

Programme CUS 2013-2016 P-2
«Information scientifique:
accès, traitement et sauvegarde»

**Axe principal «Publications»,
Mesure de mise en œuvre EP-10,
«Numérisation»**

Analyse d'adaptation de la stratégie

20 mars 2015

(Traduction de l'allemand)

Direction du programme:

roland.dietlicher@swissuniversities.ch, directeur du programme

gabi.schneider@swissuniversities.ch, directrice suppléante du programme

Site web: www.swissuniversities.ch/isci

Table des matières

1	Remarque préliminaire	3
2	Historique.....	3
2.1	Mesure de mise en œuvre EP-10, numérisation.....	3
2.2	Résultat du premier appel à projets	3
2.3	Investigations de la direction du programme.....	4
3	Rétrospective sur la formation de la stratégie	5
4	Enseignements	5
4.1	Encouragement de contenus numériques par le programme CUS P-2	5
4.2	Encouragement de nouveaux contenus par rétro-numérisation	6
4.3	Photographies, médias audiovisuels	7
4.4	Collections et environnements de recherche	7
4.5	Exigences qualitatives	8
4.6	Plates-formes de numérisation en tant que services nationaux.....	8
4.7	Centre de compétence en numérisation	9

1 Remarque préliminaire

Le présent document décrit le processus et les réflexions formant la base de la stratégie adaptée dans le domaine de la numérisation. Il porte sur la période séparant les séances du comité de pilotage CUS P-2 du 11 avril 2014 et du 20 mars 2015.

La stratégie est définie dans le document «Axe principal Publications, Mesure de mise en œuvre EP-10, Numérisation: Stratégie et mesures de mise en œuvre adaptées» du 31 mars 2015. Elle remplace les explications de la stratégie partielle e-Publishing et la mesure de mise en œuvre EP-10 prévue dans le «White Paper for a Swiss Information Provisioning and Processing Platform 2020» du 14 avril 2014.

2 Historique

2.1 Mesure de mise en œuvre EP-10, numérisation

Le «White Paper for a Swiss Information Provisioning and Processing Platform 2020» du 14 avril 2014 concrétise la stratégie du programme CUS P-2 dans le domaine de la numérisation en définissant la mesure de mise en œuvre EP-10:

«La numérisation de contenus d'importance nationale sur une plate-forme actuelle ouverte aux participants (retro-seals, e-manuscripta, e-rara, Scriptorium, rero.doc,...). Les exploitants des plates-formes de numérisation offrent leurs services au prix de revient et avec un accord de niveau de service (SLA).» (WP, p. 30, p. 55)

2.2 Résultat du premier appel à projets

Lors du premier appel à projets, sept propositions relatives à la mesure EP-10 ont été déposées, pour une demande d'aide totalisant 7449,2 kCHF. Aucune de ces propositions ne s'est tenue à la définition de la mesure de mise en œuvre figurant dans le White Paper:

Proposition	Appréciation / critique
N° 141-005: retro.seals.ch – Plate-forme pour les revues suisses numérisées: extension de l'infrastructure informatique (Bibliothèque de l'ETH, 200 kCHF)	Proposition d'extension de l'infrastructure.
N° 141-010: Di-ZAS – Collection numérisée d'extraits de journaux sur l'économie suisse (UB Basel, 950 kCHF)	La plate-forme est utilisée à l'interne et n'est accessible à aucun autre partenaire. Il s'agit du perfectionnement (numérisation et ouverture) d'une collection d'importance nationale, sous la responsabilité de l'auteur de la proposition.
N° 141-012: PSEL-SPOL – Presse suisse en ligne (Bibliothèque nationale suisse avec UB Basel, 775 kCHF)	La proposition inclut le subventionnement direct de l'infrastructure. La participation est en principe ouverte, mais elle se limite aux contenus libres de droits d'auteur jusqu'en 1920.
N° 141-013: e-manuscripta – Plate-forme pour sources manuscrites numérisées des bibliothèques et archives (ZB Zurich, 1504 kCHF)	La proposition prévoit aussi le subventionnement direct de l'infrastructure. Le modèle de partenariat limite l'ouverture de la plate-forme pour les participants disposant de faibles moyens.

