

### Institut national de physique nucléaire et de physique des particules



Préservation des logiciels, les apports du projet PRESOFT

Geneviève Romier, CC-IN2P3

contact: <a href="mailto:genevieve.romier@cc.in2p3.fr">genevieve.romier@cc.in2p3.fr</a>

Dialogu'IST, 9 juillet 2020





# Contexte et initiatives autour des plans de gestion de logiciel



## **Contexte**

- La plupart des résultats scientifiques s'appuient sur des analyses de données, simulations, ou calculs obtenus grâce à des logiciels produits par les chercheurs et les équipes de recherche. Ils font partie intégrante du processus de recherche.
- Les incitations se font de plus en plus pressantes à :
  - l'open data et aux FAIR data,
  - la réutilisation des données, la reproductibilité, la pérénisation des données de la recherche,
  - la réalisation de Plans de Gestion de Données :
    - Plan national pour la science ouverte : « Généraliser la mise en place de plans de gestion des données dans les appels à projets de recherche »
- Incitations qui s'appliquent aussi aux logiciels de la recherche :
  - Plan de gestion de données ou Data Management Plan (DMP) pour les données,
  - Plan de gestion de logiciel ou Software Management Plan (SMP) pour les logiciels.



## **SMP** versus **DMP**

Plan de Gestion de données : document formalisé (livrable du projet) :

- Contenu variable selon les projets et les agences de financement.
- Feuille de route, en ajustement constant centrée sur les données et les jeux de données (datasets) du projet de recherche.
- Le logiciel y est vu comme un outil associé aux données (tools, codes).
  - Une partie du DMP peut concerner le logiciel dans certains appels.

Plan de Gestion de Logiciel : focalisé sur le logiciel.

- peut être demandé spécifiquement pour un projet (Royaume Uni, appels particuliers).
- outil pour les responsables de logiciels :
  - document unique de référence qui concentrer les informations, document évolutif,
  - à la fois **état des lieux** et **feuille de route**, qui englobe les différents projets de financement du logiciel.
- Ne pas confondre plan de gestion de logiciel et plan de développement du logiciel !
   mais le plan de développement est inclus/référencé dans le SMP.

DMP et SMP peuvent/doivent se référencer mutuellement!



### Avant « FAIR »

- L'European Southern Observatory (ESO) :
  - « VLT Instrumentation Software Template for Software Management Plan », 2005.
- Le Software Engineering Institute :
  - « COTS and Reusable Software Management Planning: A Template for Life-Cycle Management », 2007, méthode d'évaluation de logiciels ou composants logiciels dans un contexte industriel et dans une optique de gestion à long terme.
- Les expériences du Large Hadron Collider au CERN :
  - « CMS data preservation, re-use and open access policy », 2012
    - http://opendata.cern.ch/record/411



## Initiatives à l'international

- Knowledge Exchange:
  - Workshop on Research Software Sustainability: "A study recommending practices of software sustainability to minimise the risks to reliability & reproducibility of research.", 2016
- Le Software Sustainability Institute (SSI):
  - « Checklist for a Software Management Plan », 2016, liste des principales questions se poser pour gérer un logiciel, questions associées à des questions plus précises et à des conseils.
  - Depuis, le Software Sustainability Institute a fait évoluer la *checklist*.
  - Le modèle SSI est disponible dans



- Le groupe d'intérêt Software Source Code de la RDA :
  - https://www.rd-alliance.org/groups/software-source-code-ig



# Le projet PRESOFT Preservation for REsearch SOFTware

 Projet CNRS - IN2P3 (2017-2019)



CNRS - IN2P3

- ▶ LPC V. Breton,
- CC-IN2P3 G. Romier,



CNRS - LIGM (2017-2018)

T. Gomez-Diaz

- Cadre scientifique :
  - Pérennisation du logiciel
- Objectif scientifique :
  - Etude de l'implémentation de plans de gestion de logiciel dans les unités de recherche.
- Démarche pragmatique : Adaptation aux besoins des chercheurs.
- Résultat principal : un modèle de plan de gestion de logiciel
  - sous forme d'un <u>document texte bilingue</u>
  - sous forme de modèles (FR, EN) intégrés dans <</li>



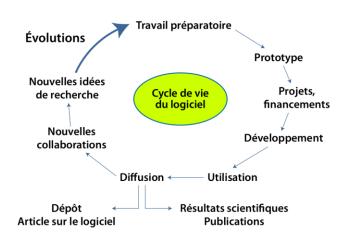


# Qu'est-ce qu'un plan de gestion de logiciel ? Exemple de PRESOFT



# PRESOFT : procédure et modèle

#### Cycle de vie du logiciel de la recherche



A chaque étape renseigner ou mettre à jour :

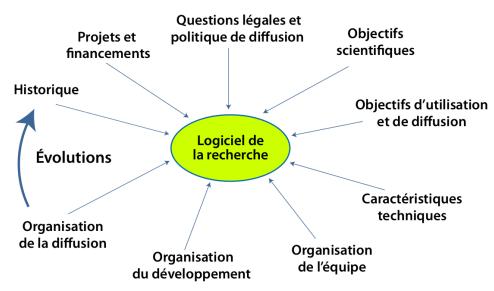
- les nouvelles informations,
- les nouvelles actions à entreprendre.

