



HAL
open science

Partage de pratiques et outils pour le calcul scientifique

Marie-Josée Cros, Claire Chabanet, Fabrice Dessaint, David Legland, Serguei Sokol, Herve Richard

► **To cite this version:**

Marie-Josée Cros, Claire Chabanet, Fabrice Dessaint, David Legland, Serguei Sokol, et al.. Partage de pratiques et outils pour le calcul scientifique. Journées Math-Info de l'INRA, Oct 2016, Mallemort, France. hal-01886412

HAL Id: hal-01886412

<https://hal.science/hal-01886412>

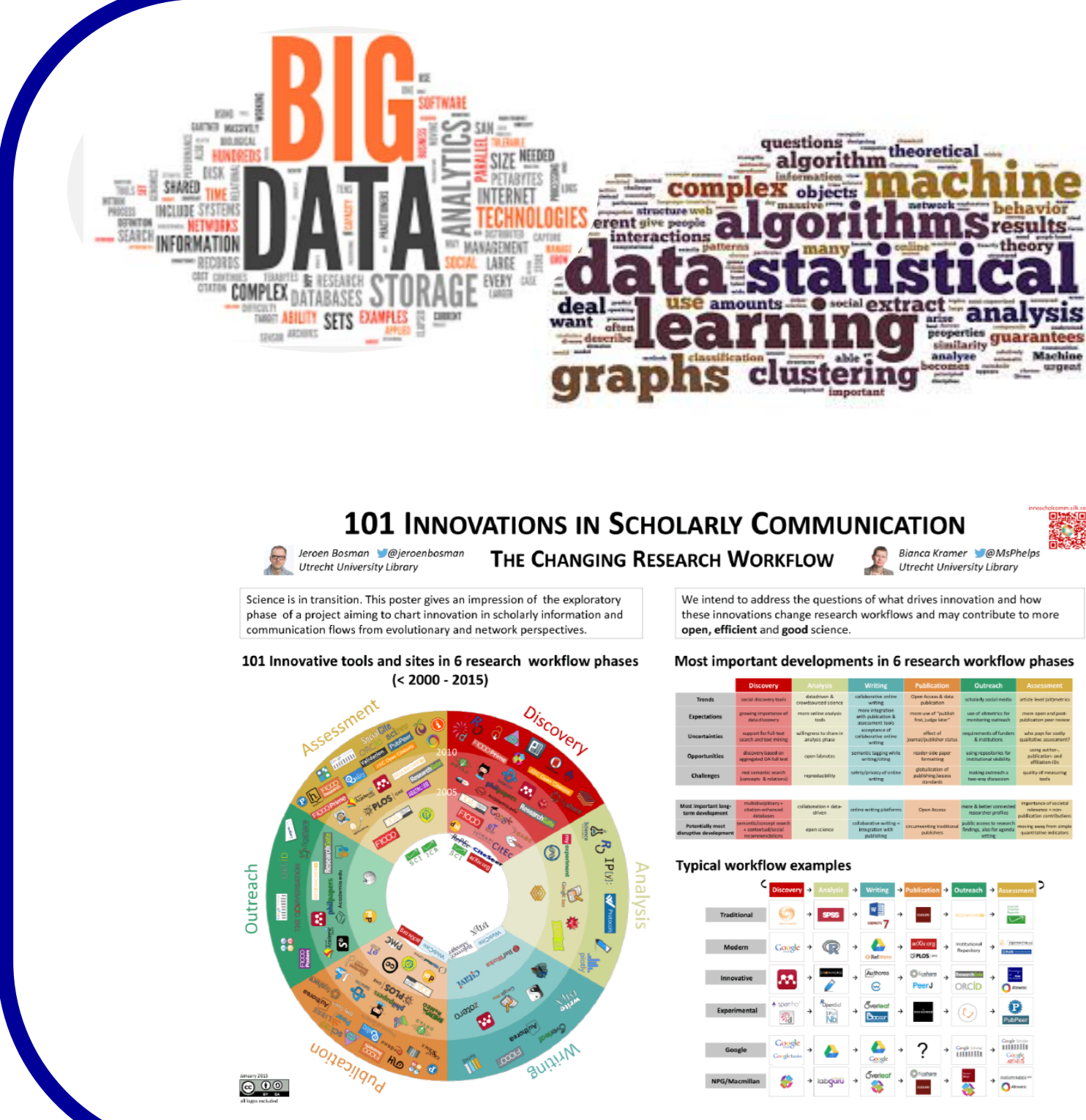
Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Partage de pratiques et outils pour le calcul scientifique

Groupe Pratiques&Outils du CATI CaSciSDI de l'INRA
pratiques-outils-calcul@listes.inra.fr



CONTEXTE

- Les scientifiques utilisent de plus en plus le calcul scientifique.
- Tous n'ont pas été formés.
- Les pratiques et les outils évoluent.