Proposition	Appréciation / critique
N° 141-023: e-rara – Plate-forme en ligne pour anciens ouvrages numérisés des bibliothèques suisses (Bibliothèque de l'ETH, 2292,5 kCHF)	La proposition prévoit aussi le subventionnement direct de l'infrastructure. La numérisation profite essentiellement aux partenaires déjà impliqués.
N° 141-031: Schweizer Kunst digital – Numérisation et publication en ligne de l'archive de photographies analogiques du SIK-ISEA sur l'art suisse, 1 ^{ère} partie (Institut suisse pour l'étude de l'art, 1238,3 kCHF)	La plate-forme est utilisée à l'interne et n'est accessible à aucun autre partenaire. Il s'agit du perfectionnement (sauvegarde/numérisation et ouverture) d'une collection d'importance nationale, sous la responsabilité de l'auteur de la proposition, à l'horizon 2020. La numérisation de collections de photographies (ou de médias audiovisuels) n'est pas prise en compte dans la stratégie du programme CUS P-2.
N° 141-014: Kompetenzzentrum Digitalisierung – Création et exploitation d'un centre national de coordination et de conseil pour la numérisation et la présentation en ligne (ZB Zurich, 489,4 kCHF)	La proposition ne correspond pas à la mesure EP-10. Elle mêle les intérêts de l'auteur de la proposition (formation d'un organe de pilotage commun pour les plates-formes des partenaires de projet) et l'intérêt que présente un pilotage national de plates-formes et de contenus.

Sur la base de ces constats, toutes les propositions ont été momentanément suspendues lors de la première série d'évaluation. Les auteurs des propositions ont été informés que le comité de pilotage se pencherait une nouvelle fois sur les modalités de l'aide à la numérisation et souhaitait conférer un profil distinct à cette forme d'aide dans le cadre du programme.

Le 14 mai, le comité de pilotage a décidé de suspendre le dépôt de projets de numérisation lors du deuxième appel à projets du 10 août 2014. La direction du programme a été chargée d'élaborer des bases de décision.

2.3 Investigations de la direction du programme

En été 2014, la direction du programme a tenu des entretiens avec différents partenaires sur les exigences liées à la numérisation:

- Susanna Bliggenstorfer, ZBZH, et Marie-Christine Doffey, BN (entretien préliminaire)
- Auteur de la proposition n° 141-014, Kompetenzzentrum Digitalisierung
- Digital Humanities Lab, Bâle
- Projet e-codices
- Fonds national suisse (FNS)
- Académie des sciences humaines et sociales (ASSH)
- Memoriav – Association pour la sauvegarde de la mémoire audiovisuelle suisse

Il s'agissait ainsi d'acquérir une vue d'ensemble et une meilleure compréhension des activités, des besoins et des points de vue des acteurs du secteur de la numérisation, afin de mieux définir le rôle du programme CUS P-2.

Le 12 novembre 2014, une première version du présent document a été discutée lors d'une rencontre avec les auteurs des propositions ainsi que des représentants du FNS, de l'ASSH, d'Infoclio et de Memoriav. Une version remaniée et incluant les mesures de mise en œuvre n'a pas encore été approuvée par le comité de pilotage le 8 décembre 2014. Une discussion approfondie avec des

membres du comité de pilotage (le 16 janvier 2015) a certes permis de confirmer le contenu de l'analyse de la direction du programme, mais les mesures de mise en œuvre ont été une nouvelle fois remise en question. Il en a résulté que le comité de pilotage et les auteurs des propositions devraient se voir soumettre un document plus bref et indiquant plus précisément les propositions que le programme CUS P-2 soutient concrètement et celles auxquelles il n'accorde pas son aide. Par ailleurs, en mars 2015, de premiers aperçus du rapport «Centre de données et de prestations de services» de l'ASSH a permis de mieux coordonner les mesures prévues de part et d'autre.

La stratégie adaptée, avec mesures de mise en œuvre, a été approuvée par le comité de pilotage CUS P 2 le 20 mars 2015.

