



GRAND PLAN D'INVESTISSEMENT

FONDS POUR LA TRANSFORMATION DE L'ACTION PUBLIQUE

Contrat de transformation

« e-Contacts Plus : assistant virtuel »

Ministère de l'Action et des Comptes publics

Direction générale des finances publiques



Ce contrat de transformation est conclu entre la direction générale des finances publiques (DGFIP) et le Secrétariat général des ministères économiques et financiers, d'une part, et la direction interministérielle de la transformation publique (DITP) et la direction du budget, d'autre part. Il définit les modalités d'exécution du projet, qui conditionnent le versement des crédits au titre du fonds pour la transformation de l'action publique. Il engage également le porteur de projet sur des résultats mesurables.

1. Présentation du projet de transformation

Avec le prélèvement à la source, la messagerie sécurisée mise à disposition des usagers particuliers dans leur compte en ligne sur le site impots.gouv.fr devient un canal structurant d'accès aux services de la DGFIP. Ce canal d'échange en ligne va encore se développer avec l'ouverture récente de cet outil de messagerie aux usagers professionnels en avril 2019. Le déploiement du contact à distance et donc du traitement des courriels est également un facteur de réussite du projet relatif au nouveau réseau à distance de la DGFIP. Dans ce contexte, la mise en place de fonctionnalités d'intelligence artificielle (IA) constitue une opportunité d'améliorer la productivité et la qualité de service (délai et qualité des réponses) de ce canal, tant vis-à-vis des usagers particuliers que des entreprises.

Au plan quantitatif, l'utilisation de l'intelligence artificielle va permettre à la fois de diminuer le flux des messages à traiter manuellement par les agents des structures d'accueil des particuliers et des professionnels, grâce à la mise en place d'un chatbot (destiné à permettre l'auto-assistance des usagers) et de gagner en productivité dans le traitement des messages par les agents, en particulier grâce à l'analyse sémantique.

1- Contexte et objectifs

Depuis 2011, la DGFIP met à disposition des usagers particuliers un outil de messagerie sécurisée, accessible dans le compte authentifié de l'utilisateur sur impots.gouv.fr. Fort de cette expérience et de l'usage croissant que les usagers en ont fait, cet outil a été profondément rénové en janvier 2016 puis complété en juillet 2017 pour en faire une messagerie complète (échange sans rupture des messages entrants et sortants entre l'utilisateur et son service local). Actuellement, cet outil permet une pré-classification des messages, permettant un traitement optimisé par les agents, voire, dans certains cas, un traitement automatisé des demandes (pour les changements d'adresse par exemple). Il est également flexible et permet un traitement des courriels selon une organisation territoriale adaptée aux temps forts des campagnes fiscales notamment (déport d'un service local à un autre, transfert vers les plateformes nationales de relations à distance par exemple).

Au fil des années, ce canal s'est donc très fortement développé ; il véhicule régulièrement 60 000 demandes d'utilisateurs par jour. Avec la mise en œuvre du prélèvement à la source au 1^{er} janvier 2019 et l'ouverture de cette messagerie aux entreprises en avril 2019, son usage a vocation à s'amplifier fortement au cours de ces prochaines années.

Si cet outil de contact apporte une souplesse appréciable pour la DGFIP (canal asynchrone, répartition de la charge sur différents services distants, pré-classification des demandes), la charge de traitement des messages par les agents est particulièrement importante alors que la technicité des sollicitations est très variable. L'usage d'une technologie d'intelligence artificielle est une opportunité majeure pour améliorer la productivité et la qualité de service (délai et qualité des réponses) de ce canal. Au plan qualitatif, la production de réponses adéquates permettra également d'améliorer la qualité de service aux usagers, aussi bien particuliers que professionnels.

À terme, la fréquence de l'utilisation des outils et la typologie des motifs de contact faciliteront la réingénierie des motifs de contact et le renforcement des outils de prévention de ces contacts. À compter de 2021, cet outil de messagerie devrait pouvoir être ouvert au bénéfice des ordonnateurs locaux correspondant au projet de porter au compte personnel des usagers les factures du secteur public local.

