

# Les plateformes des données ouvertes

15 Décembre 2021 – Séance n°4



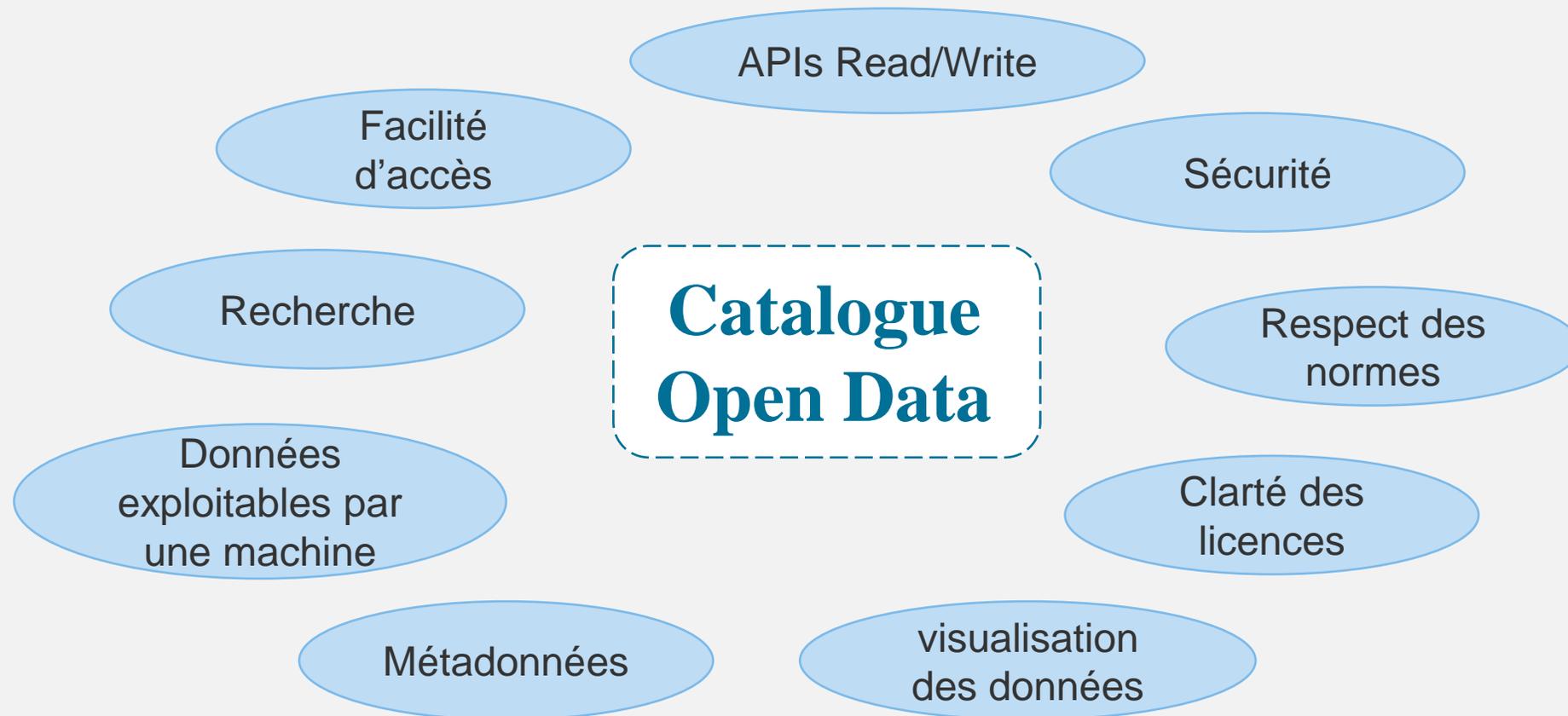
UNITED NATIONS

الاسواق  
ESCWA

Shared Prosperity **Dignified Life**



# Les plateformes Open Data



# Les plateformes Open Data

## Deux types ?

- Catalogue de données,
- Plateforme des données ouvertes,

## Trois modèles ?

- Catalogue unique,
- Catalogues décentralisés,
- Catalogues fédérés,

# Les plateformes Open Data

## Catalogue unique :

Une plate-forme nationale - initiative centralisée

(+)

Les jeux de données sont lisibles au niveau national (un seul et unique point d'entrée)

