



PRÉFÈTE
DE LA RÉGION
AUVERGNE-
RHÔNE-ALPES

Liberté
Égalité
Fraternité

DATAculturation

Sensibiliser à la donnée,
à l'intelligence artificielle
et animer un café
de la data



SOMMAIRE

Le projet DATAcculturation	5
---	----------

PARTIE 1	9
-----------------------	----------

ORGANISER UN CAFÉ DE LA DATA OU DE L'IA

Les méthodes pour animer un café de la data ou de l'IA dans votre administration

Visualisez en un coup d'œil l'ensemble des formats d'ateliers qui s'offrent à vous en fonction du contexte dans lequel vous souhaitez organiser un café de la data ou de l'IA.

PARTIE 2	52
-----------------------	-----------

POUR ALLER PLUS LOIN

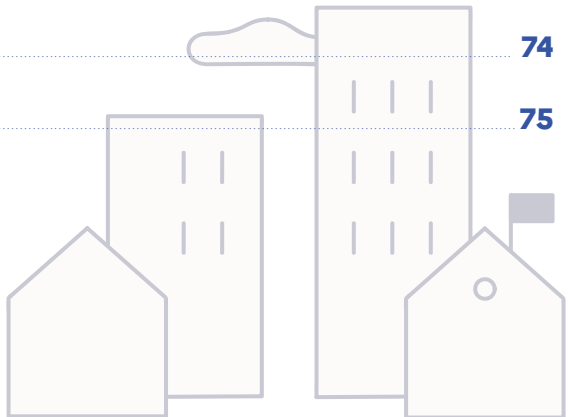
Sensibiliser à la donnée et à l'IA

Enrichissez vos connaissances sur la data et l'IA à partir de ressources complémentaires aux vidéos.

Pour aller plus loin : ressources complémentaires et bibliographie thématique.....	67
--	-----------

Lexique.....	74
--------------	-----------

Remerciements.....	75
--------------------	-----------










LE PROJET DATAACULTURATION

Les données nous entourent au quotidien, dans notre environnement professionnel et dans notre vie privée. Elles représentent un enjeu commun à toutes les administrations et concernent l'ensemble des agents publics. Les données sont également une brique essentielle de tout système d'intelligence artificielle.

DATAculturation a pour ambition d'acculturer les agents publics à la donnée et de leur permettre de développer des compétences transverses sur le sujet, transférables dans tous les environnements professionnels.

Pour cela, DATAculturation se décline via des supports simples d'utilisation et accessibles dans un kit constitué :

-  de **vidéos** présentant des concepts et exemples concrets liés à la donnée,
-  d'une méthodologie pour **organiser des cafés** de la data,
-  d'un **livret d'approfondissement** pour comprendre le sujet de la donnée.

Les vidéos définissent le sujet et abordent des exemples concrets dans nos environnements de travail et notre vie quotidienne : domaines RH, cartographie des territoires, mobilité et transports, données de santé, etc. Les exemples illustrent des environnements professionnels variés au sein des trois fonctions publiques. Ils sont compréhensibles et transposables, sur tous les postes. Le développement de nouvelles compétences sur le sujet de la donnée permet de s'adapter rapidement à tous les contextes et environnements professionnels.



Trois vidéos sont spécifiquement conçues autour de cas d'usage sur l'ouverture et la réutilisation des données afin de démontrer leur réel enjeu. Une autre vidéo (en 2 parties) est consacrée aux enjeux liés à l'utilisation de l'intelligence artificielle dans les services publics et à ses liens avec les données.

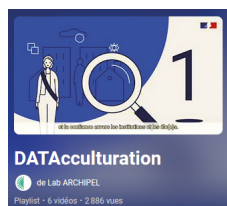
Les cafés de la data sont conçus comme des événements de sensibilisation organisés au sein des services, auprès des agents pour échanger sur la donnée, à partir des supports vidéos DATAcculturation. Ces événements pourront être animés dans les administrations en s'appuyant sur des agents et des réseaux professionnels mobilisables et volontaires.

Le sujet de la donnée vous intéresse mais :

- vous ne vous sentez pas légitime pour en parler avec vos collègues ?
- vous pensez ne pas avoir assez de connaissances ?

Le premier objectif de ce livret est de vous donner des clés de compréhension du sujet, de vous sensibiliser à la donnée à travers des supports simples d'utilisation et faciles d'accès.

ASTUCE : avant de poursuivre votre lecture, visionnez les vidéos et consultez les ressources en ligne !



Ce guide a été mis à jour en décembre 2024 !
Il a été enrichi de nouvelles formules pour animer d'autres cafés de la data, notamment à distance en visio et sur le sujet de l'intelligence artificielle.

CONTACT : justine.sperandio-martinez@auvergne-rhone-alpes.gouv.fr

COMMENT UTILISER CE GUIDE



Dans cette partie, vous accédez à des **déroulés d'animation clés en main** pour construire votre café de la data.

Retrouvez ci-après les déroulés d'animation des cafés de la data.

Dans cette partie, découvrez toutes les **ressources sur la donnée et l'IA** pour vous aider à améliorer vos connaissances sur ces sujets : articles, définitions, autres supports d'animation, jeux sérieux, etc.

Découvrez également plus d'informations sur les **vidéos DATAculturation**.



PARTIE 1



ORGANISER UN CAFÉ DE LA DATA OU DE L'IA

Les méthodes pour animer un café de la data ou de l'IA
dans votre administration





Accédez à des ressources dédiées à l'organisation d'un café de la data. Vous avez envie de partager votre curiosité et vos connaissances sur la data avec vos collègues ou dans votre service ? Les cafés de la data sont faits pour vous !

Ce livret permet de concevoir et d'animer des sessions de présentation participatives qui vous permettront d'aborder des sujets relatifs à la donnée et à l'open data avec vos collègues.

Vos atouts pour animer un café de la data ou de l'IA :

- être curieux ou curieuse du sujet de la data et avoir envie de le partager avec des collègues,
- avoir visionné les vidéos DATAcculturation,
- avoir pris connaissance de la **PARTIE 2** du livret, qui vous donne accès à des ressources complémentaires aux vidéos et notamment au détail des réponses aux quiz proposés.

Nous vous avons concocté des formules clés en main qui permettent de déguster un café chaud et servi à table, avec ou sans sucre, c'est au choix.

Chaque formule est composée d'une même base :



la tasse : c'est l'ice breaker*, ou *activité brise-glace* ; il varie selon les formules. Dans le livret, nous vous en proposons plusieurs versions, adaptées à la durée des cafés et à la composition du public.



le café : c'est l'introduction thématique du café, le cœur du sujet. Il s'agit des vidéos DATAcculturation. Selon les formules, nous vous proposons d'en diffuser une ou plusieurs.

*voir lexique page 45



Le carré de chocolat : une séquence de débriefing autour de la ou des vidéos visionnées. Selon la formule choisie, il s'agira soit d'un quiz rapide sur la vidéo visionnée soit d'une séquence participative.



Le sucre : agrémentez votre café ! Pour terminer votre animation sur une note plus douce, selon la durée de la formule choisie, nous vous proposons différentes pistes d'ouverture.

ASTUCE : si vous êtes amateur ou amatrice de café, c'est-à-dire que vous avez déjà animé des séquences d'idéation ou d'échange autour d'un sujet sur des formats similaires, vous pouvez mixer les formules entre elles !

Chaque formule proposée dans ce livret étant toujours composée de la même base (un ice breaker + une introduction thématique + une séquence interactive + une ouverture), vous pouvez tout à fait mixer les éléments et créer votre formule sur-mesure ! Et pour les baristas d'expérience, vous pouvez même vous inspirer de vos propres méthodes d'animation pour composer votre café de la data !



