'intelligence artificielle s'invite dans la musique, interview de Gérard Assayag par Nicolas Six.
- MOOC Digital Media.Paris, Mairie de Paris, EPSAA, MCD, « Recherche musicale et intelligence digitale », cycle de 6 conférences en ligne de Gérard Assayag, décembre 2017.

Note :

Le site du projet *dyci2.ircam.fr* contient :

- La liste exhaustive des publications
- La liste des événements (dissémination, spectacle vivant, enseignement, promotion, workshops, conférences invitées etc., radio-presse-tv, etc.) qui ont jalonné le projet (près d'une centaine)
- Un ensemble de ressources (collections de données, démos, tests, prototypes logiciels) librement utilisables
- Un ensemble de présentations (meetings et présentations à l'ANR)
- Une section Private contenant :
 - L'ensemble des livrables du projet sous forme de rapports
 - Les répertoires sources contenant le code des logiciels du projet

E.4 BILAN ET SUIVI DES PERSONNELS RECRUTÉS EN CDD (HORS STAGIAIRES)

Identification				Avant le recrutement sur le projet			Recrutement sur le projet				Après le projet				
Nom et prénom	Sexe H/F	Adresse email (1)	Date des dernières nouvelles	Dernier diplôme obtenu au moment du recrutement	Lieu d'études (France, UE, hors UE)	Expérience prof. Antérieure, y compris post-docs (ans)	Partenaire ayant embauché la personne	Poste dans le projet (2)	Durée missions (mois) (3)	Date de fin de mission sur le projet	Devenir professionnel (4)	Type d'employeur (5)	Type d'emploi (6)	Lien au projet ANR (7)	Valorisation expérience (8)
Nika Jérôme	H	Jerome.Nika@ircam.fr	Sept 2018	Master	France		IRCAM	thèse	6,5	15/4/2016	Post-doc	Autre privé	doctorant	oui	Oui
Nika Jérôme	H	Jerome.Nika@ircam.fr	Sept 2018	Doctorat	France		IRCAM	Post-doc	18	30/01/2018	Post-doc	Autre privé	Post-doctorant	oui	Oui
Bigo Louis	H	Louis.Bigo@ircam.fr	Sept 201	Doctorat	France		IRCAM	Post-doc	1	31/5/2016		Autre privé	Post-doctorant	oui	Oui
Fargeot Simon	H	simon.fargeot@gmail.com		Master	France		ULR	ingénieur d'études	11	30/09/2018	doctorant	recherche publique	ingénieur	oui	
Tamani Nouredine	H	nouredine.tamani@univ-lr.fr		doctorat	France	post-doc 1 an, ATER 2 ans	ULR	ingénieur de recherche	9	30/09/2018		recherche publique	ingénieur	oui	
Nika Jérôme	H	Jerome.Nika@ircam.fr		doctorat	France	Post-doc 1,5 ans	ULR	post-doc	7	30/09/2018		recherche publique	chercheur	oui	Oui
Déguernel Ken	H	ken.deguernel@ircam.fr		doctorat	France		ULR	post-doc	5	30/09/2018		recherche publique	chercheur	oui	
Ken Déguernel	H	ken.deguernel@ircam.fr	31/08/2018	Master 2	France	0	Inria	Doctorant	36	28/02/2018		recherche publique	chercheur	oui	
Nathan Liberman	H	nathan.libermann@inria.fr	31/08/2018	Master 2	France	0	Inria	Doctorant	21	15/07/2018	Doctorat en cours (3 ^e année)	recherche publique	doctorant	oui	