2-Dispositif et choix des outils

La messagerie sécurisée propose à l'utilisateur plusieurs types de formulaires selon son besoin. Parmi les demandes adressées, les formulaires « autres questions » représentent couramment plus de 50 % des demandes entrantes, dont certaines sont liées à des formulaires disponibles dans l'arborescence et, de fait, non utilisés.

Dans ce contexte, le dispositif envisagé pour la mise en œuvre de l'IA doit permettre de répondre à deux enjeux, au moyen de deux outils différents :

- lorsque l'utilisateur se connectera, un assistant conversationnel (Chatbot) l'orientera vers les bons formulaires afin de réaliser un meilleur routage des demandes vers les services compétents ;
- une fois que l'utilisateur aura formulé sa demande, un outil d'analyse sémantique permettra d'analyser les demandes de type « autres questions », et de proposer des réponses-types aux agents chargés de rédiger la réponse à l'utilisateur.

Pour le volet Chatbot (assistant conversationnel), il est retenu un outil de gestion des squelettes conversationnels et d'analyse des échanges (Sopra Stéria Botify Kit) utilisant des services cognitifs tiers (appels IBM Watson). Cette solution, déjà mise en œuvre par l'Agence Informatique Financière de l'État (AIFE), permet d'envisager une généralisation du chatbot à des cas d'usages plus étendus (portails, assistance autre).

Pour l'analyse sémantique des demandes e-Contacts, il est mis en œuvre une solution hybride associant une analyse déterministe et un dispositif de machine learning (stack Python et data science).

- l'approche déterministe s'appuie sur des graphes développés dans le cadre d'un POC permettant d'associer les mots clés d'une demande avec une réponse (association d'une réponse type à une intention). La suite Expert System COGITO a été retenue compte tenu notamment de sa connaissance de la langue française ;
- l'approche « machine learning » propose des réponses types en s'appuyant sur des modèles statistiques et logiques élaborés par des data scientists (développement de modèles d'analyse sémantique écrits en langage Python).

Ces deux approches (analyse déterministe et machine learning) sont systématiquement mises en œuvre, pour chaque demande, par un orchestrateur (CGI Smartcase) ; un dispositif de scoring de la performance des deux modèles permet à l'orchestrateur de proposer les meilleures réponses.

Cette association permet de proposer des réponses standardisées dans les meilleurs délais et facilite la gestion du flux de demandes dans l'application e-Contacts :

- l'approche déterministe réduit le risque de biais humain au début du projet, et constitue un accélérateur lorsque la machine learning n'a pas pu suffisamment entraîner les modèles d'analyse sémantique (nouveaux cas d'usage) ;
- le machine learning produit sa pleine efficacité à mesure que les modèles, élaborés par les data scientists, sont entraînés avec une boucle d'apprentissage (étiquetage et rétroactions des utilisateurs). Le taux de précision des prédictions gagne alors en efficacité.

2. Besoins et modalités de financement du projet

Le projet « e-Contacts Plus : assistant virtuel » est un grand projet réalisé dans des délais très courts. Il présente un coût global de 6,401 M€ avec 5,727 M€ de coûts directs (dont assistance, matériel, logiciel et réalisation du POC initial) et 0,675 M€ de coûts indirects.

Financement du projet :

La contribution de la DGFIP est de 0,798 M€. Le projet est co-financé à compter de 2018 par le Fonds de transformation ministérielle (FTM) à hauteur de 1,925 M€. Ainsi, la contribution demandée au Fonds pour la transformation de l'action publique (FTAP) s'élève à 3,678 M€.

Détail des dépenses financées par le Fonds :

L'apport du FTAP permet de financer le volet du projet correspondant à une partie des investissements dans l'infrastructure informatique et des prestations d'assistance aux bureaux MOA et MOE, afin de construire la solution.

