Mise en place d'un référentiel d'identités à l'échelle d'un Campus : une approche fédérative

Gautier Auburtin, Johann Holland

Mise en place d'un référentiel d'identités à l'échelle d'un Campus: une approche fédérative

Gautier Auburtin

Campus Condorcet

8 cours des humanités

93322 Aubervilliers CEDEX

Johann Holland

Campus Condorcet

8 cours des humanités

93322 Aubervilliers CEDEX

Résumé

La gestion d'identités fait l'objet depuis plusieurs années d'une préoccupation croissante au sein des établissements d'enseignement supérieur. Longtemps, la gestion des comptes utilisateurs fut limitée aux besoins d'authentification, puis élargie aux problématiques d'autorisations et enfin étendue aux fédérations d'identités. En parallèle, la gestion d'identités s'est consolidée autour de processus et référentiels métiers puis sur la base de référentiels communs. Aujourd'hui pièce maîtresse d'un système d'information dont elle assure la sécurisation, l'homogénéité et la fluidité, elle reste toutefois déterminée par des logiques et frontières institutionnelles, à l'exception des fédérations d'identités.

Le Campus Condorcet, nouveau campus de recherche en SHS, regroupe une centaine d'unités de recherche issues de onze établissements et deux sièges d'établissements. Il constitue un ensemble aux frontières floues au sein duquel les résidents quotidiens comme plus largement les usagers présentent des affiliations croisées. Dans ce contexte, la mise en place d'un référentiel d'identités commun à l'ensemble du Campus rencontre plusieurs difficultés : identifier la source ; déterminer des affiliations précises ; assurer les entrées et sorties du référentiel ; mettre en place des autorisations ; assurer l'accès aux services du Campus ; sécuriser l'authentification.

Le référentiel d'identités du Campus Condorcet cherche à apporter des réponses innovantes à ces enjeux. Il s'appuie sur des processus et des technologies fédératives (référents identités, fédération d'identités), permet à l'usager de choisir son identité source lors de l'inscription, mobilise les référents pour les invitations et les validations (entrées), fournit un référentiel de structures hébergées et d'autorisations et permet l'accès aux services du Campus par un mécanisme de délégation d'authentification auprès de l'annuaire d'origine.

Mots-clefs

IAM, gestion des identités et des accès, fédération d'identités, authentification unique, Campus, solutions libres, standards

1 Introduction

Le Campus Condorcet est un projet immobilier et scientifique visant à créer un nouveau campus de recherche en sciences humaines et sociales au nord de Paris. Dès septembre 2019, une centaine d'unités de recherche et des enseignements de master rejoindront le site d'Aubervilliers (Place du Front populaire), auxquels succéderont d'autres étudiants sur le site de Paris (Porte de La Chapelle), dès 2022. Le projet est porté par l'établissement public Campus Condorcet (EPCC) pour le compte de onze établissements membres (CNRS, EHESS, ENC, EPHE, FMSH, INED, Universités Paris 1 Panthéon-Sorbonne, Sorbonne Nouvelle Paris 3, Paris 8 Vincennes-Saint-Denis, Paris Nanterre et Paris 13).

Le projet de référentiel d'identités du Campus Condorcet a été lancé en septembre 2016, 3 années avant l'ouverture physique du Campus. Déployé en décembre 2018, le référentiel d'identités a permis d'anticiper l'installation des résidents du Campus et de leur fournir des services au travers d'un nouveau portail livré en septembre 2019. Conduit en deux phases successives (preuve de concept et conception théorique sur 18 mois, puis réalisation technique sur 6 mois), ce projet de gestion d'identités a très fortement déterminé la conduite d'urbanisation des systèmes d'information de ce nouveau Campus en sciences humaines et sociales.

À l'heure de l'ouverture du Campus et des premiers usages du SI Campus, il est possible de tirer quelques enseignements sur la pertinence de l'approche fédérative au cœur de ce modèle de gestion d'identités, tant sur le plan technique qu'organisationnel. En effet, si le modèle adopté pouvait sembler pertinent au vu des contraintes et des objectifs connus avant l'ouverture du campus, l'afflux des usagers présentait un risque majeur pour la maîtrise de la plateforme technique, des processus de gestion d'identités et l'adoption finale par les utilisateurs.

Cet article décrit le processus de conception (POC) du modèle fonctionnel de gestion des identités et sa mise en œuvre technique, les caractéristiques fonctionnelles et techniques de la plateforme d'inscription et d'authentification des utilisateurs en lien avec la fédération d'identités RENATER, puis rassemble les derniers retours d'expérience sur les usages plus généraux de cette plateforme de gestion d'identités, par ses gestionnaires et ses utilisateurs finaux.

2 Processus de conception

En septembre 2016, le projet est initié au sein du pôle numérique du Campus Condorcet composé de deux agents (Directeur et Chargé d'urbanisation numérique / Chef de projet). Il associe les DSI des 10 premiers établissements membres, au sein desquelles une équipe projet volontaire est constituée avec les administrateurs d'identités de trois établissements (EPHE, PARIS 1, INED). Cette équipe resserrée sera par la suite élargie à toutes les DSI pour tirer le bilan de la preuve de concept (janvier 2018), puis aux référents fonctionnels (RH, directions de la recherche) pour la mise en œuvre de la plateforme (avril 2018). L'apport du GIP RENATER et de l'équipe Fédération d'identités à ce projet a permis sa réalisation, en apportant l'expertise nécessaire sur la technologie SAML et un soutien régulier dans la mise en œuvre de la plateforme (paramétrages, tests).

2.1 Objectifs et caractéristiques de la population

Le référentiel d'identités (REFID) du Campus est positionné comme le fondement du Système d'information et sa première étape. À la différence d'un établissement de recherche ou d'enseignement porté par des activités et des flux métiers structurants, le SI du Campus se caractérise d'abord par une approche centrée usager (à qui doit-on fournir des services ?). Mais aussi, à la différence d'une réunion d'établissements (COMUE) dont le périmètre est relativement aisé à définir, la population cible du Campus repose sur plusieurs dimensions :

1. l'approche « institutionnelle » des établissements membres (EPST, Universités, Fondation), dont une partie très variable du personnel est concernée par le Campus (de la totalité d'un établissement - l'INED - à une fraction très réduite d'une université ou du CNRS) ;

2. l'approche « factuelle » des unités de recherche résidentes, qui représentent les acteurs principaux du Campus, mais dont la composition et le rattachement sont elles-mêmes complexes – les « résidents » devant occuper les espaces du campus ne regroupent pas la totalité des membres ou associés d'une unité et ces unités sont composées de personnes aux affiliations disparates, parfois employés par des établissements non membres du Campus (IRD ou université Paris Diderot pour ne nommer qu'elles).
3. l'approche « imprévisible » des usagers du Campus, qui au-delà des seuls résidents disposant d'un bureau, réunit les chercheurs invités, les étudiants de master, les futurs lecteurs de la bibliothèque dénommée GED (Grand équipement Documentaire), et enfin les habitants du territoire ou visiteurs d'un jour (le Campus étant « ouvert sur la ville »).

2.2 Méthode de la preuve de concept (POC)

2.2.1 Objectifs fonctionnels et méthode de travail

Le référentiel d'identités est positionné dès le début du projet au centre du SI, comme le décrit la figure suivante présentée en comité de projet POC en septembre 2016.