Identifiez bien les profils des agents qui composeront votre café : sont-ils novices sur le sujet ou déjà un peu (voire très) avertis ?

Posez la question à vos collègues lorsque vous les invitez à participer à votre café de la data pour adapter votre animation et ajuster les éléments présentés.

Retour d'expérience et conseil pour animer un café de la data

La convivialité importe plus que le lieu ! Prévoyez du café (ou du thé) et si possible des choses à grignoter pour vous assurer que tout le monde soit à l'aise.



ANIMER UN CAFÉ DE LA DATA OU DE L'IA EN UN COUP D'ŒIL...

Vous souhaitez organiser un café de la data ou de l'IA avec vos collègues ?

OUI

C'est parti ! Voyons ensemble comment vous aider à choisir le format adapté...

NON

Ce livret peut quand même vous intéresser...
Rdv page 52 !

Avez-vous plus de 30 minutes pour animer votre café ?

OUI

Les formules Americano et Cappuccino sont faites pour vous !
Rdv pages 25 à 38

NON

Les formules Ristretto sont faites pour vous !
Rdv pages 17 à 24

Le public qui assistera à votre café sera plutôt...



NOVICE SUR LA DATA OU L'IA

Les formules « avec sucre » sont faites pour vous !
Rdv pages 17 (formule Ristretto) et 25 (formule Americano)



DÉJÀ AVERTI SUR LA DATA OU L'IA

Les formules « sans sucre » sont faites pour vous !
Rdv pages 21 (formule Ristretto) et 28 (formule Americano)

DE NIVEAU INTERMEDIAIRE SUR LA DATA ET L'IA

Les formules Cappuccino sont faites pour vous – *Rdv pages 30 et 38*

ASTUCE : vous souhaitez organiser un café de la data à distance et en visio ? Rendez-vous page 13 pour un aperçu des menus disponibles et page 39 pour un déroulé adapté.

LES FORMULES TOUT COMPRIS : SYNTHÈSE DES FORMATS



RISTRETTO



De 15 à 30
minutes



De 5 à 20
personnes



Un écran
Une connexion internet

Avec sucre (page 17)

Ice breaker : *Ma donnée, Madonna ou les deux.*

Introduction thématique : visionnage de la vidéo sur la donnée et l'open data.

Séquence collective : quiz de debriefing.

Ouverture

Sans sucre (page 21)

Ice breaker : *Ma donnée, Madonna ou les deux.*

Introduction thématique : visionnage d'une des vidéos **Cas d'usage**.

Séquence collective : quiz de debriefing.

Ouverture



AMERICANO



De 30 à 60
minutes



De 5 à 20
personnes



Un écran
Une connexion internet

Avec sucre (page 25)

Ice breaker : placement dans l'espace à partir de questions sur la data.

Introduction thématique : visionnage de l'une des deux vidéos suivantes, au choix : *L'open data par et pour les agents publics* ou *Mes données personnelles sont-elles publiques ?*.

Séquence collective : *Ma journée d'agent.*

Ouverture

Sans sucre (page 28)

Ice breaker : les **Crazy jobs** sur la donnée.

Introduction thématique : visionnage de l'une des trois vidéos cas d'usage, au choix : *La Base adresse nationale, la base Géo'DAE* ou *L'open data des collectivités territoriales*.

Séquence collective : débat *d'accord ou pas d'accord ?*.

Ouverture



ASTUCE : adaptez le café de la data à votre structure, service, administration !

Chaque café de la data peut être adapté à votre environnement professionnel. Vous pouvez ainsi présenter des cas d'usage, des exemples de réutilisations de la donnée issus de votre périmètre ministériel, de votre établissement, service ou de votre collectivité.

Par exemple : vous organisez un café de la data pour acculturer vos agents ou vos collègues à l'importance de partager ses données en open data. Vous pouvez présenter l'une des vidéos DATAcculturation en fonction du menu d'animation choisi et adapter les exemples à votre contexte professionnel.

Si vous choisissez de diffuser la vidéo **La donnée et l'open data**, vous pouvez présenter des exemples ou des cas d'usage propres à votre structure :

- présentez une plateforme de partage de donnée de votre ministère ou de votre collectivité,
- montrez un exemple de partage et de réutilisation de données dans votre structure.

Vous pouvez aussi trouver un jeu de données ou une réutilisation adaptée à vos besoins sur data.gouv.fr !

Bon à savoir : il existe des plateformes de données publiques ouvertes dans chaque région de France. Vous pouvez donc trouver des exemples adaptés à votre territoire presque partout !

Pour en savoir plus sur les plateformes territoriales de données, vous pouvez consulter le rapport de l'Afigéo, Association française pour l'information géographique, ici : <https://www.afigeo.asso.fr/publication-les-plateformes-territoriales-de-donnees-geographiques-etat-des-lieux-enjeux-et-perspectives-2023/>

CAPPUCCINO



Entre 1h
et 1h15



De 5 à 20
personnes



Un écran
Une connexion internet
Un tableau blanc ou un
paper board

Formule à prendre avec ou sans sucre, pour un public mixte de novice à averti.

Ice breaker : *Crazy jobs* (voir page 28) ou *Ma donnée, Madonna ou les deux* (voir page 18).

Introduction thématique : quelques notions de base sur le big data et notions de vocabulaire et visionnage de la vidéo *L'open data par et pour les agents publics*.

Séquence collective : précision des notions de vocabulaire suite à la vidéo, séquence d'échanges et de débat et de démonstration d'outils en ligne.

Ouverture

CAFÉ CAPPUCCINO À LA MOUSSE D'IA



Environ 1h



De 5 à 20
personnes



Un écran
Une connexion internet
Un tableau blanc ou un
paper board

Formule à prendre avec ou sans sucre, pour un public mixte de novice à averti (voir des experts et expertes, page 34).

Ice breaker : *Ma donnée, Madonna ou les deux* (voir page 18). Photo langage avec des représentations de l'IA issues de la pop culture (voir page 34)

Introduction thématique : quelques notions de base sur l'IA et/ou le big data et notions de vocabulaire et visionnage de la vidéo sur la data et l'IA (épisode 6 de DATAculturation) en 2 parties.

Séquence collective : précision des notions de vocabulaire suite à la vidéo, séquence d'échanges et de débat et de démonstration d'outils en ligne

Ouverture



CAFÉ VIRTUEL



De 30 à
60 minutes



De 5 à 15
personnes



Un outil de visio adapté
Un support de présentation

Vous souhaitez réunir des collègues partout en France, dans votre région ou en télétravail ? Cette nouvelle formule vous propose d'adapter les menus existants pour un café de la data 100 % en ligne !

Formule à adapter en fonction de votre public, page 39.

Astuce : lors de l'inscription à votre café en ligne, demandez aux participants et participantes leur niveau pour mieux adapter votre animation.




Ice breaker : *Crazy jobs* (voir page 28) ou *Ma donnée, Madonna ou les deux* (voir page 18).

Introduction thématique : en fonction de votre public, vous pouvez mixer tous les éléments présentés dans les autres formules ! Retrouvez en pages 39 à 42 un exemple d'un déroulé possible adapté en visio.




Séquence collective : les quiz et les séquences **Débat** sont entièrement transposables dans cette formule 2.0 !