3 Rétrospective sur la formation de la stratégie

Le chapitre 4.3 du White Paper (p. 24-30) fixe les recommandations d'origine du groupe stratégique e-Publishing pour la numérisation (p. 26-27). La formulation de la mesure de mise en œuvre EP-10 était basée sur l'appréciation suivante émise par le programme:

«E-lib.ch a soutenu la création d'infrastructures de numérisation, de plates-formes de présentation pour différents types de documents (e-rara, retro.seals, e-manuscripta, e-codices) et le développement de standards. Les projets E-lib.ch avaient pour objectif d'établir des modèles de gestion durables. Dans un premier temps, il est judicieux de lier un nouveau soutien à une coordination accrue et à l'ouverture à d'autres participants: il faut encourager des projets de numérisation de collections de portée scientifique et non des infrastructures de numérisation.» (WP, stratégie partielle e-Publishing, p. 29)

Cette appréciation part du principe que la Suisse compte déjà suffisamment de plates-formes de numérisation établies et pouvant accepter des clients. L'ouverture à de nouveaux participants et clients ainsi que le calcul et la facturation des coûts totaux doivent permettre d'assurer l'exploitation durable et le perfectionnement des plates-formes; ces éléments doivent également servir de base à l'aide octroyée pendant la période 2013-2016.

Parmi les plates-formes mentionnées dans le White Paper, e-codices (Bibliothèque virtuelle des manuscrits en Suisse) a été désignée d'emblée comme bénéficiaire du programme par la CUS: un montant de 2000 kCHF a été attribué à e-codices pour la durée du programme. Le programme a été chargé d'intégrer e-codices au sein d'une stratégie nationale.

4 Enseignements

4.1 Encouragement de contenus numériques par le programme CUS P-2

Le programme CUS P-2 veut encourager la mise à disposition d'informations scientifiques numériques et offrir des services pour leur traitement et leur archivage. L'objectif réside dans «la concentration des forces» en Suisse, afin d'affronter les grands défis en optimisant la mise en œuvre des moyens disponibles (cf. la «Stratégie nationale»).

P-2 évalue et soutient des projets visant à créer, étendre et exploiter des services nationaux gérant les informations scientifiques numériques. Les contenus numériques sont soutenus dans l'axe principal «Publications», et ceci dans trois champs d'action: les licences, l'Open Access et la rétro-numérisation.

Le champ d'action des **licences** (EP-1) vise un approvisionnement permanent avec une sélection de produits d'édition «classiques» licenciés au niveau national. Le programme assure le regroupement et le renforcement du pouvoir de négociation, favorise la plus-value stratégique et veille à la transparence de la procédure de sélection. L'utilisation ciblée des moyens du programme doit viser à

créer des synergies optimales avec les investissements courants des bibliothèques des hautes écoles, afin d'obtenir des collections complètes aux meilleures conditions possibles. La sélection des produits eux-mêmes est du ressort des bibliothèques des hautes écoles (Consortium).

Le champ d'action **Open Access** (EP-2, EP-4, EP-5, EP-6, EP-7, EP-8, EP-9) encourage le perfectionnement des infrastructures du Green Open Access et les projets pilotes Gold Open Access. Ici également, le programme n'évalue pas des contenus, mais crée des incitatifs pour le développement de modèles de publication et de financement: EP-8 permet d'évaluer la pertinence du procédé (exemple concret: clé de financement pour SCOAP3), EP-9 vise à créer des incitatifs sans effet sur les contenus encourageant les revues de l'univers universitaire clairement établies à opter pour un modèle Open Access. Les créations de nouvelles revues n'entrent donc pas en ligne de compte.¹

Le champ d'action **Rétro-numérisation** (EP-10) intervient au niveau de l'encouragement de plates-formes déjà établies. Il doit veiller à l'ouverture et au perfectionnement de ces plates-formes.

Pour la rétro-numérisation, à la différence des licences et de l'Open Access, la mesure de mise en œuvre formulée prévoit l'évaluation des contenus des publications par des experts et des groupes de spécialistes, alors que les caractéristiques techniques des plates-formes ne font pas l'objet de l'examen. Lors de l'attribution aux experts des propositions mixtes du premier appel à projets, il s'est vite avéré que l'évaluation des aspects techniques et organisationnels d'une part et des contenus d'autre part nécessitait des profils de compétence différents. Le processus d'évaluation aurait ainsi débouché sur une impasse même pour les propositions s'inscrivant correctement dans le cadre de la mesure de mise en œuvre EP-10.

Ainsi, également dans le domaine de la rétro-numérisation, le programme CUS P-2 ne peut évaluer que les bases techniques et organisationnelles. Les contenus doivent être sélectionnés par des partenaires, le programme veillant ici au respect des «règles du jeu» et à la transparence du processus.