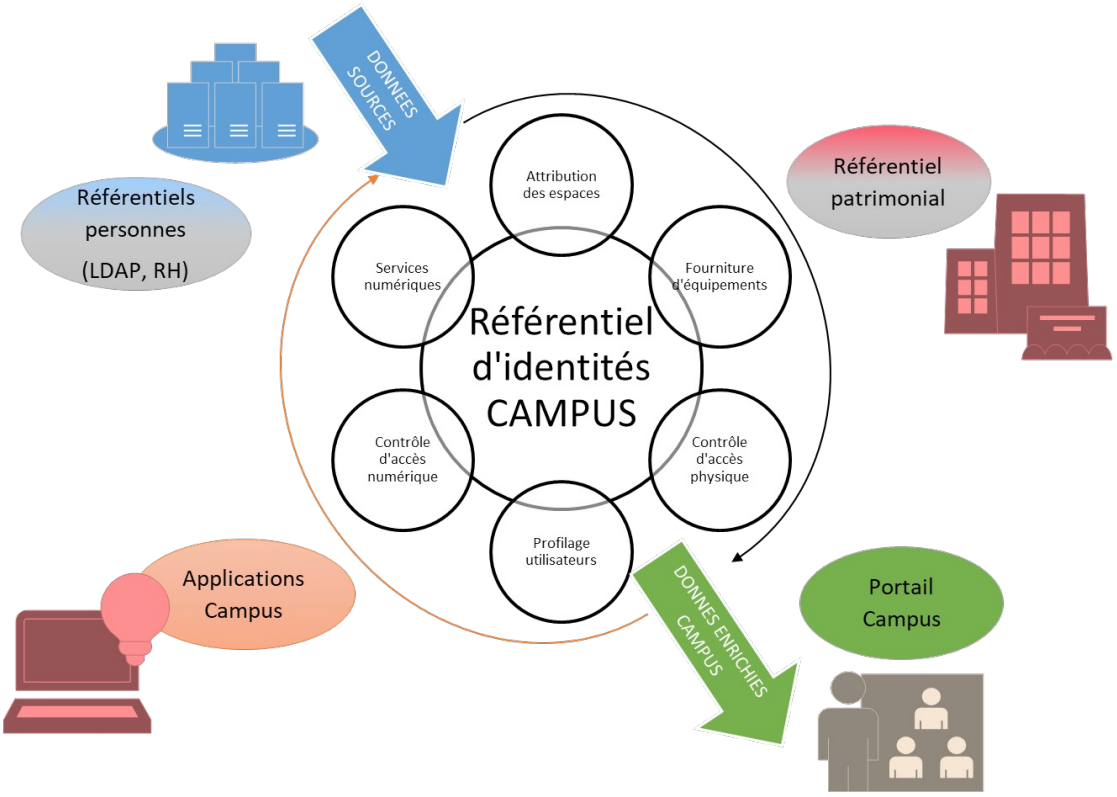


Figure 1 - Modèle fonctionnel du référentiel d'identités

Le référentiel d'identité doit :

1. obtenir ses données des référentiels des établissements (annuaires LDAP, bases RH) ;
2. répondre à des besoins de services dits physiques, en lien avec le référentiel de données patrimoniales (ou immobilières) en construction : attribution physique des espaces (macro et micro-zoning des bâtiments), contrôle d'accès par badge ;

3. permettre l'accès aux futures applications du SI Condorcet à destination des usagers (réservation de salles, demandes de support informatique, demandes d'exploitation-maintenance ou de services, …) ;
4. consolider les informations sur les usagers au travers d'un portail Campus permettant le profilage des individus pour une diffusion d'informations sélective et un enrichissement des profils.

Tirée par l'objectif d'une preuve de concept démontrable (ou non) à court terme (moins d'une année) et par l'accompagnement technique d'un prestataire[1], l'équipe projet du POC a rapidement :

1. défini les attributs indispensables d'un individu, sur la base des schémas et des données dans les établissements (EduPerson et Supann 2009) ;
2. conclu au choix technologique LDAP (OpenLDAP) pour la mise en œuvre de l'annuaire ;
3. proposé une architecture en unités organisationnelles (OU) permettant le cloisonnement des identités par établissement ;
4. proposé deux scénarios d'approvisionnement :
   - un scénario dit « déménagement », fondé sur l'import de données structurées depuis l'établissent d'origine,
   - un scénario dit « invité », dans lequel un individu extérieur à l'établissement pouvait être invité par un référent et s'inscrire de lui-même dans le référentiel.

La prestation demandée comprenait : la construction de l'annuaire LDAP ; l'analyse des données provenant des 3 établissements concernées, afin de valider le schéma cible ; la mise en œuvre d'une plateforme d'inscription.

2.2.2 Premiers constats

L'équipe projet a rapidement fait le constat que le futur Annuaire du Campus ne pouvait être construit à l'instar d'un méta-annuaire rassemblant les données des établissements membres (ex. de la COMUE PSL[2]), pour plusieurs raisons :

1. Il n'était pas possible de circonscrire de manière fiable les populations concernées au sein des annuaires des établissements : la notion d'unité de recherche n'était pas toujours précisément définie, les membres de ces unités n'étaient pas toujours connus, car hors du spectre « RH » de l'établissement.
2. Il fallait envisager un apport régulier des données concernant les usagers depuis les annuaires, ce qui soulevait des difficultés techniques insurmontables et disproportionnées au regard de la faiblesse de la population concernées (2 500 personnes à l'ouverture du Campus).
3. Au regard de la réglementation sur les données personnelles (loi Informatique et libertés et futur RGPD), les formalités à accomplir préalablement au transfert de ces données étaient très dissuasives à la fois pour le Campus Condorcet et pour les établissements membres (conventionnement d'une part, information préalable des personnes concernées de l'autre)[3].

La méthode du POC a aussi démontré son efficacité en confrontant l'équipe projet à des solutions techniques. Plusieurs problèmes sont en effet apparus :

1. Comment contrôler le cycle de vie d'une identité approvisionnée par import initial ? La plateforme devait pour cela permettre le contrôle du statut (actif ou non) de l'identité dans le

[1]Il s'agit de l'entreprise (SCOP) Entrouvert : https://www.entrouvert.com/

[2]Schéma technique de l'annuaire LDAP de la COMUE PSL : http://doc.entrouvert.org/supann/architecture.html

[3]Ces contraintes ont notamment rendu impossible l'accès, par le Campus Condorcet, aux « Services Web des données de référentiels » du CNRS, qui pourtant rassemblait une part importante et actualisée des membres des unités de recherche devant s'installer sur le Campus.

référentiel source, par un mécanisme de test ou synchronisation, ce qui ramenait au premier plan la difficulté technique de l'interopérabilité entre annuaires.

2. Comment contrôler l'inscription des invités et garantir la qualité des données lors d'une saisie manuelle sur un formulaire d'inscription ? Il fallait pour cela mettre en place des mécanismes de contrôle a posteriori, pour lesquels aucun établissement n'avait les moyens humains.

2.2.3 L'approche fédérative

C'est lors de cette expérience du POC que l'approche dite « fédérative » a éclos. Pensée comme un scénario alternatif dédiée aux invités extérieurs, elle s'est imposée comme un scénario unique d'approvisionnement des identités.

Quelques lectures issues de travaux précédents conduits en 2011 à l'Université de Californie[4] ou, plus récemment, sur le projet EduID de la fédération suisse SWITCH[5], laissaient penser que l'approvisionnement des identités sur un Campus et entre établissements était possible sur la base de deux fondements techniques :

1. l'utilisation du protocole SAML v2 pour le provisionnement d'attributs entre un fournisseur d'identités (IdP) et un fournisseur de services (SP) ;
2. l'enregistrement du SP dans une fédération d'identités[6] définissant un cadre de confiance et l'échange des métadonnées entre le SP d'inscription et les IdP concernés.