Récapitulatif (montants inscrits en k€) :

Nature de dépenses	Catégorie de dépenses	Programme budgétaire	Action/ligne budgétaire porteuse de la dépense	2018		2019		2020		2021		Cumul	
				AE	CP	AE	CP	AE	CP	AE	CP	AE	CP
e-Contacts Plus – Agents DGFIP	Titre2			127,4	127,4	270,4	270,4	276,8	276,8	0,0	0,0	674,6	674,6
dont financement DGFIP		P156		127,4	127,4	270,4	270,4	276,8	276,8			674,6	674,6
dont financement FTM		P218										0,0	0,0
dont financement FTAP		P349										0,0	0,0
Assistance MOA et MOE	Titre5			0,0	0,0	1 804,0	762,0	2 063,0	1 867,2	0,0	1 237,8	3 867,0	3 867,0
dont financement DGFIP		P156										0,0	0,0
dont financement FTM		P218				762,0	762,0					762,0	762,0
dont financement FTAP		P349				1 042,0		2 063,0	1 867,2		1 237,8	3 105,0	3 105,0
Logiciel pour analyse sémantique et ChatBot + Matériel	Titre5			0,0	0,0	663,0	663,0	573,0	573,0	0,0	0,0	1 236,0	1 236,0
dont financement DGFIP		P156										0,0	0,0
dont financement FTM		P218				663,0	663,0					663,0	663,0
dont financement FTAP		P349						573,0	573,0			573,0	573,0
Dépenses pour le POC (AMOA et AMOE)	Titre5			623,6	338,9	0,0	284,7	0,0	0,0	0,0	0,0	623,6	623,6
dont financement DGFIP		P156		123,6			123,6					123,6	123,6
dont financement FTM		P218		500,0	338,9		161,1					500,0	500,0
dont financement FTAP		P349										0,0	0,0
TOTAL				751,0	466,3	2 737,4	1 980,1	2 912,8	2 717,0	0,0	1 237,8	6 401,2	6 401,2
TOTAL Financement DGFIP		P156		251,0	127,4	270,4	270,4	276,8	276,8	0,0	0,0	798,2	798,2
TOTAL Financement FTM		P218		500,0	338,9	1 425,0	1 586,1	0,0	0,0	0,0	0,0	1 925,0	1 925,0
TOTAL Financement FTAP		P349		0,0	0,0	1 042,0	0,0	2 636,0	2 440,2	0,0	1 237,8	3 678,0	3 678,0

3. Économies prévisionnelles engendrées par le projet

Les économies prévisionnelles sont permises par le traitement automatisé des messages et les réponses instantanées proposées lors du Chat.

	Catégorie A	Catégorie B	Catégorie C
Coût moyen par emploi (en €)	42 835	35 069	30 317

	2020	2021	2022	2023
Emplois supprimés	0	138	0	0
dont catégorie A		14		
dont catégorie B		55		
dont catégorie C		69		

	2020	2021	2022	2023
dont redéployés				
dont catégorie A				
dont catégorie B				
dont catégorie C				

	2020	2021	2022	2023
Economie en €	0	2 309 402	4 618 805	4 618 805
dont catégorie A	0	295 562	591 123	591 123
dont catégorie B	0	967 904	1 935 809	1 935 809
dont catégorie C	0	1 045 937	2 091 873	2 091 873

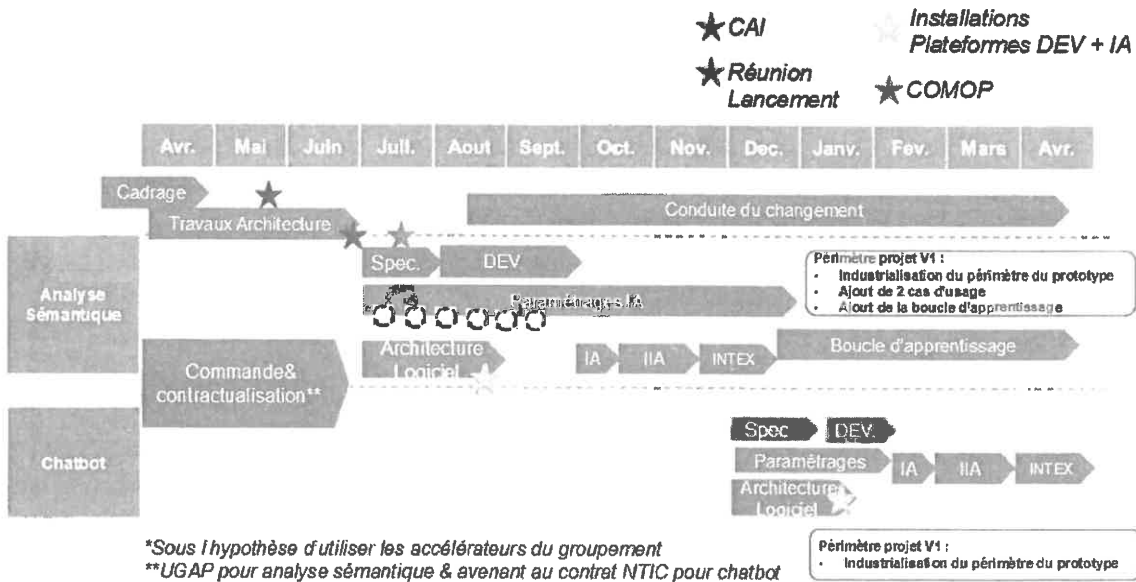
Les économies de coûts de fonctionnement (HT2) sont calculées sur une base de 2 490 € par agent.