4.2 Encouragement de nouveaux contenus par rétro-numérisation

L'aide de départ allouée via E-lib.ch a largement contribué à la mise en place de l'offre actuelle de contenus numériques. Avec E-lib.ch, l'établissement et l'exploitation des plates-formes ont été soutenus par la prise en charge de frais d'infrastructure et de personnel. La sélection des contenus était confiée aux exploitants. L'aide apportée à e-codices par le programme CUS P-2 respecte le même principe.

Les entretiens menés en été ont révélé des divergences de vue sur la question de savoir si le programme CUS P-2 doit soutenir les contenus rétro-numérisés en eux-mêmes. Dans le domaine du patrimoine culturel, il est difficile de tracer une frontière entre les compétences du programme et celles des bailleurs de fonds public – notamment la Confédération et les cantons. On a répété qu'il est toujours possible de trouver des bailleurs de fonds pour la numérisation elle-même. Toutefois, cette position a aussi des opposants résolus, selon lesquels c'était peut-être le cas il y a encore quelques années, mais plus aujourd'hui compte tenu de l'ampleur des besoins.

La nécessité d'une aide de départ est incontestée en ce qui concerne les petites institutions, qui ne disposent souvent pas des moyens nécessaires pour lancer de premiers projets de numérisation de contenus présentant un intérêt scientifique. Il serait également souhaitable d'ouvrir les plates-formes pour répondre aux besoins liés aux projets de recherche des hautes écoles. On peut en conclure que l'aide axée sur les contenus doit profiter avant tout à de nouveaux participants de petite taille et sans

¹ L'Open Access et l'Open Data constituent en outre des interfaces standard pour des contenus subventionnés. CUS P-2 doit contribuer à l'établissement d'un espace de données ouvert et réutilisable. C'est aussi l'objectif de la mesure de mise en œuvre EP-3 (établissement d'un centre de compétence pour l'Open Access et les droits d'auteur en Suisse).

grands moyens financiers. Or les propositions déposées indiquent que, sans mesures de pilotage, les aides profiteraient essentiellement aux plates-formes actuelles. Les mesures doivent donc favoriser une ouverture plus large des plates-formes et la prise en compte des coûts totaux pour tous les participants.

Comme indiqué au chapitre 4.1, le programme CUS P-2 n'a pas pour tâche de sélectionner lui-même les contenus. Ceux-ci peuvent donc rester financés à titre d'éléments des projets, à l'instar du modèle des licences et de l'Open Access, ou il faut créer pour cela un organe d'évaluation séparé.

Le groupe stratégique e-Publishing avait proposé de financer de nouveaux contenus par le biais d'un fonds géré par un comité de coordination national réunissant toutes les plates-formes. Cette solution présenterait notamment l'avantage de permettre de régler de manière consensuelle une éventuelle concurrence des plates-formes pour les nouveaux contenus (par ex. e-codices et e-manuscripta). Mais il faudrait alors définir précisément ce qu'il faut entendre par «toutes» les plates-formes. L'une des voies possibles consisterait à introduire une «certification» des plates-formes participantes en fonction des exigences mentionnées aux chapitres 4.5 et 4.6.

Au niveau de l'encouragement de contenus, des attributs tels que «scientifiques» ou «d'importance nationale» peuvent couvrir un large domaine. D'une part, la recherche est soutenue en fonction de thèmes stratégiques et non en évaluant s'il est judicieux de numériser les photographies d'une collection d'art historique, des articles de presse d'intérêt scientifique, des contenus didactiques historiques, des émissions télévisées du XXe siècle ou encore des images et des plans de jardins. La liaison entre des thèmes de recherche stratégiques et les besoins de numérisation (par ex. dans le cadre des projets d'édition soutenus par le FNS) pourrait constituer un critère déterminant. D'autre part, il existe des interactions entre l'offre et la demande. La pratique éprouvée des collections consiste à lancer de nouvelles questions de recherche par la mise à disposition proactive de contenus.