Par ailleurs, le modèle suisse EduId, centré sur la maîtrise de son identité d'affiliation par l'utilisateur, invitait à donner à l'usager un rôle prépondérant dans la création d'une identité « Campus Condorcet » sur la base de son identité d'affiliation institutionnelle.

Aidé par l'équipe fédération du GIP RENATER[7], la plateforme construite pour le POC a été interfacée avec la fédération locale Campus Condorcet[8] pour valider, au sein des dix établissements membres, un fonctionnement original selon lequel :

1. des référents « établissement » préalablement enregistrés dans la branche (OU) du LDAP Condorcet et disposant de rôles adéquats pouvaient, sur la base d'adresses courriel, inviter des futurs « résidents » à s'inscrire dans l'annuaire du campus ;
2. les invitations comportaient des liens personnalisés (token) vers une plateforme d'inscription se comportant comme un SP[9] et redirigeant l'utilisateur vers un WAYF Condorcet listant les dix IdP des établissements membres ;
3. les invités pouvaient choisir l'IdP de leur établissement, s'authentifier, constater le transfert d'attributs les concernant depuis l'annuaire de leur établissement puis consentir au dit transfert vers le Campus Condorcet ;
4. les référents initiaux, notifiés de l'enregistrement d'une personne, pouvaient valider cette inscription et finaliser la création de l'identité Campus sur la base de l'identité d'origine.

La phase projet du POC a donné satisfaction à plusieurs égards : elle a permis de s'assurer de la fiabilité technique du modèle, voire de remettre en cause le modèle d'approvisionnement initial tout en validant fonctionnellement le modèle de l'invitation. Début 2018, il restait à construire les processus organisationnels avec les établissements et à développer la plateforme prévue, mais aussi à fournir aux utilisateurs une solution plus complète de gestion de comptes et d'authentification, intégrée dans le SI.

---

[4] « UC IDM Provisioning Middleware: Technical Design Guide », University of California, August 2011.

[5] Detailed Architecture for Swiss edu-ID, November 2015, SWITCH.

[6] L'inscription d'un invité devait être initialement possible depuis l'inter-fédération EduGain, afin de permettre l'inscription de chercheurs étrangers depuis l'IdP de leur établissement.

[7] L'auteur remercie ici plus particulièrement Anass Chabli, qui nous a soutenus durant toute la durée du projet POC et au-delà.

[8] Aujourd'hui abandonnée au profit de l'intégration dans la fédération Renater.

[9] Il s'agit en l'occurrence de la plateforme *Authentik* d'Entrouvert : https://authentik.entrouvert.com/

2.3 Conception définitive

La phase de conception définitive du référentiel d'identités a suivi deux processus parallèles : d'une part, la mise en place d'une organisation regroupant des référents « inscriptions » auprès de tous les établissements amenés à transférer du personnel sur le site, que le Campus Condorcet a mobilisés dans le cadre d'une campagne d'inscription démarrée à l'automne 2018 ; d'autre part, le développement d'une plateforme dans le cadre d'un nouveau marché, sur la base des besoins identifiés pendant le POC et élargis à une gestion d'identités adapté à un SI orienté Web (portail de services et applications en mode « extranet »).

2.3.1 Objectifs

La nouvelle plateforme devait permettre l'alimentation du référentiel par des populations croissantes, en une phase principale dite « campagne d'inscription », suivi d'inscriptions « au fil de l'eau » :

Population	**Caractéristiques**	**Nombre fin 2019**	**Nombre en 2022**[10]
Usagers du Campus	Volume, origine et typologie « flous »	4 500	8 000
(dont les) Lecteurs de la bibliothèque	Volume, origine et typologie « flous »	3 000	6 000
(dont les) Résidents du Campus	Volume, origine et typologie « précis »	2 500	4 000

Les processus et formulaires d'inscription devaient pouvoir s'adapter aux différentes populations, avec plus ou moins d'attributs demandés et des mécanismes d'invitation et d'approbation différents entre, par exemple, les résidents (inscription contrôlée) et les lecteurs de la bibliothèque (inscription « libre »).

Techniquement, la plateforme devait être compatible avec les fédérations d'identités (RENATER, Campus Condorcet, Edugain) de type SAML V2 permettant à un utilisateur issu d'un autre établissement de s'authentifier via son annuaire d'origine (fournisseur d'identité) à un service fourni par le Campus (fournisseur de service).

Cela devait permettre :

1. d'assurer les inscriptions au Campus (authentification et remontée d'attributs) ;
2. de déléguer l'authentification à l'établissement d'origine pour tout service du campus ;
3. de lier l'identité locale Campus avec l'identité d'origine, au moyen d'un attribut commun (eduPersonPrincipalName ou EPPN).

La plateforme devait pouvoir, à terme, être compatible avec les protocoles d'authentification de type OAuth / OpenID Connect permettant de s'identifier via un compte réseau social, voire au travers de la fédération France Connect.

Du point de vue du SI cible, la plateforme devait s'inscrire dans le cadre d'un portail de services pour les futurs résidents, lecteurs et usagers du Campus, point d'entrée principal, sur le réseau Internet, pour l'accès aux services numériques du Campus :

1. le partage de fichiers (Drive Seafile) ;
2. la réservation d'espaces (ADE Campus : MyAdeBooking) ;
3. les demandes de support (GLPI) ;
4. les applications documentaires (SGB, SIA, Bib Num).

[10]Ce tableau ne prend pas en compte la population du site de Paris la Chapelle.

Cela passait par la mise en œuvre d'une couche SSO étant capable, outre SAML, de gérer plusieurs protocoles SSO en sortie, entre le portail d'authentification et les applications (notamment CAS).

La plateforme devait aussi permettre une gestion d'identités complète :

1. identifier les utilisateurs (notamment leur origine) ;
2. maîtriser les entrées et les sorties du référentiel ;
3. déléguer la gestion des utilisateurs par branches (ou) à des référents ;
4. construire des profils d'utilisateurs ;
5. définir des règles d'autorisation sur la base de ces profils.

Un enjeu de taille consistait à maîtriser les sorties du référentiel, telles que celles occasionnées par un changement d'affiliation (le résident n'appartient plus à un des établissements) ou par une perte de droits (le lecteur n'est plus inscrit à la bibliothèque au-delà d'un certain délai).

- Dans le premier cas, un mécanisme identifié dans Shibboleth (Back Channels / Attributes Queries)[11] pouvait permettre de tester régulièrement l'existence du compte d'origine et de verrouiller celui-ci tout en avertissant le référent pour qu'il audite le compte[12].
- Dans le second cas, un mécanisme de péremption de comptes était demandé pour qu'au-delà d'une certaine durée initiale, le compte soit verrouillé et que l'utilisateur puisse demander la prolongation de son inscription sous réserve d'une information du référent.

En vue de l'utilisation des services tiers, la plateforme devait permettre de gérer des autorisations sur le système d'information du Campus, sur la base de plusieurs données :

1. le profil de l'utilisateur (attributs) ;
2. l'appartenance de l'utilisateur à des structures ;
3. les rôles de l'utilisateur (groupes de type *OrganizationalRole*).

2.3.2 Processus organisationnels

La plateforme demandée ci-dessus devait être servie à plusieurs types d'utilisateurs :

- les administrateurs (accès complet aux différents modules) ;
- les référents inscriptions (accès aux modules d'invitation et de gestion des utilisateurs) ;
- les invités (accès aux modules d'inscription et d'authentification au SI).