Une seule technologie / Une seule Documentation / une seule API

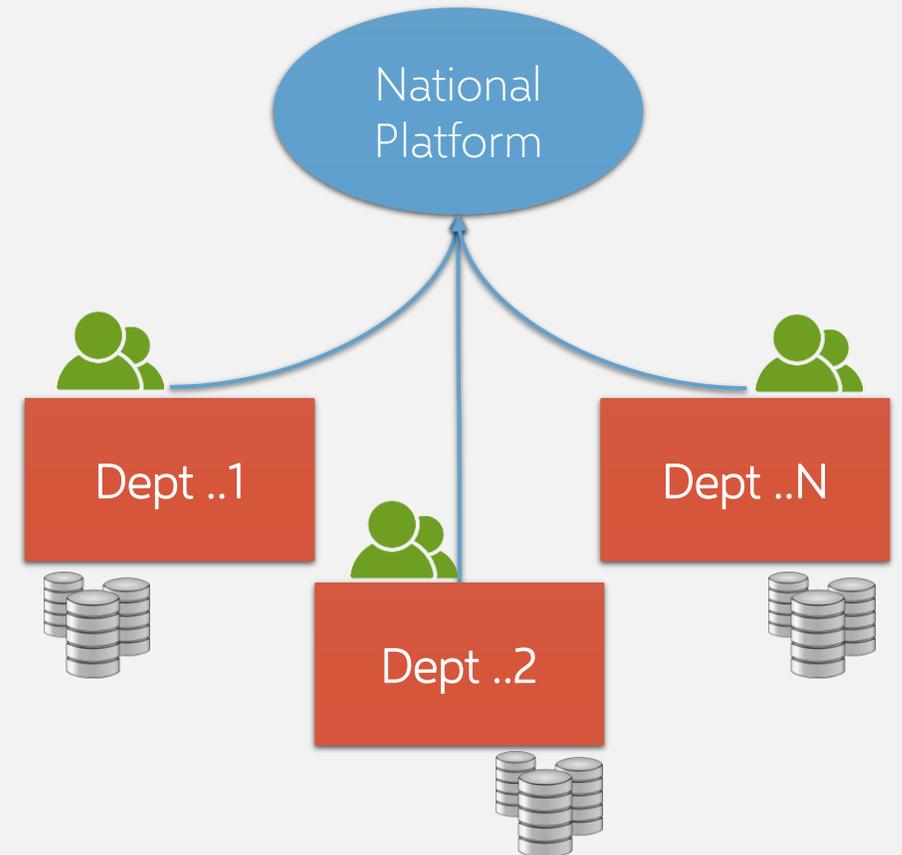
Facilité de recherche et de réutilisation

(-)

Spécificité des autres secteurs

Exigences techniques élevées

Ne favorisant pas l'innovation technologique



# Les plateformes Open Data

## Catalogues décentralisés :

Plateformes sectorielles distinctes - Décentralisées

(+)

Prise en charge de la spécificité sectorielle

Favoriser l'innovation technologique

Favoriser la compétitivité entre secteur

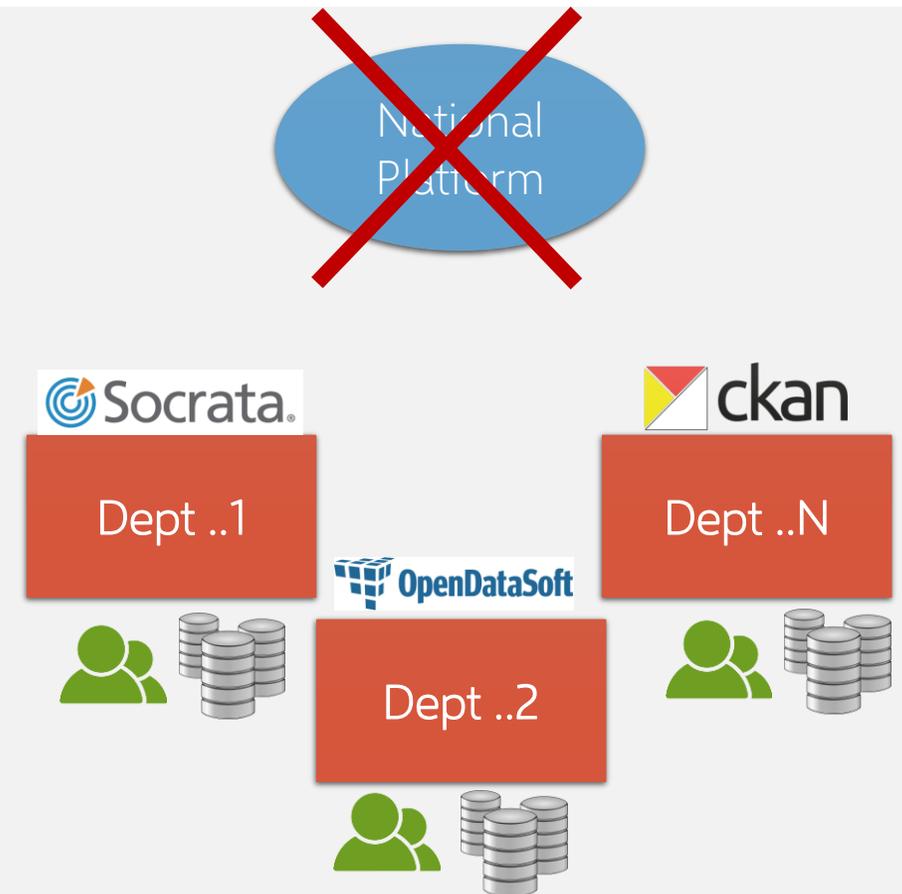
(-)

