

Proposition de Journal : SW and Computing for Big Science

Michel Jouvin

CNRS/IN2P3/LAL

jouvin@lal.in2p3.fr

Disclaimer/Crédit

- Je ne suis pas un des initiateurs du projet de journal “SW&C for Big Science”
- Idée proposée par Gunter Quast (CMS, KIT, Karlsruhe) en 2013
 - Conférence de physique des particules à Main
 - Deux années de discussions ont fait évoluer l’idée initiale vers la proposition actuelle de journal à propos de SW&C for Big Science
- L’essentiel des informations contenues dans ce transparent ont été empruntées à une présentation de Markus Elsing (CERN) au CERN en Juin
 - <https://indico.cern.ch/event/394783/contributions/2188799/attachments/1287140/1915139/GDBTalk.pdf>

Motivations

- Plusieurs communautés scientifiques produisent et analysent de grands volumes de données (PB et plus), à travers de grandes collaborations (« data-intensive sciences »)
 - Physique des particules (LHC) est l'exemple actuel le plus significatif mais des choses similaires arrivent en astrophysique, physique des particules, photon science, biologie...
- Beaucoup de challenges technologiques communs
 - Partager les expériences et savoir-faire évite de réinventer la roue
 - Peut être un catalyseur de la collaboration entre disciplines
- Développement d'infrastructure/framework informatique pour traiter le problème de la distribution des données et de leur traitement ainsi que la collaboration
 - Souvent peu (re)connus à l'extérieur de la communauté qui les a développés

Challenge de la Publication

- Peu/pas d'espace pour publier à propos des solutions à ces challenges
 - Trop informatique pour des publications liées à la science effectuée
 - Trop « ingénierie » pour des publications de recherche en informatique
 - Les publications dispersées dans différentes revues : difficile de les identifier
 - HEP : principalement les proceedings de CHEP (conférence tous les 18 mois)...
- Importance d'un journal reconnu au contenu validé
 - Journal à referee
 - Indexation, impact factor
 - Être une archive centrale pour partager l'expertise sur les outils pour les data intensive sciences
 - Participer à la reconnaissance et à la carrière « des Research SW Engineers »
 - Important pour les jeunes qui s'investissent dans ces problématiques

Le Projet

- Un journal en ligne, avec une publication continue, en partenariat avec Springer
 - Christian Caron (Executive Publishing Editor @Springer) très tôt associé au projet, fortement motivé et un de ses porteurs actifs !
 - Un business model hybride plutôt que pure Open Access
 - Trop tard pour son inclusion dans SCOAP3 phase 2 (2017-2019) mais prévoit d'y candidater pour les phases suivantes
 - Modèle hybride (si j'ai bien compris!) : open access pour les financeurs..
- Types d'articles : pas de lettre ou proceeding mais éventuellement des revues ou des « advanced tutorials »
 - Ne pas faire de remplissage avec des choses qu'on trouve ailleurs...

Sujets Possibles

- Infrastructures for large-scale, high-throughput computing
- Middleware development
- Data processing, hosting and sharing
- Distributed data analysis
- Software development Infrastructures
- Software benchmarking and Performance Assessment
- Frameworks and software integration
- Novel algorithms for efficient data reconstruction and filtering
- Deep learning algorithms
- Event and object classification
- Online/Offline data quality monitoring
- Data visualisation

Structures et Participation Française

- Scientific Advisory Committee
 - « Political support »
 - Des scientifiques reconnus dans leur discipline
 - Ex: E. Elsen, directeur de la recherche du CERN...
- Editorial Board: ceux qui « font » le journal
 - « Topical Editors » par discipline ou type de sujet
- Auteurs... sans eux pas de journal
 - Importance de faire connaître le journal et d'inciter les chercheurs/ingénieurs à y publier
 - Très en ligne avec la politique IN2P3 de pousser les ingénieurs SW&C à publier
- IN2P3 et CNRS : une opportunité pour ouvrir à toutes les communautés concernées