La mise en œuvre des processus organisationnels autour de cette plateforme a nécessité un an de travail (avril 2018 – juin 2019) pour :

1. identifier les référents inscriptions au sein des établissements (DRH, Directions de la recherche et secrétariats d'instituts) ;
2. construire avec eux le processus d'invitation, au sein d'une première campagne d'inscription (supports de communication et tutoriel vidéo, notamment) ;
3. former les référents à l'utilisation de la plateforme et notamment aux formulaires d'invitation ;
4. lancer les inscriptions : démarrées en décembre 2018, celles-ci se sont progressivement étendues aux établissements jusqu'en juin 2019.

La dynamique de la campagne d'inscription a soutenu le processus d'accompagnement au changement des futurs résidents, qui pouvaient se projeter individuellement sur le campus en procédant par eux-mêmes à leur inscription individuelle.

[11] Ces mécanismes sont détaillés par le consortium SWITCH : http://www.swissbib.org/wiki/index.php?title=Switch_Shibboleth_Backchannel_and_Attribute_Query_Plugin https://www.switch.ch/aai/support/presentations/update2016/06_Attribute-Query.pdf

[12] Trop complexe techniquement, ce mécanisme n'a pas été utilisé.

L'approche centrée utilisateur a permis une meilleure prise en compte des obligations d'information et de consentement auprès des personnes concernées : un groupe de travail composé des Correspondants Informatique et Liberté a notamment permis de borner ces obligations et de rédiger des conditions de collecte et d'utilisation des données personnelles.

2.3.3 Mise en œuvre technique

Le marché de réalisation, lancé en avril 2018, a permis d'identifier une solution technique complète de gestion d'identités composée de trois couches techniques :

- un annuaire, OpenLDAP ;
- une application de gestion d'identités, Fusion Directory ;
- une application de gestion de l'authentification et des accès, LemonLdap::NG[13].

Le choix de cette offre reposait sur deux critères principaux :

- elle était la seule à proposer une application complète pour la gestion des comptes (FusionDirectory), laquelle manquait dans le POC ;
- elle offrait une solution SSO complète multi-protocoles (SAML, CAS, OpenId Connect, Proxy HTTP).

Toutefois, des développements complémentaires ont été nécessaires :

- sur Fusion Directory (en version 1.2) : l'intégration d'un module d'invitation et de formulaires d'inscription (sur la base des API) ; le développement d'un système de péremption de comptes adapté à une authentification déléguée, sur la base des états et du cycle de vie dans Supann 2018[14] ;
- sur LemonLdap::NG (en version 1.9) : l'intégration de la fédération d'identités RENATER (moissonnage des métadonnées et diverses adaptations de LemonLdap::NG en tant que IdP et SP).

Le recours à un appel d'offres ne permet pas d'embrasser tout le panorama des solutions techniques disponibles[15], mais il peut garantir des compétences et un niveau de maintenance ici indispensables.

3 Caractéristiques de la plateforme

La plateforme de gestion d'identités a été développée en mode « Agile » (deux ateliers et de nombreuses itérations du mois de juillet à fin octobre 2018, sur 3 mois). La phase de conception fut brève, du fait du travail déjà réalisé lors du POC. En revanche, la phase de tests avec les différents IdP fut assez longue, du fait de l'adaptation de LemonLdap::NG à la fédération et des paramétrages des IdP pour la fourniture d'attributs. Elle s'est achevée par la mise en ligne de la plateforme d'inscription auprès des usagers en décembre 2018.

[13] https://lemonldap-ng.org/

[14] La révision Supann 2018 intégrait en effet une définition précise du cycle de vie de l'identité : https://services.renater.fr/documentation/supann/supann2018/recommandations2018/personnes/alimentation

[15] cf. la conclusion de cet article à propos des solutions de gestion d'organisations virtuelles.

3.1 Gestion des inscriptions

La cinématique de la gestion des inscriptions peut être représentée de la manière suivante :

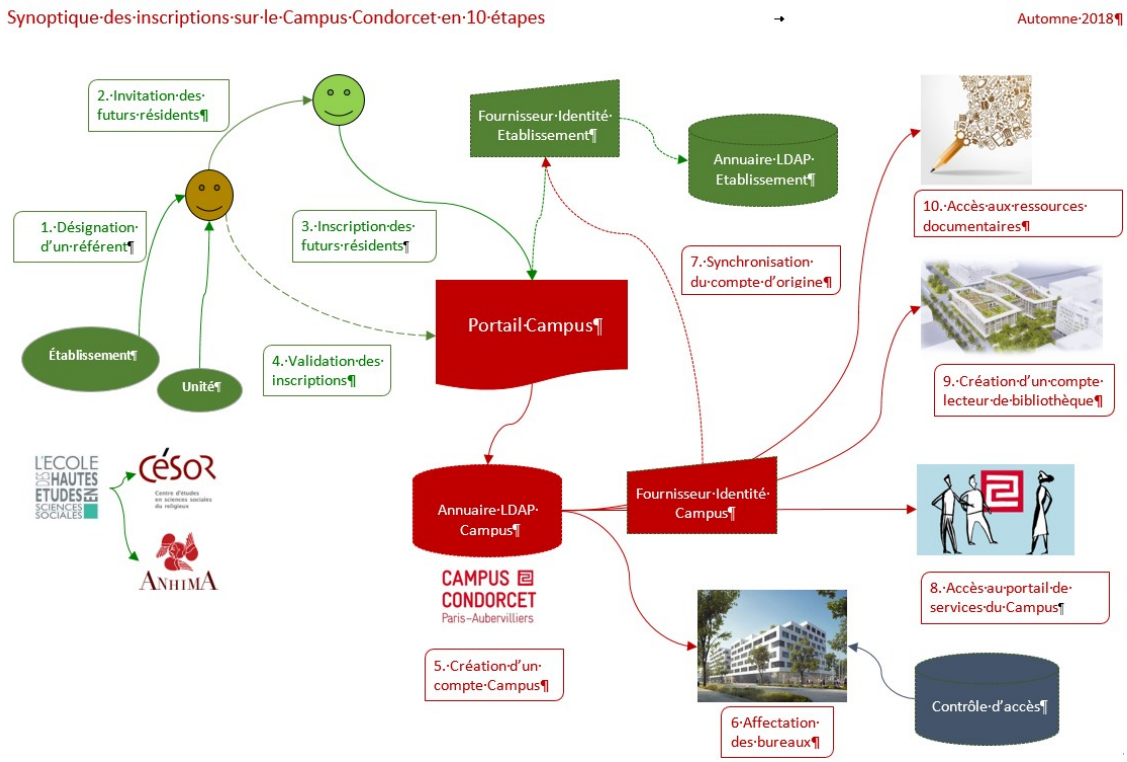


Figure 2 - Cinématique des inscriptions

Etape	Description
1. Désignation d'un référent	Chaque établissement désigne un référent habilité à identifier dans son établissement les personnes devant s'installer sur le Campus
2. Invitation des futurs résidents	Chaque référent invite progressivement une liste de personne sur la base de leur courriel depuis un plugin invitation développé dans Fusion Directory
3. Inscription des futurs résidents	Chaque personne invitée dispose d'un lien unique et peut s'inscrire depuis l'IdP de son choix via la fédération RENATER.
4. Validation des inscriptions	Le référent invitant est notifié de l'inscription et peut valider celle-ci dans Fusion Directory
5. Création d'un compte Campus	Le compte est créé sans mot de passe[16] ; le lien vers l'IdP d'origine est assuré par l'attribut

[16]Lorsque l'authentification via un IdP est impossible (5% des cas), le compte ne dispose ni de mot de passe ni d'identifiant pivot. Les comptes sont alors normalisés a posteriori via les modèles de Fusion Directory en ajoutant un mot de passe local et en lançant une demande de réinitialisation de mot de passe.