Nature de l'économie	Catégorie de dépenses	Programme budgétaire	2020		2021		2022		2023		Économies annuelles pérennes post-déploiement du projet	
			AE	CP	AE	CP	AE	CP	AE	CP	AE	CP
Gains de productivité – Allègement des tâches SIE SIP principalement (agents A, B et C)	T2	P156	0	0	2 309 402	2 309 402	4 618 805	4 618 805	4 618 805	4 618 805	4 618 805	4 618 805
Économies sur les coûts de fonctionnement liés à l'agent	HT2	P156	0	0	171 810	171 810	343 620	343 620	343 620	343 620	343 620	343 620
TOTAL			0	0	2 481 212	2 481 212	4 962 425	4 962 425	4 962 425	4 962 425	4 962 425	4 962 425

4. Calendrier, gouvernance et modalités de réalisation des projets

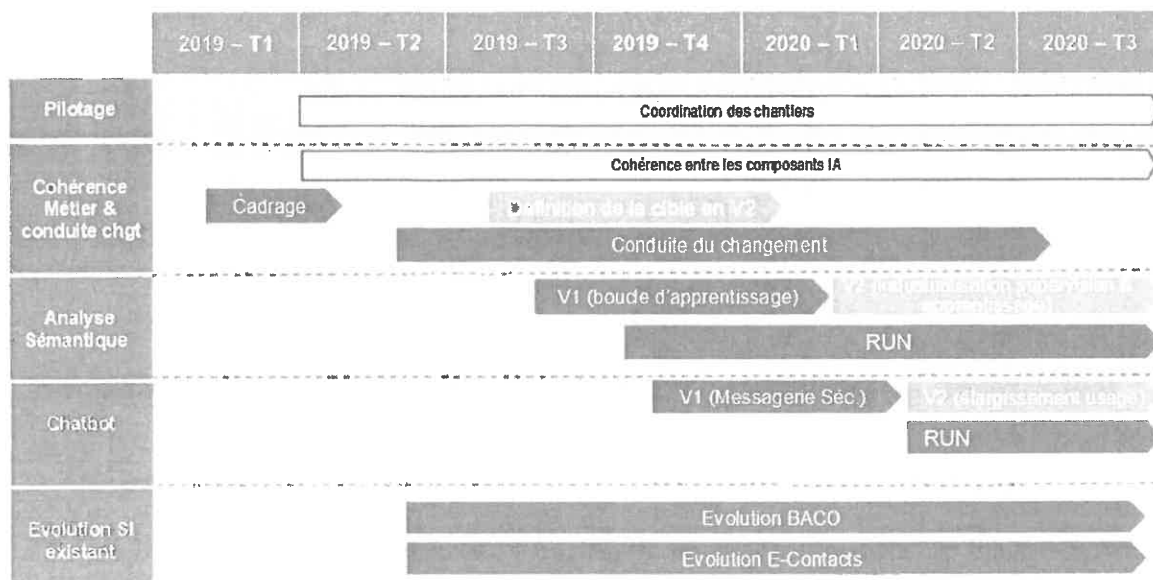
4.1. Calendrier prévisionnel

Après une étude de cadrage conduite en janvier 2019, un chantier de mise en place d'analyse sémantique destiné à optimiser les temps de traitement des agents dans la réponse apportée sur un cas d'usage a été réalisé jusqu'en octobre 2019. Ce pilote sera poursuivi jusqu'à la fin de l'année 2019 pour consolider les enseignements liés au 1^{er} cas d'usage et consolider le paramétrage du dispositif d'IA, en parallèle le projet de chatbot sera lancé. A compter de 2020, le nombre de cas d'usage sera enrichi en vue du déploiement de la solution.