Absence d'une vision unifiée de l'initiative nationale

Technologies multiples

Difficulté d'accès et de découvertes des données

Plusieurs documentations & API



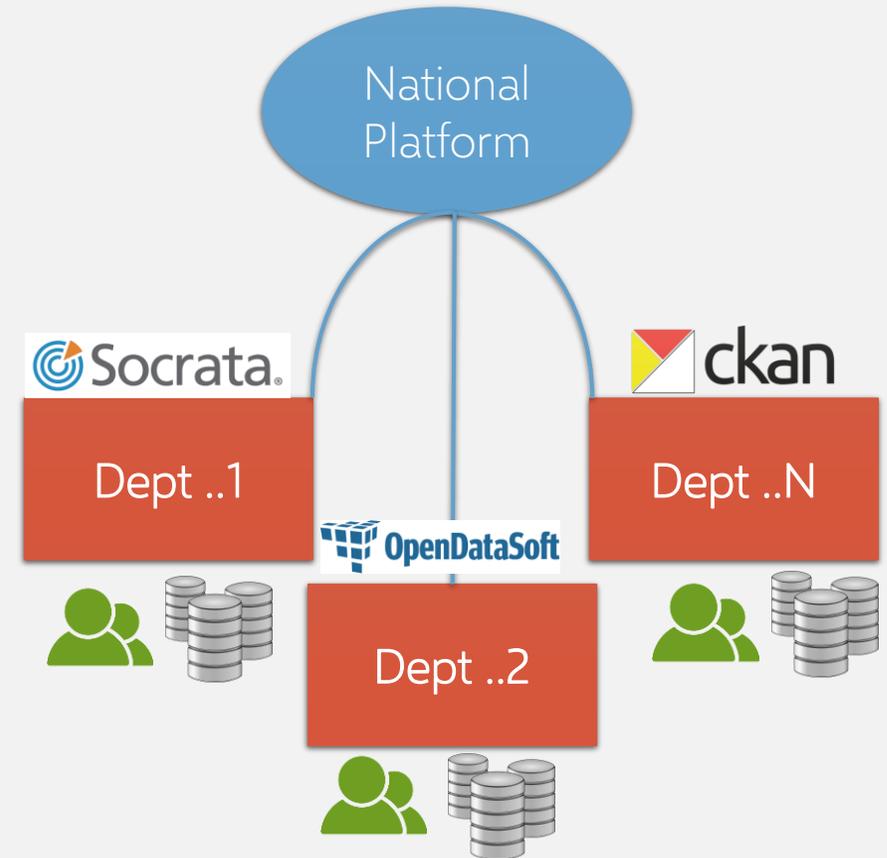
# Les plateformes Open Data

## Catalogue fédéré :

Une plateforme nationale fédérée et plusieurs plateformes sectorielles et locales.

Mettre en place un model commun des métadonnées

Un Moissonnage automatique pour le référencement des jeux de données eu niveau de la plateforme nationale



# Les plateformes Open Data

## Quel catalogue choisir ?

- ✓ L'acteur public doit disposer des ressources humaines et financières nécessaires à la mise en place et à la gestion d'une plateforme open data,
- ✓ La plateforme doit garantir le référencement sur la plateforme nationale,
- ✓ Il est recommandé d'utiliser une solution open source,
- ✓ La plateforme doit proposer des mécanismes de participation et d'interaction avec les utilisateurs.