ENJEUX

Aider les permanents et les non-permanents de l'INRA et de ses partenaires à s'informer sur, se former à, adapter, améliorer les pratiques et outils mis en œuvre pour le calcul scientifique.

Ceci afin d'améliorer la qualité (fiabilité, robustesse, traçabilité...) et la reproductibilité des calculs (effet cumulatif des recherches).

Un site web participatif pour partager des pratiques, des liens, des présentations, des introductions aux outils, un état de l'art ...

<http://wiki.inra.fr/wiki/cascisdi/>

(Le dernier '/' doit être tapé pour un accès sans authentification)

Pratiques et outils pour le calcul scientifique
Modifié par Marie-Josée CROS le 2016/09/21 09:31

Ce wiki est réalisé par le CATI CaSciSDI de l'Inra. [poster de présentation du wiki]

Pour faire du calcul scientifique de qualité et reproductible, il est nécessaire d'associer aux résultats de calcul tout ce qui est nécessaire pour reproduire les résultats :

- A - Les données
 - 1. Planifier la gestion de données
 - 2. Gérer les données brutes
 - 3. Prétraitements
 - 4. Gérer les données utilisées
 - 5. Préserver les données
 - 6. Partager les données
- B - Les codes de calcul
 - 1. Ecrire les programmes pour des hommes pas pour des machines
 - 2. Laisser l'ordinateur faire le travail
 - 3. Faire des changements incrémentaux
 - 4. Ne pas (se) répéter
 - 5. Prévoir des erreurs
 - 6. Optimiser le code uniquement après qu'il ait marché correctement
 - 7. Documenter l'objectif et la conception, pas la mécanique
 - 8. Collaborer
- C - Processus et environnement
 - 1. Enregistrer un processus
 - 2. Conserver l'environnement
- D - Calcul

Articles liés : bonnes pratiques, Calcul matriciel, Conditionnement, conventions, codage, données, extrêmes, Emacs, gestion de versions, gestionnaire de version, git, Matlab, Matrices de Hilbert, Org, outlier, planification, programmation, orientée objet, puissance, reproductibilité, régression.

Articles recommandés :

- Déterminer la taille des échantillons notion sous-jacente : puissance d'un test (Claire Chabanet, Fabrice Dessaint, 10 mai 2016)
- Conseils pour la saisie des données (Modifié par Claire CHABANET le 2016/03/14 19:02)
- Outliers ou données extrêmes, comment les détecter? que faire de ces observations? (Claire Chabanet, Fabrice Dessaint, 10 juillet 2015)
- Conventions de codage pour Matlab (Modifié par David LEGLAND le 2015/07/17 10:40)
- Programmation orientée objet sous Matlab (Modifié par Marie-Josée CROS le 2015/09/13 13:48)
- Utiliser un gestionnaire de version (Modifié par Gwen MORIN le 2016/04/07 10:55)
- Versions avec Git (Modifié par Marie-Josée CROS le 2015/07/19 11:43)
- Gérer des versions avec Git (Modifié par Marie-Josée CROS le 2015/07/19 11:43)
- 1 Utilisation de Git sous RStudio (Fabrice Dessaint, 13 mai 2015, Version 4)
- Conditionnement d'une matrice (Modifié par Sylvain ASSON le 2015/07/20 11:03)
- Régression linéaire sous Matlab (Modifié par David LEGLAND le 2015/07/17 10:09)
- Recherche Reproductible en calcul scientifique (Modifié par Marie-Josée CROS le 2016/03/25 12:00)
- Plateformes de partage (Modifié par Marie-Josée CROS le 2016/09/20 15:38)
- Automatiser les tests (Modifié par Marie-Josée CROS le 2016/03/28 15:00)
- Tests unitaires sous R (Modifié par Marie-Josée CROS le 2015/07/19 11:43)
- EditeursEnLigne (Modifié par Marie-Josée CROS le 2015/07/19 11:43)
- Org-mode Emacs (Modifié par Marie-Josée CROS le 2016/01/09 16:26)

Ressources extérieures à ce wiki :

- Plateforme Logicielle pour l'Analyse des Signaux et leur Traitement, l'Intégration de données et des Connaissances (PLASTIC) du département Cépia de l'Inra
- Site sur connaissances et pratiques de l'environnement R (R4ciam) du cati CaSciSDI de l'Inra
- Site gestion et partage des données scientifiques de l'Inra

Le site a été initié en 2015 par 8 personnes des départements ALIMH, BAP, CEPIA, MIA, MICA, SPE. Il est ouvert en lecture à tous et pour contributions / commentaires aux personnes ayant un ldap INRA. Le site évolue principalement par itérations lors de périodes définies par le groupe et dans le cadre d'un CATI INRA.