Ouverture

FORMULE 1 : CAFÉ RISTRETTO

-  Combien de temps ? Entre 15 et 30 minutes
-  Combien de personnes ? Maximum 20 personnes
-  Matériel : un écran

Trois formules au choix

-  Ristretto avec sucre, pour les novices
-  Ristretto sans sucre, pour un public averti
-  Ristretto virtuel, avec ou sans sucre rdv pages 39 à 42

ASTUCE : pour introduire les vidéos dans les différentes séquences et consulter des informations et des ressources complémentaires, rendez-vous à la **PARTIE 2** du livret, à partir de la page 52.



Ristretto avec sucre, pour les novices

Ice breaker *Ma donnée, Madonna ou les deux*

Durée : entre 5 et 10 minutes, selon la durée totale du café.
Cet ice breaker se présente comme un quiz simple où, à chaque affirmation, il faut déterminer s'il s'agit d'un sujet relatif à la donnée, à Madonna ou aux deux à la fois (ma donnée et Madonna).

Retour d'expérience et conseil pour animer un café de la data
Brisez la glace dès les premiers moments de l'échange !

*Les premières prises de paroles sont toujours les plus difficiles.
N'hésitez pas à commencer votre échange avec une activité ludique.*



- On en parlait moins dans les années 80 (réponse : ma donnée)
- Elle est née dans le Michigan (réponse : Madonna)
- Elle est de plus en plus souvent dans les nuages (réponse : les deux)
- Il faut régulièrement la contrôler (réponse : ma donnée)
- Elle est unique (réponse : Madonna)
- Elle est en vogue (réponse : les deux)
- Elle prend de la place (réponse : ma donnée)
- Il faut faire attention aux traces qu'elle peut laisser (réponse : ma donnée)
- Elle fait partie de ma vie (réponse : ma donnée ou les deux)
- Elle génère beaucoup de chiffre (réponse les deux)
- A rapporté plus d'un milliard de dollars de recette (réponse : les deux)

Introduction thématique : visionnage de la vidéo sur la donnée et l'open data ou de la vidéo sur les données et l'intelligence artificielle.

Durée : 5 minutes maximum si vous annoncez quelques éléments introductifs.

Quiz de débriefing de la vidéo sur la donnée et l'open data.

Posez les questions suivantes au groupe et prenez quelques instants pour expliquer les réponses et échanger. Durée : 5 à 10 minutes, à adapter en fonction de la durée totale du café.

QUESTION 1

Toutes les données que nous produisons ou que nous utilisons, au travail et dans notre vie privée, sont des données ouvertes et publiques. Vrai ou faux ? (réponse : faux).

Pour aller plus loin, vous pouvez introduire la notion de données à caractère personnel et de données sensibles et demander au groupe s'il peut citer des exemples de données sensibles et à caractère personnel (ex : numéro de carte vitale, date de naissance, etc.).

QUESTION 2

Citez une donnée ou type de donnée que vous produisez ou manipulez dans votre vie professionnelle (elle peut être sensible ou pas).

Pas de réponse vraie ou fausse, on peut aider le groupe à donner des réponses en se référant à des exemples comme : pointage quotidien

et horaires de travail, données RH pour les agents gestionnaires par exemple, données financières (ex. : suivi des dotations de l'État en préfecture), nombre d'accidents de la route dans le département, etc.

QUESTION 3

Parmi les enjeux de l'open data, identifiés dans la vidéo et présentés ci-après, se cache un intrus, lequel ?

- Renforcer la transparence de la vie publique.
- Trouver de nouvelles sources de financement des services publics (bonne réponse).
- Simplifier le service rendu aux citoyens et aux citoyennes.
- Améliorer les services publics.

Pourquoi est-ce une fausse réponse ? Car l'ouverture et le partage des données publiques est un principe qui repose sur le libre accès des données concernées et donc leur gratuité !

QUESTION 4 « BONUS »

La plateforme nationale des données ouvertes et publiques s'appelle...

- opendata.gouv.fr,
- donneesouvertes.gouv.fr,
- data.gouv.fr (bonne réponse).

ASTUCE : vous pouvez présenter la plateforme en quelques clics au groupe.

Quiz de débriefing de la vidéo sur les données et l'intelligence artificielle

QUESTION 1

L'invention de l'IA peut être datée...

- du film 2001 l'odyssée de l'espace avec le redoutable HALL en 1968,
- des années 50 avec les premières recherches lancées dans le domaine (réponse exacte),
- de 2015 avec la fondation d'OpenAI, la société qui a lancé ChatGPT.

QUESTION 2

L'IA est capable d'appeler un chat « un chat » sans se tromper parce que...

- L'IA est comme on dit une « cat person », elle adore les chats.



- On lui a appris à le faire en l'entraînant avec un grand nombre de photos de chats pour limiter ses erreurs (réponse exacte).

QUESTION 3

Qu'est-ce que le big data ?

(Posez la question aux participants/participant·es)

Réponse : le big data signifie littéralement « données massives » et décrit, dans le contexte du développement des technologies numériques, la production de plus en plus importante de données numériques dans le monde.

Sur cette question : n'hésitez pas, si vous avez le temps, à citer ou faire deviner les « 5 V » du big data : volume, variété, vitesse ou vélocité, véracité et valeur.

QUESTION 4

Dans les administrations, l'IA est déjà utilisée pour beaucoup de tâches, mais pas n'importe lesquelles. Cherchez l'intrus parmi les propositions suivantes :

- l'IA peut servir à synthétiser ou résumer des documents, des rapports,
- l'IA peut être utilisée pour diminuer les temps de traitement de certaines tâches répétitives et chronophages pour les agents,
- l'IA prend des décisions pour soulager les responsables de service dans cette tâche (réponse exacte),
- l'IA peut aider à classer des données produites en grande quantité et ainsi mieux les trier.

Fin du café de la data

Si vous avez encore quelques minutes, vous pouvez continuer à échanger sur le sujet, montrer des ressources disponibles sur le site data.gouv.fr par exemple. Vous pouvez conclure le café de la data en recommandant au groupe de visionner les autres vidéos DATAcculturation (annoncez les thématiques : géo DAE, base adresse nationale, données personnelles, collectivités locales et transparence).

Si vous avez visionné la vidéo sur les données et l'IA, servez-vous des références en partie 2 du guide pour donner des ressources complémentaires.



Ristretto sans sucre, pour un public averti

Ice breaker : identique à la formule *Ristretto avec sucre*.

Introduction thématique : visionnage d'une des vidéos *Cas d'usage*, au choix parmi celles sur la Base adresse nationale, la base Géo'DAE et l'open data des collectivités territoriales.

Durée : entre 5 et 10 minutes maximum, en fonction de la durée totale du café.

Quiz de débriefing de la vidéo

Posez les questions suivantes au groupe et prenez quelques instants pour expliquer les réponses et échanger.

Vous trouverez ci-après 3 quiz composés chacun de 3 questions pour les 3 vidéos. Choisissez bien le quiz correspondant à la vidéo diffusée pendant le café !

Durée : 5 à 10 minutes, en fonction de la durée totale du café.

Quiz pour la vidéo *Open data et collectivités territoriales : une histoire de transparence*

QUESTION 1

La loi rend obligatoire la publication en ligne de certains documents pour les communes de...

- 100 habitants et plus,
- 5000 habitants et plus,
- 3500 habitants et plus (bonne réponse).

QUESTION 2

La vidéo aborde le concept de la *marque blanche* pour présenter la solution mise en place par Mégalis en Bretagne. Qu'est-ce qu'une marque blanche ?

- Une trace blanche qui ne part pas au lavage, même à 60°.