Etape	Description
	supannRefId={RENATER}[EPPN]
6. Synchronisation du compte d'origine	Cette étape devait permettre de tester régulièrement l'existence du compte hors session utilisateur. Trop complexe, elle a été abandonnée[17].
7. Affectation des bureaux	Une opération de mise à jour du compte est effectuée a posteriori pour qualifier la personne (affiliation, localisation, etc.)
8. Accès au portail de services du Campus	L'utilisateur peut se connecter via son IdP d'origine au nouveau portail de services du Campus[18]
9. Création d'un compte lecteur de bibliothèque	Le compte de l'utilisateur est exporté vers la solution de gestion de bibliothèque ALMA (ex-libris)[19]
10. Accès aux ressources documentaires	L'utilisateur peut s'authentifier au catalogue de la bibliothèque[20].

Cinématique des inscriptions détaillée

La campagne d'inscription s'est déroulée de décembre 2018 à juillet 2019, avec quelques retards de la part de certains établissements (difficulté à désigner des référents et à circonscrire la population). En juillet 2019, nous avions près de 1700 inscrits et 800 invitations sans réponse.

Aperçu technique

Techniquement, la plateforme articule les deux composants applicatifs (Fusion Directory et LemonLdap::NG) au-dessus d'un annuaire OpenLDAP 2.4.4 :

[17]De fait, l'utilisateur dont le compte d'origine est inactif ne peut plus s'authentifier sur le SI Campus Condorcet (cf. étape 5).

[18]Ibid supra..

[19]Ibid supra.

[20]Ibid. supra.