La version 1 de l'analyse sémantique consiste en la réalisation du projet d'analyse sémantique avec la boucle d'apprentissage. Dans la version initiale, un cas d'usage lié à la gestion de la modulation du prélèvement à la source a été réalisé.

En parallèle, les travaux engagés sur une V2 auront pour objectif d'enrichir le périmètre fonctionnel de nouveaux cas d'usage, tant sur le chatbot que sur l'analyse sémantique.



La version 2 de l'analyse sémantique prévoit la mise en œuvre d'un module de supervision de la performance des modèles d'analyse sémantique afin de permettre un apprentissage « on-line », c'est-à-dire sans intervention en back-office, sinon pour sonder les prédictions et vérifier qu'il n'y ait pas de dérive des modèles.

La V2 propose de nouveaux cas d'usages adaptés au besoin des agents, en fonction de l'actualité fiscale.

Le chatbot en état version 1 consistera à développer des squelettes conversationnels permettant d'orienter le parcours usager vers des formulaires adaptés avec un traitement automatisé de l'information.

La V2 proposera la maintenance évolutive et corrective du chatbot.

Un chantier d'internalisation des compétences est également programmé dès la fin 2019 ainsi qu'une prestation de réversibilité programmée en 2022.

4.2. Gouvernance du projet

Une équipe dédiée au sein de la MOA professionnelle de la DGFIP (SCN Cap Numérique) et de la MOE du SSI de la DGFIP a d'ores et déjà été constituée, organisée en instance de pilotage opérationnel. Les équipes internes seront renforcées par une assistance externe le temps de la construction du projet et les 3 premières années suivant le déploiement de la solution (consacrées à l'amélioration du système d'une nature itérative : boucle d'apprentissage, machine learning). Au-delà, la maintenance du projet sera assurée via les ressources internes de la DGFIP.

Le projet sera suivi dans le cadre des instances de pilotage des projets informatiques DGFIP notamment s'agissant des instances stratégiques en comité de pilotage Internet, particuliers et professionnels, présidé chaque trimestre par le Directeur général adjoint des finances publiques. À noter que cette instance permet d'associer tous les chefs de services aux orientations stratégiques de la transformation publique et intègre la Mission Stratégie relations aux publics qui représente depuis 2013 « la voix de l'utilisateur » aux projets majeurs de la DGFIP en relation avec les publics.

En outre, l'avancement du projet sera suivi dans le cadre de 2 instances opérationnelles :

