

OpenTapioca et ISTEK



Antonin Delpuch

Journée ISTEK 2021

OpenTapioca en bref

- Annotation des entités nommées dans des textes
- Utilise la base de connaissances Wikidata



Un peu de contexte: Dissemin

En 2016, adoption d'une politique pour l'accès ouvert (AO) à l'École normale supérieure.



Résolution votée par le Conseil scientifique de l'École normale supérieure sur l'accès libre aux publications

17 juin 2016

L'École normale supérieure est attachée à la libre diffusion des fruits de la recherche qu'elle héberge, car ils constituent un bien commun.

Le Conseil scientifique de l'ENS demande donc aux chercheurs de l'ENS de s'assurer que le texte intégral de leurs publications soit désormais disponible dans une archive ouverte publique, au plus tard 6 mois après publication dans le domaine des sciences, de la technique

Une plateforme pour aider la mise en place de politiques AO

- Recenser les papiers des chercheurs
- Analyser leur disponibilité dans les archives ouvertes
- Proposer un dépôt simplifié des articles manquants

Politiques d'auto-archivage

Importées depuis SHERPA/RoMEO



Full text: Unavailable

 [Preprint](#): archiving allowed [Upload](#)

 [Postprint](#): archiving allowed [Upload](#)

 [Published version](#): archiving forbidden

[Policy details](#) Data provided by  SHERPA/RoMEO

Quid des politiques des financeurs et des employeurs ?

Lier les articles aux institutions

Quelles sources de données ouvertes ?

- Crossref/Fundref : couverture faible
- HAL : beaucoup de doublons dans l'ensemble d'institutions
- Extraction à partir d'articles ?

Generalized hydrodynamics in complete box-ball system for
 $U_q(\widehat{sl}_n)$

Atsuo Kuniba¹, Grégoire Misguich², Vincent Pasquier²

¹ Institute of Physics, University of Tokyo, Komaba, Tokyo 153-8902, Japan

² Institut de Physique Théorique, Université Paris Saclay, CEA, CNRS UMR 3681, 91191 Gif-sur-Yvette, France

atsuo.s.kuniba@gmail.com, gregoire.misguich@ipht.fr, vincent.pasquier@ipht.fr

L'API ISTE

```
author : [
  {
    "name": "Qiu hao Qu",
    "affiliations": [
      "Department of Neurosciences, Center for Gene Expression and Drug
Discovery, 1500 E. Duarte Road, Duarte, California 91010, USA.",
      "These authors contributed equally to this work."
    ]
  },
  {
    "name": "Guo qiang Sun",
    "affiliations": [
      "Department of Neurosciences, Center for Gene Expression and Drug
Discovery, 1500 E. Duarte Road, Duarte, California 91010, USA."
    ]
  }
]
```

Métadonnées extraites des PDF disponibles sous
License Ouverte

Annotation du jeu de données

- Alignement manuel de 1000 affiliations à Wikidata
- Représentation en format RDF/NIF

Department of Electrical Engineering and Computer Science, The
Laboratory for Information and Decision Systems, [Massachusetts
Institute of Technology], [USA]

[United States of America \(Q30\)](#)
sovereign state in North America
Rank: 17.42, phrase: 10.23
Statements: 1423, sitelinks: 377
Score: 0.979183912732802

[Union of South Africa \(Q193619\)](#)
state in southern Africa from 1910 to
1961, predecessor to the Republic of
South Africa
Rank: 10.10, phrase: 10.23
Statements: 47, sitelinks: 55