- Un service ou un produit créé par une entreprise ou un organisme et mis à disposition d'autres utilisateurs gratuitement pour leur compte (bonne réponse).
- Une marque qui donne la possibilité à un autre utilisateur d'utiliser son logo, contre rémunération.

QUESTION 3

La vidéo illustre le sujet de l'open data comme enjeu de renforcement de la transparence de l'action publique. Cette exigence découle du principe défini à l'article 15 de la Déclaration des droits de l'homme et du citoyen de 1789. Oui, mais que dit cet article ?*

- *Les citoyens ont le droit de demander la publication de toutes les informations détenues par les administrations publiques.*
- *L'open data est un principe fondateur de la démocratie et sur lequel reposent les droits des citoyens.*
- *La société a le droit de demander compte à tout agent public de son administration (bonne réponse).*

** Pour en savoir plus, vous pouvez consulter le Guide pratique rédigé par la CNIL et la Cada : https://www.cada.fr/sites/default/files/guide_pratique.pdf*

Quiz pour la vidéo **La Base adresse nationale**

QUESTION 1

Qui est responsable de la mise à jour des bases adresses locales, qui alimentent la BAN ?

- Les communes (bonne réponse)
- La Poste
- Les impôts

QUESTION 2

La BAN est accessible...

- aux administrations uniquement,
- aux entreprises et aux particuliers ayant réglé un abonnement,
- à tout le monde et gratuitement (bonne réponse).

QUESTION 3**Où est précisément localisée géographiquement une adresse ?**

- Au débouché du local sur la route la plus proche
- À la boîte aux lettres
- À la porte d'entrée du bâtiment à localiser

(toutes les réponses sont correctes)

Quiz pour la vidéo *La base Géo'DAE***QUESTION 1****Y a-t-il un DAE sur votre lieu de travail ?**

Réponse : voir si les personnes répondent et leur rappeler qu'il s'agit d'une obligation légale pour les établissements recevant du public.

QUESTION 2**La loi oblige depuis 2018 les établissements et espaces publics disposant d'un DAE à les maintenir en bon état de marche et à les déclarer dans la base Géo'DAE dans un format de données standard. Selon vous, pourquoi est-ce nécessaire d'employer ce format standard ?**

- Pour diminuer le temps de vérification et de correction des données saisies dans la base ; c'est beaucoup de travail !
- Pour fiabiliser le schéma de transmission des données ; c'est quand même plus pratique, non ?
- Pour faciliter la création de services.

(les 3 réponses sont bonnes !)

QUESTION 3**L'utilisation d'un DAE lors d'un arrêt cardiaque augmente les chances de survie de...**

- 5 %
- 85 %
- 40 % (bonne réponse)



Fin du café de la data

Si vous avez encore quelques minutes, vous pouvez continuer à échanger sur le sujet, montrer des ressources disponibles sur le site data.gouv.fr par exemple.

Vous pouvez enfin conclure le café de la data en recommandant au groupe de visionner les autres vidéos du projet DATAculturation (annoncez les thématiques abordées).

Retour d'expérience et conseil pour animer un café de la data.

Osez vous lancer sans douter de votre légitimité ! Animer un temps d'échange autour de la technologie n'est pas seulement réservé aux personnes expertes. Si vous trouvez un intérêt dans ce sujet, ne doutez plus et lancez-vous !



FORMULE 2 : CAFÉ AMERICANO



Combien de temps ? entre 30 minutes et 1 heure



Combien de personnes ? de 5 à 20 personnes



Matériel : un écran (TV, ordinateur)

Trois formules au choix :



Americano avec sucre, pour les novices



Americano sans sucre, pour un public averti



Americano virtuel, avec ou sans sucre rdv pages 39 à 42



Americano avec sucre, pour les novices

Ice breaker : placement dans l'espace à partir de questions sur la data.

Durée : 5 à 10 minutes maximum, selon le nombre de participants/participant.es.

Séparer l'espace en 2 zones distinctes : une zone **Plutôt d'accord** et une zone **Plutôt pas d'accord**. Vous allez alors présenter plusieurs affirmations et les personnes vont y répondre en se plaçant dans la zone correspondant à leur avis. Une fois chaque participant placé, n'hésitez pas à lancer le débat en leur demandant pourquoi ils s'y sont placés.

ASTUCE : avant de vous lancer, distinguez *données personnelles* et *données sensibles* avec votre groupe.

Vous pouvez soit leur demander s'ils connaissent la différence et peuvent citer des exemples. Si non, vous leur en donnez.

Pour vous aider, rendez-vous page 57.



Les affirmations

- J'ai déjà tapé mon prénom et mon nom dans une barre de recherche Google.
- Je suis du genre à accepter facilement les cookies.
- Les données sensibles, je n'en produis pas dans le cadre de mon travail.
- C'est principalement le service informatique qui doit gérer le sujet de la donnée.
- J'ai déjà bloqué l'accès aux données de géolocalisation sur mon téléphone.
- Je connais le concept de publicité ciblée.
- Je n'ai rien à cacher.
- J'ai déjà googlé une/un ex.
- Mes comptes sur les réseaux sociaux sont en mode « privé ».

Introduction thématique. Visionnage des deux vidéos suivantes :

- l'open data par et pour les services publics,
- mes données personnelles sont-elles publiques ?

Durée : 10 minutes environ si vous introduisez les vidéos par un petit contexte à chaque fois.

Séquence collective : *ma journée d'agent*

L'objectif est de reconstituer le parcours des agents dans une journée de travail type, depuis le trajet vers le bureau jusqu'au retour à la maison, pour identifier quand ils croisent des données.

Durée : 10 à 15 minutes avec un temps d'introduction et des échanges avec les participants et les participantes.

Matériel : un tableau blanc ou un paper board ou des feuilles de post-it géantes à accrocher au mur, des post-it, des stylos.

Écrivez en gros les grands moments de la journée à identifier :

- 1) le trajet domicile/travail,
- 2) journée de travail : tâches quotidiennes,
- 3) pause déjeuner,
- 4) activités extérieures : sportives, culturelles, familiales, etc.,
- 5) trajet travail/domicile (si différent et si éléments à ajouter).

Travaillez le profil de l'agent avec vos participantes et participants : prénom, âge, situation professionnelle et personnelle. Cela permet de mieux s'approprier le personnage.

Sur une table, disposez des post-it de 4 couleurs différentes en identifiant les types de données auxquels ils correspondent. Par exemple : les post-it bleus pour les données personnelles, les post-it roses pour les données sensibles, les post-it jaunes pour les données publiques



non ouvertes et les post-it verts pour les données ouvertes (open data). Demandez aux participantes et participants de réfléchir, en fonction de chaque moment de la journée de l'agent, aux données créées ou partagées en fonction de leur typologie, en respectant le code couleur des post-it. Elles et ils peuvent coller les post-it au fur et à mesure de leurs idées sur le panneau.

Tout au long de l'atelier, aidez les agents à compléter le tableau en leur donnant des exemples de données type ou en participant vous-même à l'exercice.

Exemples de données : consulter les horaires de bus/métro, se connecter à son compte Facebook, pointer, utiliser la fonction de géolocalisation, passer à la bibliothèque, au restaurant administratif, etc.

Synthèse et ouverture

À la fin de l'exercice, vous pouvez demander aux agents de faire le point sur l'exercice et ensuite sur le café.

Posez-leur les questions suivantes :

- qu'est-ce que vous surprend ?
- qu'est-ce qui ne vous surprend pas ?
- qu'avez-vous appris ?

Durée : 5 à 10 minutes, en fonction de la durée totale du café.