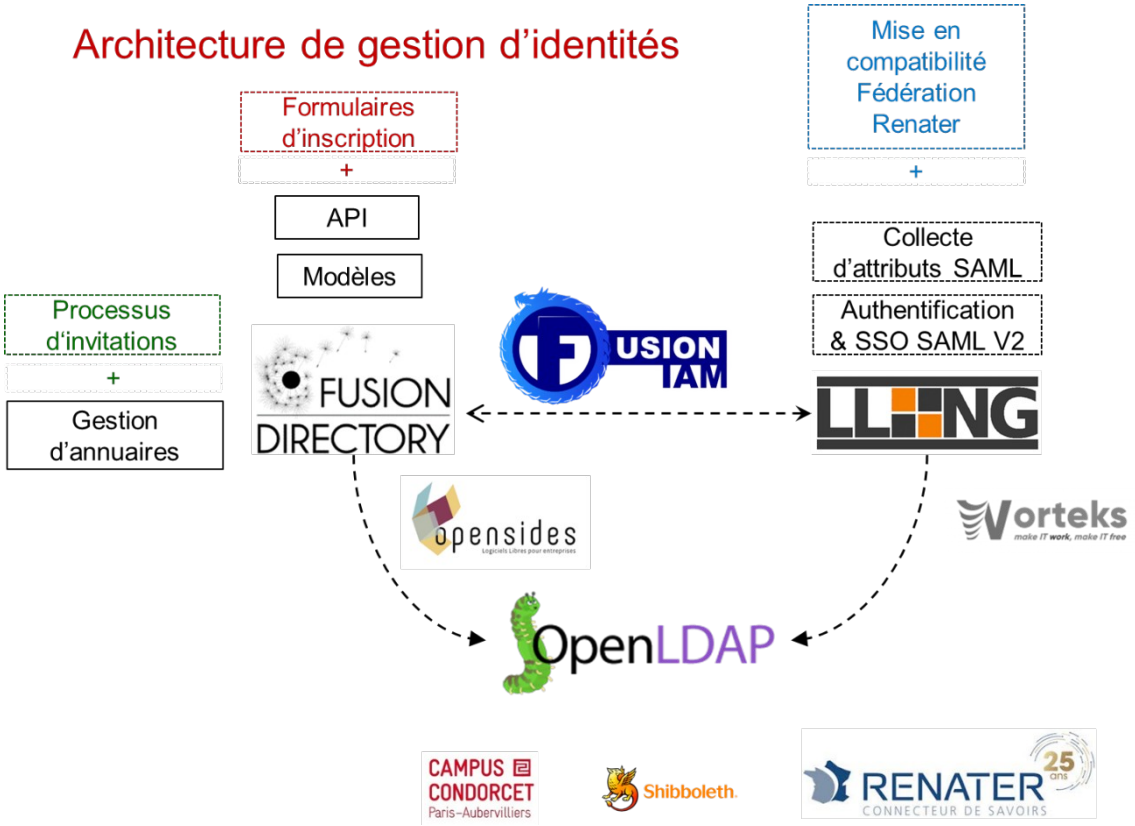


Figure 3 - Architecture technique

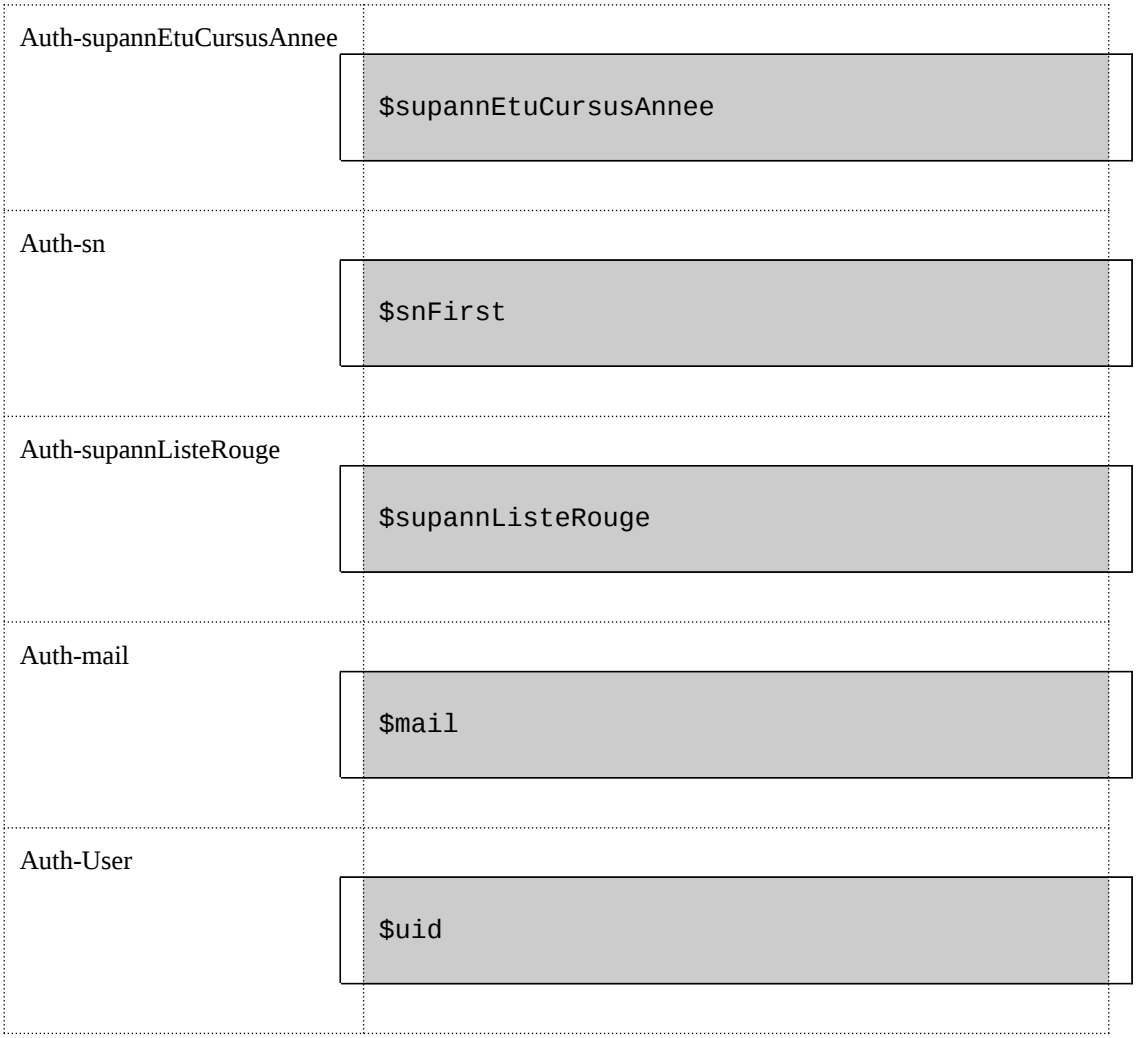
1. Les invitations sont gérées depuis le référentiel via un plugin Fusion Directory développé à l'occasion (pool d'invitations et gestion des inscriptions).
2. Les formulaires d'inscription (« Fédération » et « Libre » sont fondés sur un plugin Fusion Directory, lequel utilise les API et les Modèles de Fusion Directory.
3. Dans le cas du processus d'inscription via la Fédération, LemonLDAP::NG (LLNG) intercepte le processus et se comporte comme SP SAML V2 (via une mise compatibilité de LLNG pour la Fédération RENATER : moissonnage des métadonnées, notamment) :
    a. LLNG collecte les attributs depuis la transaction IdP > SP et les enregistre (via PostgreSQL) dans une session SSO ;
    b. LLNG est activé en mode Proxy NGINX « devant » Fusion Directory : il retravaille les attributs de sessions via des macros[21] et les transmet dans des entêtes HTTP[22] ;
    c. Fusion Directory collecte via Apache les entêtes HTTP et les transmet au formulaire d'inscription ;
4. L'utilisateur visualise les données et valide son inscription : le compte est créé.

[21]Les macros ont été rendues nécessaires pour « normaliser » les valeurs envoyées par les IdP à LLNG, soit lorsque l'attribut eduPerson ou Supann n'était pas présent, soit lorsque la valeur comportait des caractères inacceptables.

[22]Cf. le mode Handler HTTP de LemonLDAP::NG : https://lemonldap-ng.org/documentation/latest/start#handlers

Les attributs gérés dans la transaction sont les suivants :

Entête HTTP	**Variables de session LLNG et macros**
Auth-displayName	`$displayName`
Auth-cn	`$cnFirst`
Auth-eduPersonAffiliation	`( $eduPersonAffiliation =~ /registered-reader/ ) ? $eduPersonAffiliation : ( $eduPersonAffiliation ? $eduPersonAffiliation.\"; registered-reader\" : \"registered-reader\" )`
Auth-eduPersonPrimaryAffiliation	`$eduPersonPrimaryAffiliation ? $eduPersonPrimaryAffiliation : ( $eduPersonAffiliation =~/; / ? \"\" : $eduPersonAffiliation )`
Auth-supannRefId	`$eduPersonPrincipalName ? \"{RENATER}\" . $eduPersonPrincipalName : \"\"`
Auth-givenName	`$givenNameFirst`



Attributs détaillés

Certains problèmes ont été rencontrés lors de la phase de mise en œuvre, concernant notamment des filtres d'attributs trop restrictifs ne permettant pas la diffusion d'attributs demandés. Ils ont été corrigés par des échanges directs avec les responsables des IdP d'origine. Cela étant, les inscrits pouvant s'enregistrer depuis n'importe quel IdP, il est important que la politique de diffusion des attributs soit conforme aux métadonnées diffusées par le SP d'inscription.

3.2 Gestion des identités

Les données

La gestion des identités dans un contexte fédératif soulève un certain nombre de difficultés :

- les données concernant les utilisateurs sont relativement pauvres (*cf.* les attributs de base ci-dessus) ;
- il est impossible de s'interfacer avec un système d'information existant (RH, Recherche, Scolarité) pour mettre à jour ces données, puisque seul l'annuaire d'origine peut être consulté, et ce uniquement lors de l'inscription (pas de mise à jour ultérieure).

Or, les besoins d'interopérabilité concernant le référentiel du Campus sont relativement similaires à ceux d'un établissement :

- Contrôle d'accès (UNICAMPUS)
- Impression (PAPERCUT)
- Téléphonie (AVAYA)
- Demandes de support (GLPI)
- Compte Lecteur (ALMA)
- Authentification Web (LLNG)
- Portail de services (KSUP)
- Réservation de salles (ADE Campus)
- Drive fichiers (Seafile)
- Messagerie (Partage de RENATER)
- Listes de diffusion (SYMPA)
- Autres outils en développement (Forum, Chat)

Comme le représente ce schéma :

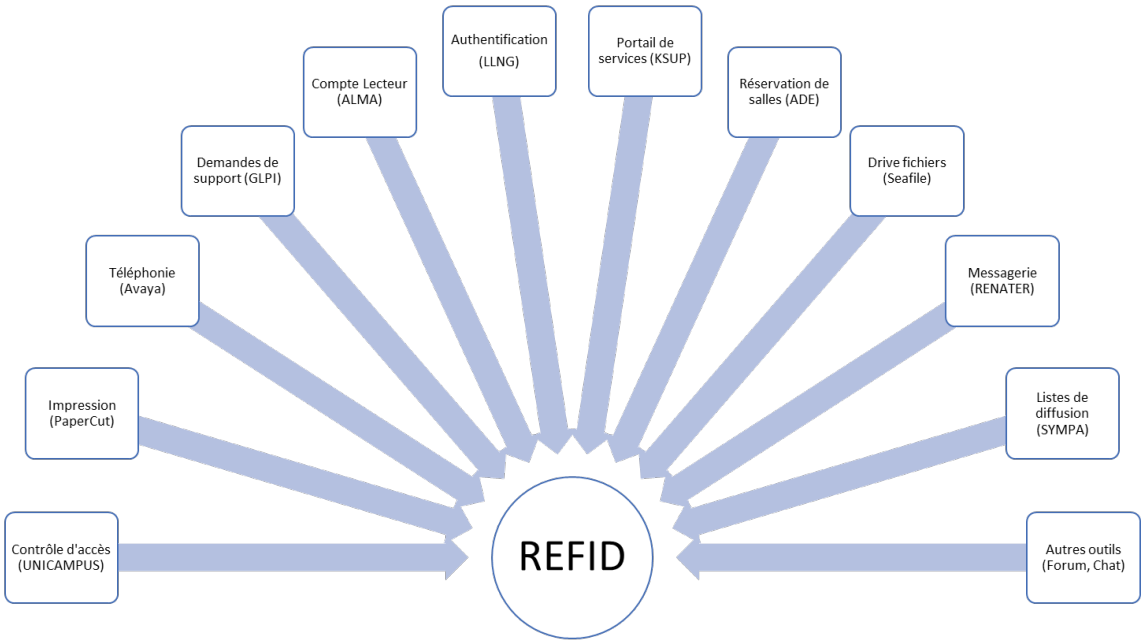


Figure 1 - Schéma synoptique du référentiel d'identités dans le SI Condorcet

Des corrections ultérieures (*cf. infra*) doivent donc être apportées aux données du référentiel, afin que celui-ci apporte les données nécessaires aux règles de gestion minimales des applications (rôles et autorisation, par exemple) :