4.3 Photographies, médias audiovisuels

Jusqu'à présent, l'aide à la numérisation de médias audiovisuels n'était pas prise en compte dans la stratégie du programme CUS P-2. Si une telle aide devait être jugée importante et si l'intérêt de services nationaux devait se faire jour, il faudrait commencer par réunir des bases d'élaboration de mesures de mise en œuvre en collaboration avec les acteurs des secteurs BAM (bibliothèques, archives et musées), resp. GLAM (Galleries, Libraries, Archives, Museums).

Avec les supports de données modernes, l'aspect de mesure de sauvetage offert par la numérisation gagne de l'importance: ces supports de données sont plus fragiles que le papier et souvent gravement menacés de décomposition. Le programme ne nie pas cette menace mais n'est pas en mesure pour l'instant de répondre également à ces attentes (cf. proposition n° 141-031 «Schweizer Kunst digital»).

Une stratégie possible pourrait être élaborée par ex. avec Memoriav (Association pour la sauvegarde de la mémoire audiovisuelle suisse). Le programme CUS P-2 devrait rester ouvert et attentif à l'évolution future de ce domaine.

4.4 Collections et environnements de recherche

Les plates-formes e-rara, retro.seals.ch, e-codices et Kartenportal.CH, dont l'établissement a été soutenu par E-lib.ch, avaient toutes pour élément central la numérisation de collections des partenaires participants. Toutes répondent à une même prise de conscience: des collections sélectionnées doivent être rendues plus visibles, des originaux précieux doivent être protégés et des documents doivent être diffusés sous forme électronique («vitrine», pérennisation, utilisation 24/7).²

² Pour d'autres exemples, voir le site d'archive du portail e-lib.ch: <http://www.e-lib.ch/copy/fr/Angebote2/Digitale-Sammlungen.html>

Ces plates-formes sont souvent axées sur des types de documents (imprimés, manuscrits, cartes, photographies, etc.).

Ce concept ne répond que partiellement aux exigences de la recherche. Les chercheurs sont certes reconnaissants de disposer ainsi des sources nécessaires sous forme numérique et ils en font un usage intensif. Mais le traitement de questions de recherche spécifiques nécessite maintenant l'agrégation de documents numériques de différentes collections, leur utilisation dans des environnements propriétaires tels que les portails spécialisés ou les environnements de recherche virtuels, ou la création de nouveaux documents numériques. C'est pourquoi de nouvelles plates-formes (consacrées par ex. aux éditions numériques) et des environnements de recherche (par ex. SALSAA au DHLab Bâle) travaillant avec des documents numérisés apparaissent dans les domaines spécialisés en relation avec les sciences humaines numériques (Digital Humanities). Ces applications sont souvent thématiques.

Du point de vue du programme, il est souhaitable que les chercheurs tirent parti de l'expérience réalisée par des plates-formes déjà soutenues avec des projets et des environnements de numérisation et puissent le cas échéant solliciter des prestations de services. Inversement, les plates-formes devraient réagir aux nouvelles exigences de la recherche et évoluer en conséquence.

Dans ce contexte, il faut tenir compte du fait que les projets de recherche soutenus par le Fonds national suisse comprennent des volets de numérisation: par exemple, la numérisation de sources analogiques et leur mise à disposition pour d'autres applications scientifiques (métadonnées, etc.) fait partie du processus de recherche. En outre, l'Académie des sciences humaines et sociales (ASSH) prend part à des projets d'édition de longue durée. Il serait indiqué de réfléchir aux moyens de favoriser la haute valeur technique et la durabilité de tels projets dans le domaine des hautes écoles via des ponts vers le programme CUS P-2.

4.5 Exigences qualitatives

Les entretiens menés ont permis de cristalliser les principales caractéristiques d'une plate-forme de numérisation.

Dans cette approche, un **environnement de numérisation** constitue un «ensemble de services permettant de définir des flux de travail pour différents environnements d'utilisation capables de communiquer entre eux».

Un **projet de numérisation** réunit les six éléments suivants:

1. Standards techniques du processus de numérisation
2. Métadonnées complètes et standardisées (niveaux technique, administratif, structurel et descriptif)
3. Adressage persistant
4. Conservation pérenne, archivage
5. Établissement des droits d'utilisation
6. Interfaces standardisées et ouvertes (selon LOD et TEI).

Ces caractéristiques assurent la durabilité et l'interopérabilité; elles doivent être utilisées pour le développement et l'évaluation des plates-formes de numérisation.