Americano sans sucre, pour un public averti

Ice breaker : les *crazy jobs* sur la donnée.

Durée : 5 à 10 minutes maximum, selon le nombre de participants/participant.es.

Les participants/participant.es doivent choisir un ***crazy job*** ou ***travail fou*** parmi ceux proposés dans la liste ci-après. Donnez leur quelques instants pour imaginer leur travail quotidien ou justifier leur choix, puis invitez les à s'exprimer tour à tour.

Conseil : cet ice breaker fait appel à l'imagination et suppose de faire un pas de côté car les jobs proposés n'ont a priori aucun sens sur le papier. Donnez l'exemple en choisissant vous-même un *crazy job* et en passant au tableau en premier ! Un même *crazy job* peut être choisi par plusieurs personnes.

ASTUCE : une liste enrichie de *crazy jobs* est disponible en annexe de ce guide. Vous pouvez même imprimer et découper les *crazy jobs* sous forme de cartes ! Rendez-vous pages 43-44 et sur le site du Lab Archipel pour télécharger le support.

Liste des *crazy jobs* :

- passeur/passeuse d'aspirateur à données,
- gardien/gardiennne de coffre-fort de données personnelles,
- inventeur/inventrice de mots de passe inviolables,
- sculpteur/sculptrice de profil en ligne,
- inspecteur/inspectrice à cookies,
- historien/historienne d'historiques internet.

Introduction thématique : visionnage de deux vidéos *Cas d'usage*, au choix, parmi celles présentes sur la Base adresse nationale, la base Géo'DAE et l'open data des collectivités territoriales.

Durée : 10 à 15 minutes avec un temps d'introduction et des échanges avec les participants et les participant.es.

Séquence débat

L'objectif est de susciter des échanges en réaction à des affirmations à partir desquelles il est demandé au groupe de se positionner (d'accord/pas d'accord). Nous vous proposons, au choix, les affirmations ci-dessous.

Selon la durée de votre café et les échanges au sein du groupe, vous pouvez aborder plus ou moins de thèmes. Vous pouvez également enrichir la liste des affirmations ! Pendant l'exercice, si possible, notez les arguments principaux sur une feuille ou un tableau.

	D'accord	Pas d'accord
Je me sens concerné/concernée par le sujet de la donnée		
La donnée est un sujet à la mode		
Les données publiques doivent être accessibles gratuitement et pour tout le monde		
L'open data est utile dans mon quotidien		
Si c'est gratuit, c'est que vous êtes le produit		
La donnée devrait être un moyen pour l'État et les collectivités publiques de trouver des sources de financement complémentaires		

Synthèse et ouverture

À la fin de l'exercice, synthétisez les principaux arguments échangés au cours des échanges. Reformulez-les avec le groupe si besoin. Vous pouvez demander aux participants/participantantes ce que l'atelier leur a apporté.

Durée : 5 à 10 minutes, en fonction de la durée totale du café.



FORMULE 3 : CAFÉ CAPPUCCINO



Combien de temps ? Entre 1h et 1h15



Combien de personnes ? De 5 à 20 personnes



Matériel : un écran, une connexion internet, un tableau blanc ou un paper board

Formule à prendre avec ou sans sucre, pour un public mixte de novice à averti (voir quelques expertes et experts, etc.).

ASTUCE : pour bien identifier les niveaux de chaque participant et participante, installer un panneau ou une feuille sur laquelle ils et elles pourront se positionner.

Exemple :

La donnée et vous... Quel est votre niveau ? Mettez une croix / une gommette / un aimant en face de votre niveau

Novice :

Averti ou avertie :

Expert ou experte :

Ice breaker : au choix entre

- **Ma donnée, Madonna ou les deux :** voir page 18
- Les **crazy jobs :** voir pages 28 et annexes pages 43-44

Introduction thématique

Quelques propos que vous pouvez adapter en fonction de vos connaissances pour planter le décor et faire réagir.

En 2020 savez-vous à combien on estime le nombre de données créées et partagées dans le monde ?

Le chiffre estimé est 64,2 zettaoctets. Pour mieux faire réagir vous pouvez dessiner ce qu'1 zettaoctet représente : 1 suivi de 21 zéros !

On prévoit un taux de croissance pour la création et la réplique annuelle des données de 23 % entre 2020 et 2025. Selon certaines projections, ce le volume de données créées et partagées dans le monde serait de 181 zettaoctets en 2025.

Les données nous entourent et avec l'avènement du numérique et d'internet, elles se multiplient et s'échangent toujours plus vite. C'est un enjeu stratégique politique et économique. On parle maintenant de Big Data et on le définit par les 5 V (volume, vitesse, variété, valeur et véracité des données créées et partagées).

Avant de commencer, demandez au groupe de donner une définition des termes suivants : donnée, information, donnée publique, donnée ouverte. *Noter les réponses des participants et participantes au tableau sans dire si c'est juste ou faux.*

Visionnage de la vidéo sur la donnée et l'open data.

Débrief de la vidéo

Qu'avez-vous retenu de la vidéo ? Quels sont pour vous les enjeux de l'open data ?

Compléter les définitions après le visionnage et donner des définitions (vous pouvez ajuster les définitions en fonction de vos exemples) :

- **donnée** : 1 élément constitué à partir de règles, unités de mesure ou indicateurs, qui peut être observé, collecté, stocké. Exemple : la température (relevée en Celsius ou Fahrenheit), la longueur ou la hauteur d'un bâtiment (relevée en mètres), la distance (kilomètres, miles), etc.,
- **information** : c'est l'interprétation, l'analyse réalisée à partir de données et dans un contexte précis. Ex. : la combinaison des températures, du niveau d'ensoleillement et de la pluviométrie (données) peuvent permet la construction d'un bulletin météo (information).



- donnée publique : c'est une donnée produite dans le cadre de missions de service public (par une administration ou un établissement doté d'une mission de service public). Une donnée publique n'est pas forcément publiée en ligne et disponible en open data. Par exemple, les données relevant du secret de la Défense nationale ou des données à caractère personnel ne peuvent pas être rendues publiques, accessibles en open data.
- donnée ouverte (open data) : une donnée librement accessible, en général sur une plateforme (format numérique), gratuite et réutilisable.

Exemple par la grille FAIR* pour définir une donnée ouverte :

- facilement trouvable,
- accessible,
- interopérable,
- réutilisable,
- + gratuite !

Aller plus loin

Pouvez-vous citer des formats de données ouverts ?

CSV : *comma separated values* → c'est un format de fichier de base, le plus ouvert, pour collecter, stocker, transmettre et réutiliser des données. Chaque champ (cellule) est séparée par une virgule à l'encodage, une convention qui permet à un plus grand nombre d'applications d'utiliser les données contenues dans ces fichiers.

Retour d'expérience et conseil pour animer un café de la data.

Privilégiez les interactions ! Bien que de nombreuses informations soient partagées durant le café, il ne faut pas donner l'impression qu'il s'agit d'une conférence. Les interactions avec et entre les participants doivent sans cesse être encouragées.

Pourquoi les formats PDF et Excel ne sont pas considérés comme réellement *ouverts* au sens de l'open data ? Car ce sont des formats dits *propriétaires*, sous licence.

Séquence collective

Débat à choisir parmi les questions/affirmations suivantes ou celles que vous souhaitez créer !

- Toutes les données publiques doivent être accessibles gratuitement et pour tout le monde.
- L'open data est utile dans mon quotidien.
- La donnée devrait être un moyen pour l'État et les collectivités publiques de trouver des sources de financement complémentaires.

