

Notice pour les jeux de données ANR DGPIE

Avril 2021

Date de mise à jour des données

Ces tables seront mises à jour au moins deux fois par an et idéalement une fois par mois par l'ANR.

Contenu

Ces jeux de données listent les projets gérés par la Direction des grands programmes d'investissement de l'État (DGPIE) depuis 2010. Ces projets sont financés dans le cadre des Programmes d'investissement d'avenir (PIA). Seuls les projets financés sont mentionnés.

Ces jeux de données ne contiennent pas les projets gérés par la Direction des opérations scientifiques (DOS) et financés par l'Agence depuis 2005. Ceux-ci sont listés dans des jeux de données spécifiques.

Le montant indiqué est le montant de l'aide allouée au niveau du projet.

Les données générales des projets financés sont dans le jeu de données "Projets". Les données concernant les partenaires de ces projets sont dans le jeu de données "Partenaires". La clé de jointure entre les deux jeux de données est le champ "Projet.Code_Decision_ANR".

Organisation

Les données proposées couvrent la période 2010-2021.

Les données sont continuellement mises à jour et augmentées. Elles sont directement extraites des bases internes de l'Agence.

Les deux noms des fichiers sont les suivants, où `yyyymmdd` est la date d'extraction.

- Projets : `anr-dgpie-2010-2021-projets-finances-yyyymmdd-projets`
- Partenaires : `anr-dgpie-2010-2021-projets-finances-yyyymmdd-partenaires`

Formats

Chaque jeu de données est proposé en 3 formats différents.

Format XLSX Excel 2007

Le fichier est produit par la routine `"pandas.DataFrame.to_excel"` de Python3/Pandas, avec l'argument `engine="xlsxwriter"`.

Documentation : https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.to_excel.html.

Format JSON

Le fichier est produit par la routine "pandas.DataFrame.to_json" de Python3/Pandas, avec l'argument "orient=split" pour privilégier la rapidité de lecture.

Documentation : https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.to_json.html.

Format CSV

Le fichier est produit par la routine "pandas.DataFrame.to_csv" de Python3/Pandas. Les arguments suivants sont utilisés pour la production du fichier CSV. Notez que le séparateur est ";" et que toutes les chaînes non numériques sont protégées par des "double quotes".

- encoding="utf-8"
- sep=";"
- quoting=csv.QUOTE_NONNUMERIC
- doublequote=True

Documentation : https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.to_csv.html.

Champs du jeu de données Projets

Nom du champ	Définition	Notes et informations complémentaires	Exemple
Projet.Code_Decision_ANR	Code unique d'identification du projet financé par l'ANR	Ce code est de la forme "ANR-nn-xxxx-nnnn", où n est un chiffre et x est un caractère imprimable. Pour quelques projets membres d'actions divisées en sous-actions, il est de la forme "ANR-nn-xxxx-nnnn-mm"	ANR-19-P3IA-0001 ANR-10-CAMP-0128-01
Projet.Date_debut	Date de début du projet	Vide si la date n'est pas connue	2019-09-01 00:00:00
Action.Edition	Année d'édition de l'appel à projet pour l'action		2019
Action.Acronyme	Acronyme de l'action		P3IA
Projet.Acronyme	Acronyme du projet déterminé par le porteur lors du dépôt de son projet		PRAIRIE
Projet.Titre.Francais	Titre français du projet déterminé par le porteur au moment du dépôt		PaRis Artificial Intelligence Research InstitutE
Projet.Resume.Francais	Résumé français du projet déterminé par le porteur au moment du dépôt	Vide si pas de résumé	(Facultatif)
Projet.Aide_allouee	Aide totale allouée au projet au moment du conventionnement. Montant en euros	0 si non connu	21500000

Champs du jeu de données Partenaires

Projet.Code_Decision_ANR	Code unique d'identification du projet financé par l'ANR	Ce code est de la forme "ANR-nn-xxxx-nnnn", où n est un chiffre et x est un caractère imprimable. Pour rojets membres d'actions divisées en sous-actions, il est de la forme "ANR-nn-xxxx-nnnn-nn"	ANR-19-P3IA-0001 ANR-10-CAMP-0128-01
Projet.Acronyme	Acronyme du projet déterminé par le porteur lors du dépôt de son projet		PRAIRIE
Projet.Partenaire.Code_Decision_ANR	Code du partenaire au sein du projet	Un partenaire est un organisme officiel associé au projet. Chaque partenaire se voit attribuer un numéro au moment du conventionnement. Ce numéro va de 01 au nombre de partenaires. Le partenaire mm du projet "ANR-nn-xxxx-nnnn" a pour code "ANR-nn-xxxx-nnnn-mm". Cependant, cette règle n'est pas suivie pour les projets de la forme "ANR-nn-xxxx-nnnn-nn". Les partenaires ont des codes de la forme "ANR-nn-xxxx-nnnn-nn" ou "ANR-nn-xxxx-nnnn-nn-nn"	ANR-19-P3IA-0001-01
Projet.Partenaire.Est_coordinateur	Est-ce que le partenaire est le coordinateur du projet? Ce champ est un booléen	Il y a un et un seul établissement coordinateur par projet	VRAI
Projet.Partenaire.Nom_organisme	Nom de l'organisme partenaire déclaré par le coordinateur au moment du dépôt et mis à jour durant le suivi du projet en cas de modification		INRIA siège
Projet.Partenaire.Categorie_organisme	Statut du partenaire tel que renseigné par le coordinateur au moment du dépôt		Organismes de type EPST
Projet.Partenaire.Responsable_scientifique.Nom	Nom de la personne responsable scientifique du projet		PONCE

Projet.Partenaire.Responsable_scientifique.Prenom	Prénom de la personne responsable scientifique du projet		Jean
Projet.Partenaire.Adresse.Ville	Ville de réalisation des travaux du projet		Le Chesnay CEDEX
Projet.Partenaire.Adresse.Region	Région de réalisation des travaux du projet		Île-de-France
Projet.Partenaire.Adresse.Pays	Pays de réalisation des travaux du projet		France