Conserver les versions successives!

Le cycle de vie de votre logiciel est une base pour définir vos étapes de mise à jour.

Adaptez et appropriez-vous le formulaire.

#### Modèle PRESOFT de SMP





# Le modèle PRESOFT : document texte et DMP OPIDoR

- Tient compte du retour des chercheurs du LIGM
- Modèle sous forme de document texte
  - http://www.france-grilles.fr/presoft/ et http://www.france-grilles.fr/presoft-en/
  - une version unique bilingue aux formats .odt et .pdf
  - accessible à tous (licence CC-BY 4.0).

### Modèle dans DMP OPIDoR :

- https://dmp.opidor.fr/
- une version en français et une version en anglais
- accessible à tout utilisateur de DMP OPIDoR (communauté enseignement supérieur et recherche) pour créer un plan et le partager avec ses collaborateurs et partenaires.





# Utilisation du modèle : comparatif

	Modèle texte	Modèle dans
Forme	odt et pdf (plusieurs tableaux, plusieurs questions par tableau)	formulaire en ligne (plusieurs thèmes, plusieurs questions par thème)  MAJ permanente possible
Langues	FR et EN dans le même document	FR ou EN (choix à l'initialisation du plan)
Gestion des versions	manuelle	manuelle (pour le moment - export puis stockage externe)
Ecriture collaborative	manuelle et séquentielle (un seul document)	différents rôles et droits d'accès possibles (dynamique)
Publication du plan	manuelle	export partiel possible
Accès	tout public (téléchargement)	<ul> <li>création de plan : communauté ESR</li> <li>partage possible d'un plan avec tout partenaire/collaborateur</li> </ul>



# Les apports du plan de gestion de logiciel



# Les apports du plan de gestion de logiciel

### Un document de référence :

- Centralise les informations concernant le logiciel,
- Aide à la reflexion et à la planification (actions, échéances, périodes critiques...).
- Outil pour mieux gérer la production scientifique à moyen & long terme, en période de contraintes budgétaires fortes,
- Support d'une démarche dans laquelle se poser les questions est finalement le plus important.

### Au niveau équipe :

- Vue d'ensemble de chaque logiciel,
- Suivi facile dans l'équipe,
- Amélioration des méthodes de gestion (gain de temps),
- Identification de ce qui marche bien,
- Identification des risques & opportunités
- Mise en œuvre d'une stratégie.

#### Au niveau laboratoire ou institution :

- Vue d'ensemble & meilleure connaissance des productions scientifiques,
- Identification des productions importantes,
- Mise en œuvre d'une politique.
- Collection utile pour l'évaluation de l'unité.
- Possibilité de plan de gestion des logiciels d'un laboratoire pour l'ensemble de ses logiciels.



# Les apports à la préservation

# Connaissance fine du logiciel de la recherche sur la totalité de sa durée de vie et l'ensemble de ses aspects :

 historique, projets et financements successifs, propriété intellectuelle, objectifs scientifiques, utilisation, caractéristiques techniques, documentation, gouvernance, organisation, coûts, organisation du développement et de la diffusion éventuelle...

### Aide:

- à la traçabilité des évolutions du logiciel, à la reproductibilité, au suivi des dépendances, à la prévention de l'obsolescence, à l'identification des compétences nécessaires et au coût de la préservation.
- à l'identification des risques,
- o à la décision : Quel logiciel préserver ? Combien de temps ? A quelles conditions ?

### Au niveau laboratoire

- Contribue à l'élaboration d'une stratégie logiciel au sein du laboratoire.
- Permet une meilleure gestion des actifs logiciels internes.
- Permet une meilleure connaissance des compétences et savoir-faire internes.



# Difficultés principales

- Peu d'incitations, peu de demandes dans les appels à projets ou évaluations, peu de leviers :
  - la démarche reste souvent volontaire.
- Impression pour les développeurs que cela n'apporte rien :
  - « c'est dans github ».
- Intérêt mais manque de temps.
- Appuis possibles :
  - la démarche qualité,
  - la réutilisation des données, la reproductibilité,



### Et maintenant?

- Le projet IN2P3 est terminé.
- Le modèle reste disponible :
  - sur DMP OPIDoR (Merci à l'équipe OPIDoR de l'INIST!)
  - sur <a href="http://www.france-grilles.fr/presoft/">http://www.france-grilles.fr/presoft/</a>
  - dans hal: <a href="https://hal.archives-ouvertes.fr/hal-01802565">https://hal.archives-ouvertes.fr/hal-01802565</a>
- Pour plus d'informations : <u>presoft@cc.in2p3.fr</u>



Merci de votre attention!

Questions?

Nous contacter : <a href="mailto:presoft@cc.in2p3.fr">presoft@cc.in2p3.fr</a>