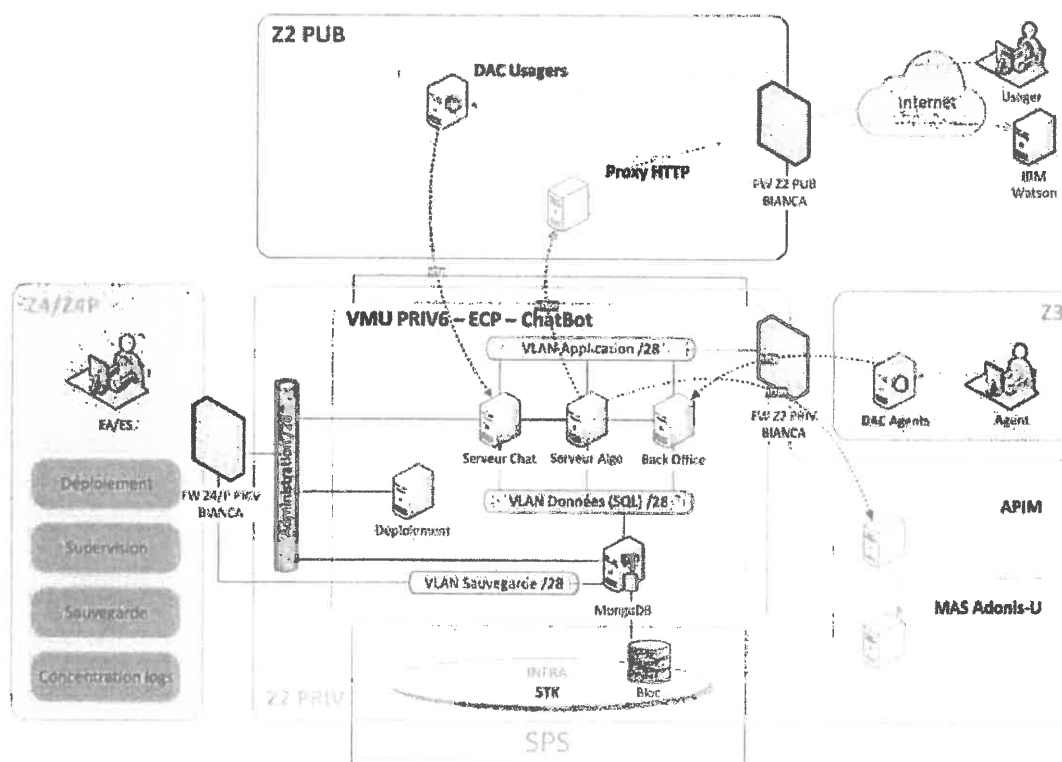
- un comité opérationnel (COMOP) relatif au domaine des applications de la relation usagers. Ce comité est présidé par les chefs des bureaux de MOA et de MOE respectives avec la participation des autres acteurs (dont les bureaux métiers, le pilotage de la production et celui de la recette d'intégration). Il vise à piloter les travaux et l'avancement des projets du domaine, arbitrer ou demander un arbitrage supra ;
- un comité de suivi du projet, présidé par les chefs de projet MOA et MOE. Il vise à suivre l'avancement des travaux, piloter le projet en identifiant les risques et difficultés, et en faisant remonter les demandes d'arbitrages.

4.3. Modalités de réalisation du projet – respect des principes de l'État plateforme

Chatbot :

Le schéma ci-dessous précise l'architecture technique sécurisée qui sera mise en place pour le ChatBot.

Architecture technique – Schéma de synthèse (1/2)



Afin d'accélérer la mise en production du ChatBot, sera utilisé le logiciel BotifyKit de ChatBot avec un paramétrage et une modification limitée pour répondre aux besoins de la DGFIP.

Le socle technologique de la solution ChatBot est basé sur l'utilisation d'openjdk8, une base de données MongoDB et plusieurs librairies open source conformes à la matrice technologique de la DGFIP dont les principales sont :

- ReactJS est le nom du framework visuel de l'application ChatBot
- SuperAgent est l'outil de réalisation des appels API, notamment utilisé pour les appels à IBM/Watson
- Webpack est un framework permettant la modularisation d'une application JavaScript
- Babel est un compilateur JavaScript
- Express est une infrastructure Web minimaliste, souple et rapide pour Node.js
- Mongoose est la partie ORM dans Node.js pour accéder à une base MongoDB