## Les mécanismes de participation ?

- Une rubrique pour les demandes d'ouverture des nouveaux jeux de données,
- Une fonctionnalité pour signaler un jeu de données
- Les commentaires pour chaque jeu de données,
- Une rubrique pour les cas de réutilisation des données ouvertes,
- Une rubrique pour les propositions d'amélioration de la plateforme open data,

# CKAN

**CKAN** est un catalogue de données open source le plus utilisé par les gouvernements. Cette solution peut être installée sur un serveur Linux.

Le catalogue CKAN est conçu pour publier et gérer des données au moyen d'une interface IHM ou d'une API (write).

Code en Python pour le back-end et de JavaScript pour la partie interface.

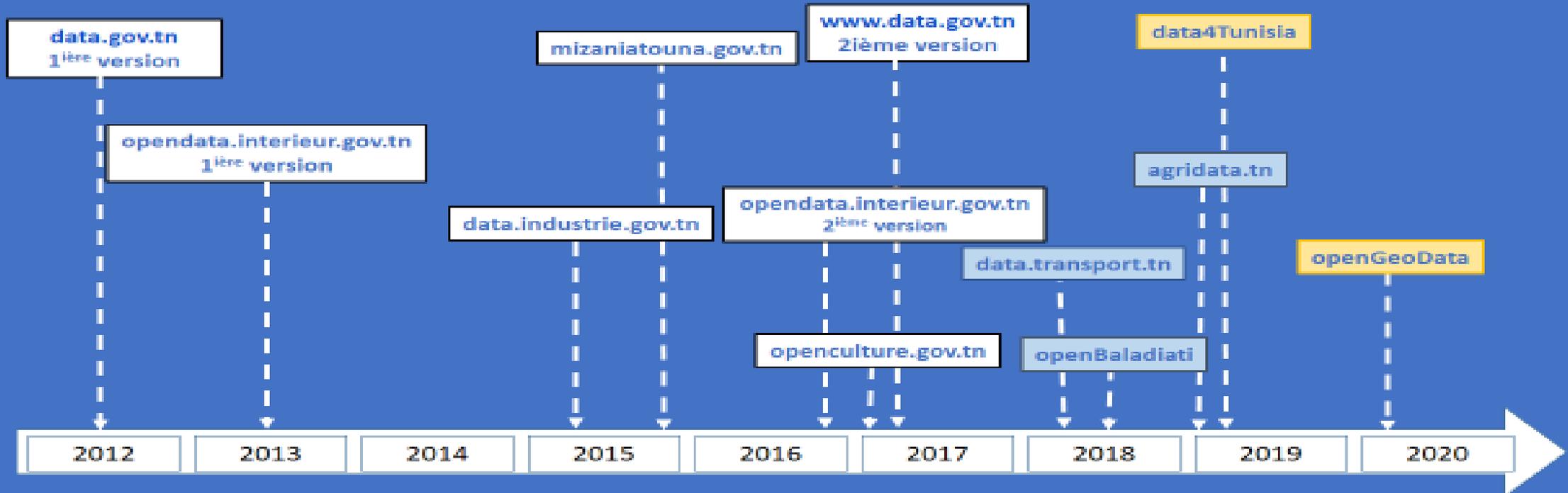
La mise en place de catalogue de données ouvertes à base de CKAN peut être rapide mais la personnalisation et l'optimisation demandent plus d'investissement qu'une solution clé en main.

Existence d'une communauté de développeurs conséquente assure une évolution constante de la technologie et des outils.

En dehors des outils de recherche, de visualisation, de cartographie classique, et d'interrogation API, CKAN propose aussi des extensions externes accessibles sur ce site : <https://github.com/ckan/ckan/wiki/List-of-extensions>

# CKAN

## Les portails open data en Tunisie



# DKAN

**DKAN** est une variante de CKAN qui repose sur DRUPAL , un système de gestion de contenu très populaire écrit en PHP au lieu de Python.

DRUPAL, possède sa propre architecture modulaire, avec des milliers de modules disponibles en téléchargement. Ce logiciel permet également de personnaliser les modules et possède une grande communauté de développeurs.

# uDATA

**uDATA** c'est la plateforme développée par EtatLab l'agence de service numérique de l'Etat français. Elle est développée en Python et facilement adaptable au besoin.