4.6 Plates-formes de numérisation en tant que services nationaux

À la différence des «services nationaux» potentiels dans d'autres domaines, les plates-formes de numérisation ont ceci de particulier que l'aide octroyée dans le cadre d'E-lib.ch a servi à encourager la création de plusieurs plates-formes. Et l'orientation de ces plates-formes diffère au niveau des contenus, de la technique et de l'exploitation.

Du point de vue technique, au-delà de l'appareillage physique, les fonctions de base de la numérisation comprennent les quatre composants informatiques suivants:

1. Dépôt des documents numérisés et des métadonnées avec interfaces
2. Fonctions pour le contrôle d'accès et la conservation pérenne
3. Présentation (navigation, recherche)
4. Outils de traitement

Les deux premiers composants sont indépendants de la forme concrète de l'original analogique et du type de l'utilisation. C'est souvent le cas aussi des composants 3 et 4, soit les logiciels de base servant à la présentation et à la recherche ainsi que les outils de traitement. La standardisation et la consolidation des plates-formes permettent de générer des synergies. Le programme ne met donc pas l'accent sur la création de nouvelles plates-formes. Celles-ci seraient confrontées aux mêmes problèmes que les plates-formes existantes: le financement du perfectionnement et de l'exploitation. C'est pourquoi le programme mise sur le perfectionnement de plates-formes existantes et leur ouverture à de nouveaux clients. Même des institutions ou des projets de petite envergure doivent pouvoir faire numériser et déposer leurs objets dans des environnements bénéficiant d'un appui et d'un soutien professionnels.

En règle générale, l'entretien de ces plates-formes ou environnements ne peut être assuré à long terme que par des prestataires solides des hautes écoles (par ex. les grandes bibliothèques des hautes écoles en collaboration avec les centres de calcul et SWITCH).

La concentration des plates-formes sur certains types de médias (codes, manuscrits, anciens imprimés, cartes, journaux, etc.) ou certains thèmes constitue un obstacle majeur à leur consolidation. Le programme peut intervenir à ce niveau et favoriser par ex. l'ouverture de plates-formes actuelles pour différents types de documents numérisés, le découplage des composants informatiques susmentionnés, l'extension vers une infrastructure multi-clients ou la réunion de plusieurs plates-formes.

4.7 Centre de compétence en numérisation

Les expériences réalisées lors du premier appel à projets confortent les recommandations d'action originales de la stratégie partielle e-Publishing pour la numérisation, qui accordent un rôle important à la coordination nationale, à assurer par un comité de coordination. Citation: «Ce comité coordonne les projets de numérisation et les demandes de nouveaux partenaires, fixe et échange des standards et de meilleures pratiques, et réunit les différents services concernés» (WP, p. 26s, C.2).

Jusqu'à présent, le conseil en matière de technique de numérisation appropriée, de plates-formes ou de financement est fourni par tous les exploitants des plates-formes plus ou moins selon les besoins. En outre, la Bibliothèque nationale suisse propose Digicoord, plate-forme d'information sur les projets de numérisation suisses (<http://www.digicoord.ch>) sur tous ses sites. Aujourd'hui, Digicoord s'adresse aux collaborateurs professionnels des bibliothèques et archives. Digicoord aide à se faire une idée de l'état d'avancement des projets suisses de numérisation, propose des informations sur l'organisation des projets ainsi que sur la technique et favorise la mise en œuvre en fournissant des adresses et un forum pour les échanges d'expérience.

L'occasion s'offre ainsi de viser une extension du conseil et de la coordination des projets de numérisation au niveau des responsables des projets dans la recherche ainsi qu'une collaboration avec les spécialistes des portails spécialisés, des environnements de recherche et des centres de données, notamment dans le domaine des sciences humaines. Cela aiderait les chercheurs à s'orienter et améliorerait la qualité et les rendements à long terme tout en réduisant les charges globales.

Un comité de coordination ou un centre de compétence devrait posséder un angle de vue national, mais connaître et soutenir les préoccupations de toutes les régions. Ce service devrait également disposer d'une direction largement soutenue et acceptée. Il reste à définir comment la durabilité d'une telle institution pourrait être assurée. Si l'ouverture et la consolidation des plates-formes actuelles réussissent, l'importance d'un tel centre de compétence diminuera d'autant.