Ouverture

Si vous avez le temps, vous pouvez visionner une des trois vidéos.

Exemples (la BAN, la base Géo'DAE ou l'open data des collectivités) pour ouvrir la réflexion sur les usages de la donnée ouverte sur le territoire.

**Les principes FAIR (Findable, Accessible, Interoperable, Reusable) décrivent comment les données doivent être organisées pour être plus facilement accessibles, comprises, échangeables et réutilisables. Source : CCSD, Centre pour la communication scientifique directe. <https://www.ccsd.cnrs.fr/principes-fair>*



FORMULE 3 BIS : CAFÉ CAPPUCCINO À LA MOUSSE D'IA



Combien de temps ? Environ 1 heure



Combien de personnes ? Entre 5 et 20 personnes



Matériel : un écran, une connexion internet, un tableau

Formule à prendre avec ou sans sucre, pour un public mixte de novice à averti (voir quelques expertes et experts, etc.).

Ice breaker : au choix entre

- **Ma donnée, Madonna ou les deux, page 18**
- Photo langage avec des représentations de l'IA issues de la pop culture : imprimez ou affichez sur un écran des images d'intelligences artificielles extraites de films ou de la pop culture et demandez aux participant/participantes de choisir une image qui les inspire ou reflète leur relation ou leur sentiment par rapport à l'intelligence artificielle (ou un souvenir ?).

Idées d'images à reprendre : le réseau Skynet (Terminator), HAL (2001, l'Odyssée de l'espace), Samantha (HER), Viki (I, Robot), David (A.I. intelligence artificielle), Rachael (Blade Runner), J.A.R.V.I.S. (Iron Man), la Matrice (Matrix), la cyberpolicier Motoko Kusanagi (Ghost in the shell) ou encore le robot Maria (Metropolis).

Conseil : n'oubliez pas de présenter un résumé des films proposés (soit au verso des cartes soit à l'oral si vous projetez les images).

Introduction thématique

Vous pouvez reprendre des éléments cités dans le déroulé d'animation page 31 pour introduire sur les notions de données massives (big data) et donner des éléments de définition.

Visionnage au choix

- De la vidéo sur la donnée et l'open data.
 - Puis de la vidéo (en 2 parties) sur data et IA.
 - Ou de la vidéo sur data et IA seule (toujours en 2 parties).
- Organisez le visionnage des vidéos en fonction du temps dont vous disposez et des échanges. Les vidéos contiennent beaucoup d'informations, ne négligez pas les temps de réflexion, de pause, voire de reVISIONNAGE des vidéos pour rythmer les échanges et améliorer la compréhension de votre public.

Débriefing de la ou des vidéo(s)

Des éléments de débriefing peuvent être piochés dans le déroulé d'animation page 31.

Pour ensuite introduire la vidéo sur la data et l'IA (à visionner en 2 parties), voici quelques éléments :

- éléments de définition : sauriez-vous donner la définition de **intelligence artificielle** ? **Algorithme** ? **Machine learning** ou **apprentissage automatique** ? **IA générative** ?
- Comme pour le déroulé page 31, notez au tableau ou de façon visible les réponses données par les participants/participant.es, sans dire si elles sont justes ou fausses.
- vous pouvez aussi introduire en retraçant une brève frise chronologique de l'IA (quelques exemples de dates) :
 - seconde guerre mondiale : le test de Turing ou les prémices de l'IA,
 - 1956 : l'IA est reconnue comme un domaine de recherche et une discipline scientifique (conférence de Dartmouth),
 - années 50-70 : les recherches progressent autour des courants symbolique et numérique,
 - années 80 : les recherches connaissent des avancées dans les systèmes experts (reproduisent les capacités cognitives humaines),
 - 1997 : victoire de l'IA d'IBM Deep Blue contre Garry Kasparov (échecs),
 - années 90 et 2000 : nouvelles performances en reconnaissance d'images, avancées de l'IA générative, etc.



- Vous pouvez retrouver des détails sur ces éléments historiques :
- dans le module pédagogique **Qu'est-ce-que l'IA ?** sur le site des Cafés IA : <https://cafeia.org/wp-content/uploads/2024/09/Copy-of-Module-pedagogique-Quest-ce-que-l-IA-.pdf>
 - dans les ressources du jeu **Aïe Aïe IA**, séquence **Sur les traces de l'IA** : https://resin.grandlyon.com/shared/jeuAieAieIA/impression/CARTES_INSPECTION_sur-les-traces-de-l-IA.pdf

Visionnage de la vidéo > Partie 1

Débriefing de la vidéo

- Donnez le temps aux participants/participantes de réfléchir après le visionnage, demandez-leur comment ils et elles réagissent, si ils ou elles ont des questions.
- Si vous avez choisi d'introduire avec des éléments de définition, demandez aux participants/participantes de corriger/ajouter des éléments aux définitions données avant la vidéo.

Donnez les définitions exactes au fil des échanges :

- **algorithme** : un ensemble d'instructions conçues pour résoudre un problème. C'est une suite d'étapes pour aboutir à un résultat à partir d'informations d'entrée.
Question : pouvez-vous citer des algorithmes ?
Réponse très simple : une recette de cuisine fonctionne comme un algorithme.
- **intelligence artificielle** :
 - la définition donnée dans la vidéo : l'IA est un ensemble d'algorithmes conçu pour imiter l'intelligence humaine,
 - on peut simplifier encore : l'IA c'est *faire faire à un ordinateur des fonctions typiques de la cognition*. Cette définition montre surtout que l'IA reste un outil produit par les humains pour reproduire des fonctions de l'intelligence humaine : on ne parle donc pas d'une IA autonome comme dans Terminator.
- **machine learning** ou **apprentissage autonome** : certains algorithmes sont *évolutifs* ou *apprenants*, leur comportement

va évoluer dans le temps en fonction des données fournies. Vous pouvez détailler les 3 grands types d'apprentissages automatiques : non supervisé, supervisé (celui qu'on décrit dans la vidéo) et par renforcement (voir le module **Qu'est-ce-que l'IA ?** cité plus haut).

- **IA générative** : ah c'est celle dont on entend le plus parler ces dernières années...

Mais c'est quoi réellement ? Ce sont des types d'IA qui produisent du contenu (texte, audio, vidéo, image) en réponse à un prompt ou une requête, c'est-à-dire la demande d'un utilisateur ou une utilisatrice.

Vous pouvez simuler la formulation d'un prompt pour aider les participants/participant·es à comprendre ou illustrer cette définition par des exemples.

➔ Ne négligez pas la quantité d'informations contenue dans cette première partie. Pour expliquer les concepts, les définitions, n'hésitez pas à faire des arrêts sur image à des moments clés de la vidéo, comme l'explication de l'apprentissage supervisé avec les photos de chats.

Visionnage de la vidéo > Partie 2

Débriefing

Cette partie se concentre sur les enjeux de l'utilisation de l'IA dans les services publics et les administrations. Vous pouvez répondre aux éventuelles questions des participants/participant·es ou les laisser partager des retours d'expérience, puis enchaîner sur l'une des séquences participatives proposées ci-après.

Séquence participative

Proposition 1 : lancer un débat à partir d'idées reçues sur l'IA

- **L'IA sait tout faire.** Qu'en pensez-vous ?
- **L'IA pourra bientôt (ou peut déjà?) accomplir des tâches et prendre des décisions comme les humains.** Qu'en pensez-vous ?