Pour faire tourner une plateforme uData en mode standalone, vous avez besoin d'un simple serveur linux ou MacOS et de suivre la documentation en ligne. Par la suite, il faut juste travailler sur la personnalisation. Il est aussi possible de déployer la plateforme avec Docker.

Le code source est accessible sur : <https://github.com/opendatateam/udata>

Exemples : <https://www.data.gouv.fr> ; <https://data.public.lu> ; <https://dados.gov.pt>

# GeoNode

**GeoNode** est une plateforme open source pour développer des systèmes d'informations géospatiales (GIS) et pour déployer des infrastructures de données géo-spatiales.

Cette plateforme a été conçue pour être facilement extensible et modifiable, et peut être intégrée dans une autre plateforme existante.

Exemples : <https://geonode.wfp.org/>



Shared Prosperity **Dignified Life**



# Le catalogue CKAN

# CKAN et les extensions

<https://github.com/open-data/ckanext-scheming>

Scheming est une extension permettant de personnaliser des schémas des jeux de données, des groupes et d'organisation dans CKAN.

L'extension des schémas se fait à l'aide de fichiers JSON ou YAML qui incluent des règles de validation, la modification d'extraits de modèle de formulaire et des extraits d'affichage.

# CKAN et les extensions

<https://github.com/open-data/ckanext-fluent>

Fluent est une extension permettant d'ajouter du texte, des démarques et des balises multilingues aux jeux de données, aux groupes et aux schémas d'organisation de CKAN.

Fluent fonctionne avec Scheming en fournissant des validateurs et des extraits de code pouvant être utilisés dans les schémas personnalisés.

Fluent affiche des zones d'édition pour toutes les langues pour chaque champ dans le formulaire d'édition, mais par défaut n'affiche que la valeur de la langue de l'utilisateur lors de l'affichage d'un jeu de données.

Toutes les valeurs multilingues sont renvoyées et créées/mises à jour via l'API à l'aide d'un objet JSON.

Par exemple. un champ Fluent "label" peut avoir la valeur :

```
... "label": { "ar": "كتب", "fr": "Livres"}, ...
```

# CKAN et les extensions

## DataStore + Data Pusher / xloader

L'extension CKAN DataStore fournit une base de données ad hoc pour le stockage de données structurées à partir de distributions d'un jeu de données CKAN. Les données peuvent être extraites des fichiers et stockées dans le DataStore.

Lorsqu'une distribution est ajoutée au DataStore, on obtient :

- Aperçu automatique des données sur la page de la distribution, à l'aide de l'extension Data Explorer
- API Data : permettant de chercher, filtrer et mettre à jour les données, sans avoir à télécharger et uploader l'intégralité du fichier de données

Le DataStore est intégré à l'API CKAN et au système d'autorisation.

➔ Le DataStore est généralement utilisé avec le DataPusher, qui téléchargera automatiquement les données vers le DataStore à partir de fichiers appropriés, qu'ils soient téléchargés via le FileStore de CKAN ou via des liens .

# CKAN et les extensions

## Les autres extensions :

webpage\_view / pdf\_view / officedocs\_view / basic-charts

Spatial\_metadata spatial\_query navigablemap geojson\_view

Harvest / dcat

Pages / googleanalytics / showcase / datarequests / disqus / rating / contact

# CKAN et les extensions

**La Version 2.9.X** : migration de Python 2.X à Python 3.X

Plusieurs extensions ont été migrées vers python 3

D'autres sont en phase de progression de migration

<https://github.com/ckan/ckan/wiki/Python-3-migration-guide-for-extensions>

Extension	Python 3 support
<a href="#">ckanext-scheming</a>	✓
<a href="#">ckanext-harvest</a>	✓
<a href="#">ckanext-dcat</a>	✓
<a href="#">ckanext-xloader</a>	✓
<a href="#">ckanext-spatial</a>	✓ :
<a href="#">ckanext-pages</a>	✓
<a href="#">ckanext-geoview</a>	✓
<a href="#">ckanext-googleanalytics</a>	✓
<a href="#">ckanext-qa</a>	☹️
<a href="#">ckanext-showcase</a>	✓
<a href="#">ckanext-pdfview</a>	✓
<a href="#">ckanext-hierarchy</a>	☹️
<a href="#">ckanext-archiver</a>	✓
<a href="#">ckanext-fluent</a>	✓



Shared Prosperity **Dignified Life**



# Discussion