Type de donnée	Valeur attendue	Problème constaté	Action
Profil *eduPersonAffiliation* *eduPersonPrimaryAffiliation*	student | faculty | staff | employee | member | teacher | registered-reader | affiliate | alum	Les inscrits CNRS sont tous qualifiés d'employés et ne distinguent pas les chercheurs des ITRF Aucune des personnes inscrites librement n'a de profil défini	Aucune : les données restent identiques aux annuaires d'origine
Affiliation *supannEntiteAffectation* *supannEntiteAffectationPrincipale*	Établissement Unité	Aucune des personnes inscrites n'a d'affiliation définie[23]	Corriger ces données (par lots) selon les structures du REFID
Établissement *supannÉtablissement (UAI)*	Établissement	Aucune des personnes inscrites n'a d'établissement[24]	Corriger ces données (par lots) selon les structures du REFID
Id de référence *supannRefId*	EPPN d'origine	Aucune des personnes inscrites librement n'a d'EPPN défini	Corriger ces données au cas par cas ou affecter des mots de passe locaux
Liste rouge dans l'annuaire *supannListeRouge*	Oui | Non	Toutes les personnes sont sur liste Rouge	Proposer à l'utilisateur de définir sa propre politique de publication de données personnelles via le portail
Fonction *Title*	Libre multivalué	Aucune des personnes inscrites n'a de fonction connue	Proposer à l'utilisateur de définir sa propre fonction via le portail
Téléphone professionnel *telephoneNumber*	Libre	Tous les n° de téléphone seront réaffectés sur le Campus	Synchroniser la valeur avec la solution de téléphonie

[23]En effet, même si les attributs peuvent techniquement être transférés, deux problèmes se posent : ils sont rarement renseignés et lorsqu'ils le sont, les règles d'affiliation aux structures dans les annuaires sources (valeurs), définies localement à ceux-ci, ne peuvent être utilisées telles quelles dans le REFID du Campus.

[24]Idem que *supra*.

Type de donnée	Valeur attendue	Problème constaté	Action
Adresse postale professionnelle *postalAddress*	Libre	Toutes les adresses seront affectées sur le Campus	Corriger ces données (par lots) avec les adresses des structures
n° bureau *roomNumber*	Libre	Toutes les n° bureau seront affectés sur le Campus	Corriger ces données (par lots) avec la liste des occupants

L'application de gestion d'identités Fusion Directory apporte de ce point de vue des fonctionnalités intéressantes via les modèles, qui permettent des mises à jour pré-paramétrées sur des comptes utilisateurs sélectionnés.

Toutefois, certaines mises à jour n'ont pu être réalisées avec les modèles (ex : affectation globale de n° de bureaux et de n° de téléphone, notamment). Pour ces besoins précis, l'équipe s'est tournée vers la solution LDAP Synchronisation Connector (LSC[25]).

La maîtrise des données de structures (supannEntiteAffectation, notamment) a représenté un travail important de recensement et qualification et ordonnancement (parent/enfant, tutelles multiples des établissements et unités de recherche. Opéré dans FusionDirectory et appliqué a posteriori aux individus, ce travail permet la publication sur le portail d'un annuaire des structures en parte publique[26] et d'un annuaire des personnes en partie privée.

Les processus d'arrivée

L'arrivée quotidienne des usagers sur le Campus a rendu l'écart entre les données source et les données souhaitées de plus en plus problématique. Le processus d'invitation a donc été revu pour combler la perte de données à l'inscription et raccourcir le délai de mise à jour.

Un processus GLPI a donc été mis en place :

1. un formulaire GLPI est mis à disposition des référents via le portail de services, comportant les champs suivants :
    - Nom + Prénom,
    - Établissement (porteur),
    - Unité (d'accueil),
    - Date de fin de contrat,
    - Courriel,
    - N° de bureau,
    - Adresse MAC ;

[25]Solution libre permettant la synchronisation des données entre plusieurs sources et destinations (LDAP, BDD, CSV) : http://lsc-project.org/. Par ailleurs, LSC est utilisé pour l'export de données du référentiel vers ALMA (export des données des comptes au format XML)

[26]L'annuaire des structures est synchronisé (LDAP > XML) dans KSup et publié sur le portail : https://www.campus-condorcet.fr/annuaire-des-structures-1.

2. Le formulaire est soumis au pôle numérique, qui :
    - invite la personne,
    - valide l'inscription,
    - complète les données dans le référentiel (profil, affiliation),
    - prépare ses droits sur le SI (cf. réservation de salle, Seafile, etc.,
    - enregistre son adresse MAC (accès réseau),
    - confirme l'inscription de la personne sur le ticket GLPI ;
3. Des notifications parviennent :
    - à la personne concernée,
    - au référent demandeur.

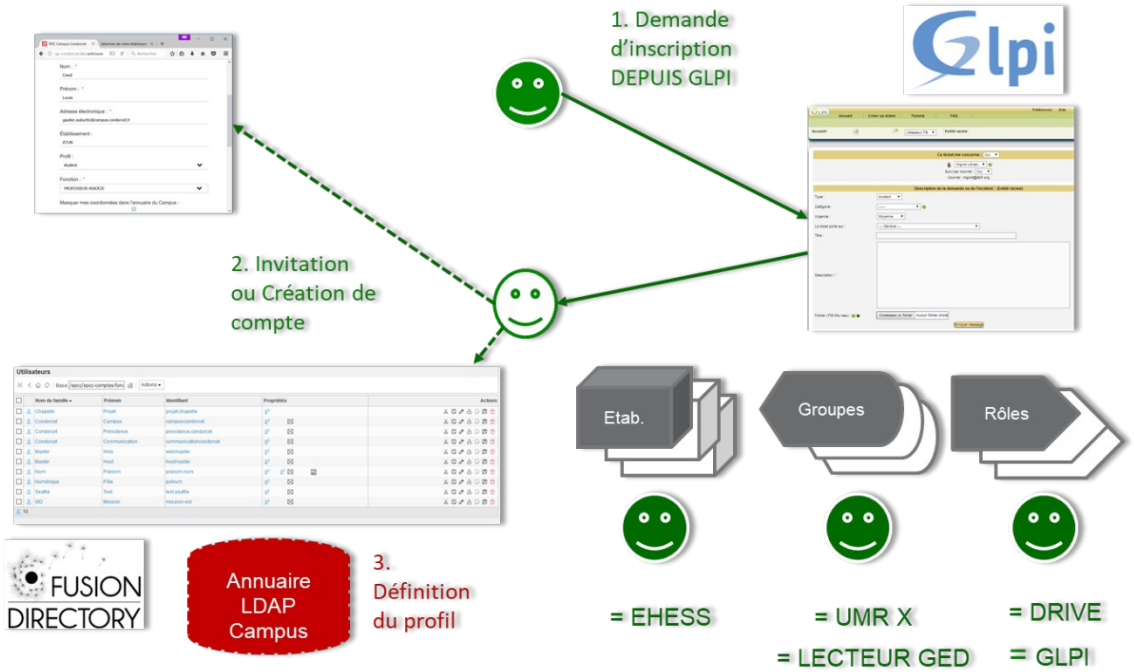


Figure 4 - Processus de demande d'inscription via GLPI

Le processus de départ

Les comptes sont automatiquement verrouillés lorsqu'ils atteignent la date limite de validité du compte définie lors de l'inscription. L'utilisateur peut prolonger son compte, le pôle numérique est prévenu en parallèle. En outre, un processus de départ manuel est assuré par mail / formulaire GLPI.