- *L'IA va faire disparaître le travail des humains.*
Qu'en pensez-vous ?
- *L'IA peut-elle créer de l'art ?* Qu'en pensez-vous ?
- *Chat GPT va-t-il tuer notre imagination ?* Qu'en pensez-vous ?
- *Préférez-vous être recruté/recrutée par une IA ou par un humain ?* Qu'en pensez-vous ?

ASTUCE : proposez votre sujet de débat si vous avez des idées !

Proposition 2 : pour tester la fiabilité des réponses apportées par des outils d'IA génératives, faites tester ComparIA : www.comparia.beta.gouv.fr/

Objectif : montrer qu'il y a autant de réponses possibles que d'outils d'IA existants et que la règle, comme pour toute recherche, c'est de comparer, diversifier ses sources. Vous pouvez profiter de cet atelier pour sensibiliser les participants/participantes. ComparIA vise aussi à enrichir les outils d'IA générative avec des données en langue française, car la majorité des requêtes et des données d'entraînement de ces outils (largement américains) sont formulées en langue anglaise.

ASTUCE : mettez à disposition un ordinateur ou une tablette pour que vos participants/participantes testent ComparIA en direct (ou un autre outil d'IA de votre choix).

FORMULE CAFÉ VIRTUEL

**Vous souhaitez organiser un café de la data à distance ?
Toutes les formules proposées dans ce guide sont transposables
en visio, à condition de maîtriser un outil de visio
et éventuellement d'autres outils collaboratifs en ligne
ou de préparer quelques supports d'animation à diffuser.**



Combien de temps ? Durée maximale recommandée : 45 minutes



Combien de personnes ? Jusqu'à 20 personnes



Matériel : un ordinateur, une connexion internet, un support de présentation, éventuellement un outil collaboratif en ligne si vous en maîtrisez un (post-it, tableau blanc, etc.).

Formule à prendre avec ou sans sucre, pour un public spécifique ou mixte, de novice à expert.

AVANT DE VOUS LANCER

Organiser un café de la data en visio demande un peu plus de temps préparation que pour un café de la data classique : choix et test des outils diffusion des liens de visioconférence, préparation de supports à diffuser à l'écran, etc.

Ne négligez pas ces étapes ! Réfléchissez aussi à la manière dont vous allez diffuser la ou les vidéos pendant le café virtuel : tous les outils de visioconférence ne permettent pas de diffuser des vidéos. Vous pouvez aussi partager le lien vers la ou les vidéos dans la messagerie instantanée de l'outil et demander à tout le monde de couper son micro et de signaler par un message quand il ou elle a fini le visionnage (prévoir entre 2 et 5 min).



Enfin, les cafés de la data organisés en visio sont en général moins conviviaux que les cafés en présentiel : tout le monde n'active pas sa caméra, les participants/participantes ont tendance à réagir dans le chat plutôt qu'à prendre la parole. En tant qu'animateur ou animatrice, vous devrez sans doute vous montrer encore plus dynamique que pendant un café en présentiel !

ASTUCE : identifiez le niveau de votre audience en posant la question dans votre formulaire d'inscription avant la visio !
Vous adapterez ainsi mieux votre animation.

Adaptation des ice breakers

Toutes les animations proposées dans le guide, sauf le placement dans l'espace, sont transposables à distance ! Beaucoup d'outils de visio permettent de créer des sondages interactifs en direct, comme webinaire de l'État, et ainsi de poser des questions en direct.

Vous pouvez par exemple utiliser ces modules pour l'ice breaker **Ma donnée, Madonna ou les deux**. Pour le ice breaker **Crazy jobs**, astuce : écrivez tous les crazy jobs sur un support projeté à l'écran, laissez les participantes et participants choisir le job qui leur convient le mieux et donnez-leur la parole tour à tour pour s'exprimer.

Adaptation des introductions thématiques : là encore, tout le contenu est transposable en ligne.

ASTUCE : comme vous n'aurez pas à votre disposition de tableau blanc ou de paper board comme en salle, préparer des éléments de présentation sur un support projeté en séance.
Le contenu présenté sera ainsi plus interactif et visuel !

Pour visionner une ou plusieurs vidéos : si l’outil de visio que vous avez choisi permet le partage d’écran et de diffuser en direct une vidéo, sans trop en dégrader la qualité sonore et visuelle, vous pouvez visionner une vidéo DATAcculturation.

Nous vous recommandons néanmoins de partager le ou les liens Youtube des vidéos disponibles sur la page Archipel, ou tout autre lien si vous hébergez vous-mêmes les vidéos.

Diffusez le lien dans le chat de votre outil de visio et demandez aux participantes et participants, en coupant leurs micros, de regarder la vidéo choisie. Vous pourrez en parler ensuite.

Lien vers les vidéos en ligne : <https://www.youtube.com/playlist?list=PLXxl21OaojlyAwX-Tg1m-z4NYKx11cI3l>

Séquences collectives

- Pour les formules ristretto, les quiz sont facilement adaptables en ligne : utilisez des sondages interactifs si votre outil de visio en propose (ex : webinaire de l’État) ou préparez les questions sur un support projeté à l’écran et demandez aux participants et participantes de répondre dans le chat ou à l’oral.
- Pour les formules americano, plusieurs possibilités :
 - si vous choisissiez une séquence **Débat** (page 29), vous pouvez inscrire la question ou l’affirmation débattue sur un support projeté à l’écran, ou dans le chat de la visio, ou l’énoncer à l’oral et ensuite animer la séquence d’échanges,
 - si vous choisissez la séquence **Ma journée d’agent** (page 27), il vous faudra préparer une séquence interactive sur un outil collaboratif en ligne, permettant par exemple d’utiliser un tableau blanc ou des post-it. La séquence vous demandera plus de préparation en amont du café, pour préparer les différents panneaux et post-it en couleurs.



Ouverture

Là encore, à vous de préparer le contenu que vous souhaitez présenter. Vous pouvez également partager votre écran pour naviguer sur des plateformes de partage de données et montrer des exemples aux participantes et participants.

ASTUCE : consulter la **PARTIE 2** du guide pour trouver des ressources complémentaires et des liens à partager !

CARTES CRAZY JOBS





Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Historien ou historienne d'historiques internet



Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Pleur ou plieuse de courriels



Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Libérateur ou libératrice de données enfermées



Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Plombier ou plombière de flux de données



Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Cafés de la data - Crazy jobs

Trafiquant ou trafiquante de données frelatées



TOUS LES CONSEILS ET RETOURS D'EXPÉRIENCE DES ANIMATEURS ET ANIMATRICES DE CAFÉS DE LA DATA

« La convivialité importe plus que le lieu !

Prévoyez du café (ou du thé), voire des choses à grignoter pour vous assurer que tout le monde soit à l'aise.

Brisez la glace dès les premiers moments de l'échange !

Les premières prises de paroles sont toujours les plus difficiles. N'hésitez pas à commencer votre échange avec une activité ludique.

Osez vous lancer sans douter de votre légitimité !

Animer un temps d'échange autour de la technologie n'est pas seulement réservé aux personnes expertes. Si vous trouvez un intérêt dans ce sujet, ne doutez plus et lancez-vous !

Privilégiez les interactions !

Bien que de nombreuses informations soient partagées durant le café, il ne faut pas donner l'impression qu'il s'agit d'une conférence. Les interactions avec et entre les participants doivent sans cesse être encouragées.

Appuyez-vous sur les conseils des autres animateurs et animatrices.