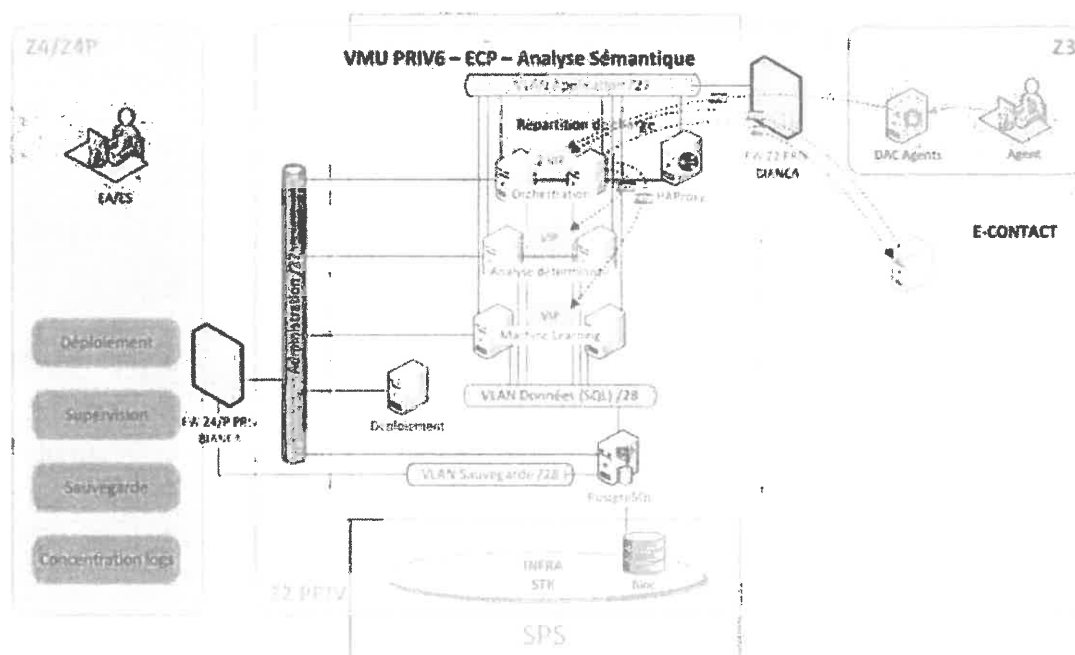
Pour accéder à la console d'administration et au Chatbox, plusieurs navigateurs peuvent être utilisés : Firefox 56, Chrome 66, IE11, Safari 10 et EDGE.

Les recommandations du RGAA (Référentiel général d'accessibilité pour les administrations) seront prises en compte pour le développement des squelettes conversationnels et la charte graphique du ChatBot.

Analyse Sémantique :

Le schéma ci-dessous précise l'architecture technique sécurisée qui sera mise en place pour l'Analyse Sémantique.

Architecture technique – Schéma de synthèse (2/2)



Afin d'accélérer la mise en production de l'Analyse Sémantique, nous utiliserons plusieurs logiciels (Smart Case pour orchestrer le traitement des messages et Cogito Studio et Express) avec le développement de modèles d'apprentissage (Machine Learning) utilisant des librairies Python et le paramétrage de « règles Déterministes » basé sur le logiciel Cogito.

Le socle technologique de la solution Analyse Sémantique est basé sur l'utilisation d'openjdk8, de la version 3.7 de Python, une base de données Postgres (pour le serveur Base de Données d'Apprentissage) et plusieurs librairies open source conformes à la matrice technologique de la DGFIP dont les principales sont :

- Web et Flask : pour gérer les appels REST
- Pandas : construction et lecture des CSV (input de l'exploration)
- Numpy : sur les modèles, traitement vectoriel
- Scikit Learn : sur les modèles, gère le machine learning et de la vectorisation
- Keras : deep learning simple
- NLTK : outils de text processing
- Spacy : outils de text processing, plus forts en FR, utilisé sur l'anonymisation
- PyTorch : Deep learning plus avancé
- Tensor Flow : Deep learning plus avancé
- Gensim : text processing, vectorisation, word 2 vec, doc 2 vec, LDA
- GloVe : espace vectoriel de mots

Afin de permettre une confrontation (et une appropriation) rapide de la solution d'Analyse Sémantique avec les agents qui traitent les messages des usagers, il sera mis en place une expérimentation et un déploiement partiel sur un seul centre de contact.

Le développement et l'accompagnement sont réalisés dans le cadre d'une prestation avec l'attributaire du marché NTIC de la DGFIP, régie par le CCAG conforme aux exigences des projets des administrations de l'État.

5. Modalités de suivi et critères d'évaluation du projet

Le présent contrat donne lieu à un suivi du projet financé. Des indicateurs d'avancement et de résultats sont suivis dans le cadre du financement du projet. Ces indicateurs sont communiqués, à sa demande et au moins une fois par an, au secrétariat du fonds. Des réunions de suivi pourront être organisées à la demande d'une des parties lors de la communication de ces indicateurs.