3.3 Gestion de l'authentification

Troisième brique de la gestion d'identité, la plateforme de SSO LemonLdap::NG fournit un service d'authentification multi-protocoles[27] sur les applications Web du SI Condorcet :

Services	**Application**	**MODE SSO**
Fournisseur d'identités	LLNG 2.0, rattaché à la Fédération RENATER	IdP SAML V2
Portail de services	KSUP 6.7	CAS V2
Portail d'assistance services	GLPI 9.43	CAS V2
Planification, consultation et réservation de salles,	ADE Soft et MyAdeBooking 6.7	CAS V2
Annuaire Pages Blanches	White Pages[28]	Handler HTTP
Gestion des comptes	Fusion Directory	Handler HTTP
Partage de fichier	Seafile Pro 7.09	SP SAML V2
Catalogue de bibliothèque	PRIMO EX LIBRIS (SAAS)	SP SAML V2 / CAS V2

L'intérêt de la plateforme SSO ne se limite pas à une authentification centralisée. LemonLDAP::NG permet aussi le contrôle des accès sur la base de règles fondées sur les caractéristiques de l'utilisateur au sens LDAP (attributs, groupes et rôles).

Exemple d'accès à une application en fonction du groupe de l'utilisateur :

```
default:groupMatch($hGroups, 'cn', 'referents') or
groupMatch($hGroups, 'cn', 'annuaire')
```

En mode Handler HTTP, ces règles peuvent s'appliquer à des fragments d'url :

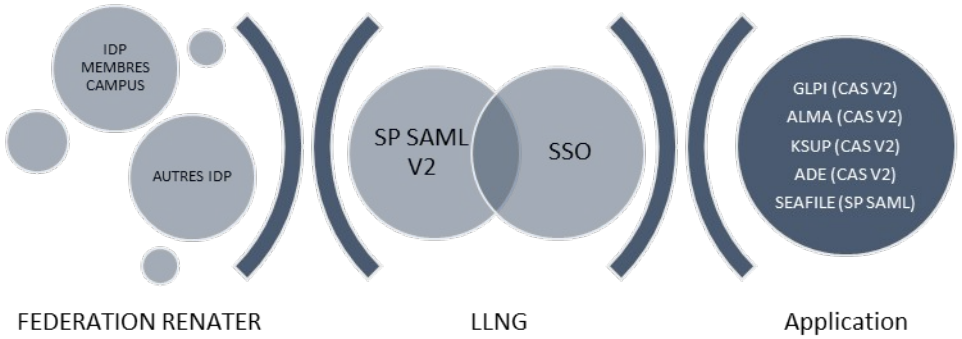
```
default:$_userDB eq "LDAP"
```

[27] L'ensemble des protocoles supportés, qui ne se limite pas au domaine du Web, est décrit dans la documentation officielle de LemonLDAP::NG : https://lemonldap-ng.org/documentation.

[28] L'outil White pages est une application en PHP développée en Open Source : https://ltb-project.org/documentation/white-pages

```
^/publicform:skip
```

Mais plus original encore, LemonLDAP::NG peut se comporter comme un pont d'authentification entre plusieurs protocoles[29]. Intégré en tant que SP SAML V2 dans la fédération d'identités RENATER, il peut déléguer l'authentification à l'IdP d'origine[30] dès lors qu'un utilisateur tente de se connecter sur une application du SI.



Ce modèle d'authentification déléguée présente les avantages suivants :

- le Campus Condorcet ne gère pas le mot de passe de l'utilisateur ;
- l'utilisateur reste rattaché à son établissement d'origine et sa propre politique de gestion d'identités.

En revanche, il repose sur un principe d'affiliation unique de l'utilisateur à un établissement (qui peut ne pas être l'établissement employeur) et présente quelques défauts (marginaux) liés à la maîtrise relative de l'utilisateur de son propre compte d'établissement (cf. les 5% de personnes inscrites sans IdP).

4 Conclusion

Le référentiel d'identités du Campus Condorcet vise à construire un modèle technico-organisationnel de gestion d'identités adapté à des projets fédératifs (Campus, etc.) Il s'apparente à bien des égards aux expériences de gestion d'organisations virtuelles portées notamment par le consortium Géant[31].

Sur le plan technique, les solutions pour répondre au besoin sont nombreuses[32], mais elles nécessitent le plus souvent des compétences techniques très poussées. L'externalisation[33] et la gestion de projet en deux

[29] https://lemonldap-ng.org/documentation/latest/start#authentication_users_and_password_databases

[30] Cela se fait sur la base de l'EPPN de l'utilisateur enregistré à l'inscription.

[31] Par exemple, le service *Eduteams* de Geant https://www.eduteams.org/.

[32] Pour en citer certaines : Grouper et Comanage (Internet2). De nombreuses solutions sont décrites dans : GEANT, « Market Analysis for Virtual Organisation Platform as a Service (VOPaaS) », 2015.

phases (POC puis réalisation) a permis à l'équipe du Campus de pallier ces difficultés et de monter en compétences progressivement.

Sur le plan organisationnel, la démarche de délégation de responsabilité auprès des établissements présente des difficultés réelles (méconnaissance de la population, processus non maîtrisés) mais qui peuvent être surmontées avec le temps et le développement d'un climat de confiance et de co-responsabilité.

Enfin, du point de vue de l'utilisateur, la commodité que présente l'utilisation de son compte d'origine s'ajoute à un sentiment d'appartenance multiple. Le Campus Condorcet est un lieu de résidence, mais aussi un facilitateur d'accès.

Il resterait beaucoup à dire sur les ambitions, les enjeux et les difficultés de la mise en œuvre d'un véritable référentiel d'identités, à l'heure où des projets comme SINAPS émergent, mais il nous faudra un peu plus de temps et de recul pour en tirer les expériences.

Bibliographie

[1] Gautier Auburtin, Référentiel d'identités du Campus Condorcet, Journée Fédération RENATER – 26 septembre 2o18 : https://services.renater.fr/_media/federation/formations/journee-fedid18/journees_federation_2018_cc_refid_fed.pdf

[2] Tangui Coulouarn, « eduTEAMS : gestion des autorisations pour les organisations virtuelles », JRES 2017 : https://conf-ng.jres.org/2017/document_revision_2907.html?download

[3] Youssef Ghorbal, « Retour sur expérience sur la mise en place d'un système centralisé de gestion des identités numériques », JRES 2017 : https://conf-ng.jres.org/2017/document_revision_1978.html?download

[4] GEANT, Market Analysis for Virtual Organisation Platform as a Service (VOPaaS), 2015

[5] SWITCH, « Detailed Architecture for Swiss edu-ID », November 2015.

[6] GIP RENATER, « Constitution d'un référentiel d'identités - Best Practice Document », 2014

[7] University of California, « UC IDM Provisioning Middleware: Technical Design Guide », August 2011.

[33] L'externalisation n'est pas non plus chose évidente dans un marché de niche comme l'enseignement supérieur (Annuaires Supann et Fédération Renater).