



PROJET SEMESTRE 6 - PV

Séance 1 (kick-off)	1
20 février 2019 / 8:30	1
Séance 2 (follow-up)	2
27 février 2019 / 8:30	2
Séance 3 (visit at e-sculape)	3
5 mars 2019 / 9:00	3
Séance 4 (follow-up)	5
6 mars 2019 / 9:00	5
Séance 5 (analyse financière)	6
13 mars 2019 / 9:00	6
Séance 6 (follow-up)	7
20 mars 2019 / 9:00	7
Séance 7 (projet existant)	8
27 mars 2019 / 9:00	8
Séance 8 (analyse techno)	9
3 avril 2019 / 8:30	9
Séance 9 (séance mandant)	10
10 avril 2019 / 9:00	10
Séance 10 (follow-up)	11
17 avril 2019 / 9:00	11
Séance 11 (réalisation)	12
1 mai 2019 / 9:00	12
Séance 12 (final)	13
8 mai 2019 / 9:00	13

Séance 1 (kick-off)

20 février 2019 / 8:30

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Présentation du projet
- Discussion des objectifs pour le cahier des charges
- Questions générales sur le machine learning
- Analyse de marché (I&C)

NOTES

- Application existante (e-sculape) en WinDev, output en CDACH, application python qui viendra la compléter
- Système OCR déjà en place avec reconnaissance du type et informations (noms, dates...) avec code QR
- Objectif : scannage de masse, extraction de données (dactylo et manuscrit) et classification pour eHealth au moyen du machine learning
- E-sculape possède une DB des médecins et clients (HSFQS)
- Faire une étude de marché et analyse de concurrence (state of the art)

DÉCISIONS

- M. Buntshu n'est plus responsable du projet

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Cahier des charges	26.2	x			
Envoyer doc + DB	26.2				x
Planning	26.2	x			

Séance 2 (follow-up)

27 février 2019 / 8:30

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Présentation du cahier des charges
- Discussions sur la suite du projet

NOTES

- Différencier milestones/objectifs et activités/cahier des charges
- Une activité contient des étapes et produit un livrable
- Le plan définit les étapes
- Plusieurs activités mènent à l'objectif
- Introduction personnelle pas nécessaire
- Contexte : ce qui a déjà été fait
- Nécessaire de définir ce qui va être fait concrètement, une finalité, c'est l'objectif

DÉCISIONS

- Début de séance à 9h00
- Envoie convocation mardi midi

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Visite chez e-sculape	6.3	x			
Nouveau cahier des charges corrigé	1.3	x			

Séance 3 (visit at e-sculape)

5 mars 2019 / 9:00

PARTICIPANTS

Jérôme Clément, Samuel Frehner, Patrick Audriaz

ORDRE DU JOUR

- Comprendre l'input donné à l'application (document scanné) ainsi que tout le processus actuel
- Repartir avec des exemples pour commencer l'analyse technologique et des potentiels tests (plusieurs types de documents de plusieurs médecins de personnes décédées comme convenu)
- Comprendre le format de données souhaité en output afin de faire le match document/patient/médecin

NOTES ÉCONOMIQUE

- Reprendre l'analyse qu'a fait M. Frehner, la juger dans un contexte où le scannage est fait de manière indépendante. Proposer des alternatives.
- Deux vecteurs : Médecins qui partent à la retraite et qui veulent numériser leurs archives ou cabinet qui veut passer du papier au numérique
- Le service de scannage ferait office de produit d'appel
- Pour le moment, e-sculape propose un ERP (logiciel qui permet de gérer l'ensemble des processus opérationnels d'une entreprise en intégrant plusieurs fonctions de gestion) de la santé pour les CABINETS

NOTES TECHNIQUES

- Le résultat est important, le moyen et la technologie pour y arriver l'est moins
- Les documents doivent être triés par ordre chronologique (plus récent au plus vieux)
- Le scannage sort un document PDF/A qui contient l'OCR

- L'OCR est effectué grâce au logiciel "Capture Perfect"
- Chaque document correspond à un seul fichier PDF, ils sont ensuite réunis par type dans un seul PDF (rassemble les documents qui n'ont pas de QR code qui suivent un document avec un QR code)
- Le PDF des documents rassemblés est encapsulé en base64 dans un XML
- Ordre de conception proposé : reconnaissance du type de documents, puis client, puis du médecin (sender), puis cabinet
- Trouver un moyen de segmenter (séparer) les types de documents

DÉCISIONS

- Les PV doivent être transmis à M. Clément
- Séances régulières avec le mandant pour qu'il puisse suivre l'avancement

Séance 4 (follow-up)

6 mars 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Feedback sur le cahier des charges
- Retour sur la visite chez e-sculape et informations utiles pour la suite du projet

NOTES

- Modifications cahier des charges : version et filière sur première page et mise à jour des meta data
- Ajouter segmentation des documents au cahier des charges
- Potentiel problème des carrés noirs mis pour assurer l'anonymat des documents, pour de la reconnaissance d'image
- Descripteur de texte pour classifier les documents
- Chercher un moyen de trouver la position des éléments dans un PDF (Sugarcube)