5.1. Indicateurs d'avancement

Les indicateurs d'avancement permettent de sécuriser la mise en œuvre du projet. Ils sont définis comme suit :

- Montant des crédits consommés en AE et CP et respect des enveloppes allouées à chacune des tranches, par nature de dépense, par rapport aux besoins de financement présentés dans le point 2 du présent contrat ;
- Respect du calendrier prévisionnel de déploiement du projet, par rapport au calendrier présenté dans le point 4.1 du présent contrat ;

5.2. Indicateurs de résultat et d'impact

Les indicateurs de résultat permettent d'évaluer l'atteinte des objectifs du projet :

- Montant d'économies générées (€) et répartition par nature de dépenses. Les économies réalisées seront comparées aux économies prévisionnelles présentées dans le point 3 du présent contrat ;

- **Ratio de messages pré-traités par analyse sémantique** : /questions posées dans la rubrique « autres questions » de la messagerie sécurisée (couverture en augmentation progressive) ;
- **Ratio de satisfaction des utilisateurs du chatbot** (à mettre en place sur le parcours).

Indicateur	Valeur actuelle	Cible fin 2020	Cible fin 2021	Cible fin 2022
Ratio de messages pré-traités par analyse sémantique/ pour la rubrique « autres questions »	0% (en 2019)	10%	25%	40 %
Ratio de satisfaction des utilisateurs du ChatBot	0% (en 2019)	50%	55%	65%

6. Modalités et calendrier de versement des aides

Les crédits sont mis à la disposition **de la secrétaire générale des ministères économiques et financiers**. La secrétaire générale procède aux diligences nécessaires pour permettre l'ordonnancement des crédits du FTAP par les directions concernées.

La mise à disposition des crédits s'effectue par tranche. Le montant de chaque tranche sera définitivement arrêté par le secrétariat du fonds, les dépenses annuelles détaillées dans la présente convention étant prévisionnelles, à l'exception de la première année de financement (2019). À partir de 2020, le secrétariat du fonds décide à échéance régulière, *a minima* au 1^{er} trimestre de chaque année, du montant des nouvelles tranches de financement au regard de l'avancement du projet, du suivi des indicateurs et de l'avis rendu par le DINSIC sur le projet.

S'agissant en particulier des opérations d'investissement (titre 5 majoritaire), le montant des AE nécessaires pour le financement d'une phase fonctionnelle du projet¹ devra faire l'objet d'une affectation au sens de la comptabilité budgétaire, en cohérence avec les phases du projet décrites au 4.1.

Les crédits sont mis à disposition dans le cadre de gestion BOP-UO décrit en annexe. La consommation des crédits (AE et CP) sur le programme 349 est opérée en référant la nomenclature budgétaire d'activités annexée au présent contrat.

7. Matérialisation des économies réalisées

La matérialisation des économies liées au projet est suivie annuellement, conformément aux indicateurs définis au paragraphe 5.2. Le porteur de projet communique au secrétariat du fonds les économies effectivement réalisées et explicite les raisons des éventuels écarts avec les prévisions exposées dans le présent contrat.

¹Au sens de l'article 8 de la LOLF

8. Modifications du contrat de transformation

Le présent contrat peut être modifié par voie d'avenant.

Toute difficulté majeure dans la réalisation du projet sera portée à la connaissance du comité de pilotage qui pourra suspendre ou interrompre les financements initialement définis.

9. Communication liée au projet

Dans toute communication relative à son projet, le porteur est invité à préciser qu'il a reçu le soutien financier du Fonds pour la transformation de l'action publique.

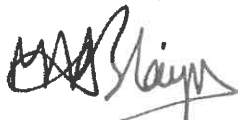
14 JAN. 2020

Le directeur général des finances publiques



Jérôme FOURNEL

La secrétaire générale des ministères économiques et financiers



Marie-Anne BARBAT-LAYANI

Le délégué interministériel à la transformation publique



Thierry LAMBERT

La directrice du budget



Amélie VERDIER

ANNEXE RELATIVE AUX NOMENCLATURES BUDGETAIRES D'EXECUTION

Cadre de gestion BOP-UO : 0349-CDBU-CEFI

Action – Domaine fonctionnel : 0349-01

Référentiel de programmation :

Code Chorus	Désignation Chorus	Commentaires
034901015201	MACP - Econtacts+ AV	Concerne toutes les dépenses HT2 relatives au projet et imputées sur le programme 0349