DÉCISIONS

- Validation du cahier des charges

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Analyse financière	13.3	x			
Documentation de la visite chez e-sculape	6.3	x			

Séance 5 (analyse financière)

13 mars 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Présentation de l'analyse financière

NOTES

- Ajouter intro et conclusion à l'analyse économique
- Se documenter sur les système DEP en Europe
- Schéma liaison entre patientdossier.ch / communautés de références / ERP e-sulape
- Comment e-sculape met en oeuvre leur ERP (serveur, cloud ?)
- Introduction pour la partie "produit développé"
- Mieux définir la solution pour les médecins qui partent à la retraite (sans ERP)
- Demander à e-sculape s'ils reçoivent les dossiers déjà séparées par clients

DÉCISIONS

- -

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Corrections de l'analyse financière	15.3	x			
Documentation du processus actuel	20.3	x			

Séance 6 (follow-up)

20 mars 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Présentation de l'analyse financière corrigée
- Discussion sur l'avancement (analyse du projet existant)

NOTES

- Il n'y a pas d'échange d'informations entre les systèmes primaires, tout transite par le DEP.
- Il faut commencer l'analyse technologique

DÉCISIONS

- -

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Documentation du processus actuel	26.3	x			

Séance 7 (projet existant)

27 mars 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Présentation de l'analyse du projet existant
- Début de l'analyse technique

NOTES

- Synthétiser DEP/LDEP/eHealth afin de pas se répéter
- Clarifier les notions de système primaire et secondaire
- Modifier le client "part à la retraite" pour "repreneur d'activité"
- Étudier deux possibilités de classification : text analysis / image analysis et demander à M. Fischer son avis

DÉCISIONS

- Conclure l'analyse financière

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Analyse technologique	3.4	x			

Séance 8 (analyse techno)

3 avril 2019 / 8:30

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Présentation de l'analyse technologique
- Agender séance avec le mandant pour présenter les analyses

NOTES

- Analyse du "triplet learning"
- Modifications du schéma d'architecture
- Clarifications des informations connues / inconnues (known doctor, unknown patient, one patient file at a time)
- Un modèle CNN par docteur
- S'intéresser aux gain de performances en localisant le texte sur le PDF (top / bottom)

DÉCISIONS

- -

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
RDV mandant	10.4	x			

Séance 9 (séance mandant)

10 avril 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz. Jérôme Clément

ORDRE DU JOUR

- Approuver PV précédent
- Présentation de l'analyse technologique corrigée
- Mandant présent pour discuter de l'avancement
- Présentation du début de la réalisation

NOTES

- Corrections sur l'analyse selon remarques M. Clément
- Pas de data-augmentation dans le projet
- Contact aide Keras : Michael Jungo
- Important de trouver la source du document en plus (HUG, CHUV...) afin de les classifier indépendamment
- Séance avec M. Clément après les vacances de Pâques
- Tester le triplet loss
- Documents en couleurs ou NB

DÉCISIONS

- -

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
CNN fonctionnel sur Tobacco	-	x			
Correction analyse	-	x			

Séance 10 (follow-up)

17 avril 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Discussions sur la suite de la réalisation

NOTES

- Faire des tests sérieux, évaluation et visualisation des résultats et de la fiabilité (train, val, test)
- Concentration sur l'analyse d'image, puis le reste
- Indiquer ce que sont les "next step" pour qqn qui reprendrait le projet
- Conversion PDF → JPG remise à plus tard, une fois le CNN fonctionnel
- Projet fini à 95% à la rentrée des vacances

DÉCISIONS

- Se concentrer sur l'analyse d'image

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
CNN fonctionnel sur Tobacco	-	x			
Correction analyse	-	x			

Séance 11 (réalisation)

1 mai 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Présentation du travail effectué durant la semaine de vacances (conception, réalisation, tests)
- Discussions sur la suite et fin du projet

NOTES

- Documenter que nous sommes dans une approche “very little data” pour un CNN classifiant des images (+ dire nombre de data)
- Trouver un moyen d'évaluer le Fuzzy String Matching (optionnel)
- Ajouter un sommaire pour la réalisation du Transfer Learning
- Ré-entraîner le modèle avec un dataset “balanced” et comparer
- Documenter pourquoi je n'ai pas utilisé RVL-CDIP

DÉCISIONS

- -

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
Finaliser tests	8.5	x			
Conclure le rapport	8.5	x			
Correction analyse	8.5	x			

Séance 12 (final)

8 mai 2019 / 9:00

PARTICIPANTS

Andreas Fischer, Nicolas Schroeter, Patrick Audriaz

ORDRE DU JOUR

- Approuver PV précédent
- Finalisation du rapport
- Exigences pour la présentation ?

NOTES

- Relancer M. Clément pour l'invitation à la présentation
- Présentation orale :
 - 1 slide sur les enjeux / environnement
 - Présentation technique
 - Expliquer le problème et la démarche
 - 1) Input, 2) Output, 3) Black Box
- Ajouter date de consultation pour les références de site web
- Envoyer rapport PDF final par email

DÉCISIONS

- Rapport prêt pour le rendu

TÂCHES (TO-DO)

TÂCHE	DÉLAI	AU	FI	Sch	e-S
-					