Il ne faut pas hésiter à demander de l'aide auprès de personnes ayant déjà de l'expérience dans l'animation de ces temps d'échange. »

ASTUCE : retrouvez la communauté DATAcculturation sur Tchapp pour échanger avec des animateurs/animatrices de cafés de la data



DATAcculturation - fil d'actualité du projet

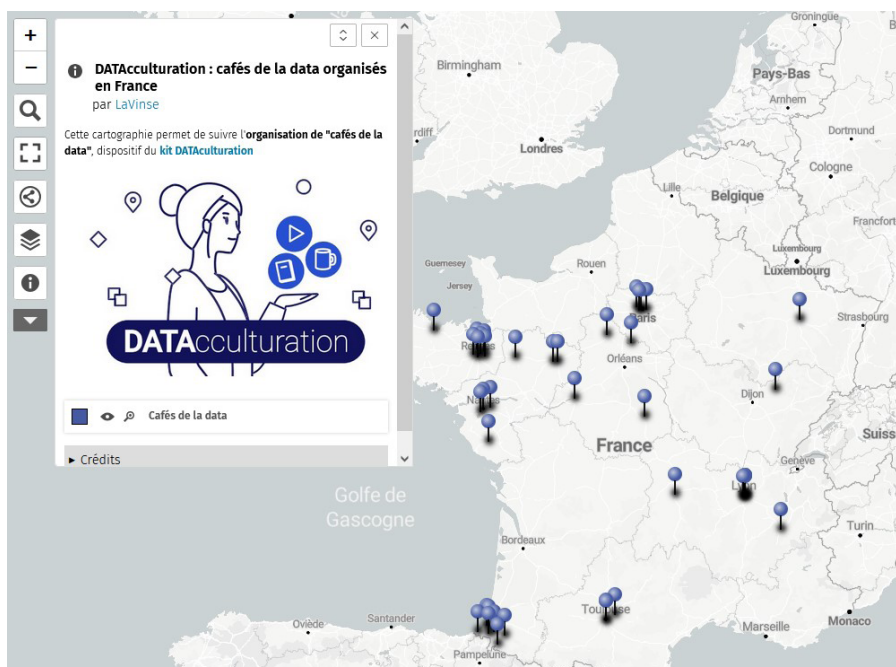
Utilisez la communication en amont pour séduire un plus large public !

Ces temps d'appropriation collective de la technologie peuvent parfois être considérés de l'extérieur comme des temps experts. Bien expliquer qu'ils s'adressent à toutes et tous, sans prérequis, est important.



DÉCOUVREZ LA CARTE DES CAFÉS DE LA DATA ORGANISÉS PEUT-ÊTRE PRÈS DE CHEZ VOUS...

La carte est collaborative, en édition libre. Si vous animez un café de la data dans votre structure, n'hésitez pas à l'ajouter sur la carte directement ou à contacter justine.sperandio-martinez@auvergne-rhone-alpes.gouv.fr pour toute demande.



https://umap.openstreetmap.fr/fr/map/dataculturation-cafes-de-la-data-organises-en-fra_1002168#6/47.160/3.494

PASSEZ À LA PRATIQUE : IMAGINEZ VOTRE CAFÉ DE LA DATA



A series of horizontal dotted lines for writing, spanning the width of the page.



A series of horizontal dotted lines spanning the width of the page, intended for handwritten notes.



PRÉFÈTE
DE LA RÉGION
AUVERGNE-
RHÔNE-ALPES

Liberté
Égalité
Fraternité

DATAcculturation

Cafés de la data

Parler **open data*** simplement
et autour d'un café, c'est possible !



Visionnez les vidéos !



**Vous êtes agent public
et le sujet de l'open data* vous intéresse ?**

Vous souhaitez en apprendre plus ?

*open data : données ouvertes

Participez
à un
Café de
la data !

Pensez à venir
avec votre tasse !



PARTIE 2



POUR ALLER PLUS LOIN

Sensibiliser à la donnée et à l'intelligence artificielle



Dans cette partie, vous trouverez une liste d'ingrédients vous permettant d'enrichir vos connaissances et d'agrémenter votre café de la data. Vous accédez à des ressources complémentaires sur les vidéos.

LES VIDÉOS DATAACULTURATION

Le projet DATAculturation s'articule autour de six capsules vidéos courtes (dont une vidéo en 2 parties), conçues pour être accessibles au plus grand nombre d'agents. Leur objectif est de sensibiliser à la donnée comme partie intégrante de notre environnement professionnel pour éclairer sous un nouveau jour nos missions mais aussi notre vie quotidienne. Elles abordent des concepts de façon pédagogique pour donner un premier niveau de connaissance sur des sujets souvent perçus comme complexes et réservés à des experts/expertes. Elles sont réalisées en images animées pour rendre les concepts faciles à appréhender.

Chaque vidéo aborde une thématique ou précise un concept sur le sujet de la donnée et de l'open data. Les vidéos peuvent être visionnées indépendamment les unes des autres et sans connaissances préalables. Chaque vidéo est cependant accompagnée de ressources et de supports complémentaires pour aller plus loin dans la compréhension des sujets abordés.

Cette partie du livret éclaire l'histoire de la conception de chaque vidéo et donne des pistes d'exploration complémentaires des concepts évoqués. L'objectif n'est pas de donner accès à des connaissances exhaustives sur chaque sujet mais de compléter les notions abordées dans les vidéos.



Vidéo - La donnée et l'open data : la donnée par et pour le service public.

Cette vidéo explore la thématique de la donnée et de l'open data en présentant des notions clé pour comprendre le sujet et les enjeux liés à la politique d'open data en France.

Objectif n°1 : prendre conscience que nous utilisons et produisons des données au quotidien, aussi bien dans notre vie professionnelle que dans notre vie personnelle.

La vidéo propose de suivre Karine, agent en préfecture, dans son quotidien au travail et dans sa vie privée. Ce point de vue rend les notions abordées concrètes. Comme Karine, nous produisons des données au quotidien : sur notre lieu de travail (pointage horaires), dans les transports (validation des titres) ou dans notre sphère (navigation internet).

Objectif n°2 : introduire des éléments sur la politique d'open data et ses enjeux.

Toutes les données produites et utilisées ne sont pas des données publiques ni même des données ouvertes, au sens de l'open data.

La vidéo présente néanmoins, à partir de certains types de données publiques, les enjeux liés à leur ouverture, leur partage et leur réutilisation.

Exemples

- L'accès aux données liées aux accidents de la route peut permettre d'adapter et d'améliorer les politiques publiques de prévention.

La préfecture du Rhône publie chaque année des bilans de l'accidentologie dans le département et la région Auvergne-Rhône-Alpes

<https://www.rhone.gouv.fr/Actions-de-l-Etat/Transport-mobilite-securite-routiere-et-fluviale/Securite-routiere/Accidentologie-dans-le-Rhone-et-en-region-Auvergne-Rhone-Alpes>

- L'accès aux données liées aux horaires des lignes de transport en commun d'une ville peut permettre d'analyser les horaires de forte affluence et ainsi renforcer les passages si besoin.

La publication des horaires des lignes de transport en commun du réseau lyonnais a permis à l'application Citymapper d'enrichir les informations proposées à ses utilisateurs/utilisatrices et d'améliorer le service d'itinéraire proposé (exemple à retrouver sur le site de la Métropole de Lyon) : <https://data.grandlyon.com/portail/fr/reutilisations/64b6b9c159c22b003284734a>.

- L'accès aux données de certaines dotations de l'État pour financer des projets dans les territoires peut permettre d'identifier certaines disparités dans leur répartition et donner lieu à des rééquilibrages.

La publication des données liées au déploiement des projets France Relance sur le territoire de Bretagne permet de créer une carte interactive pour visualiser leur mise en œuvre (consulter la carte sur la plateforme GeoBretagne : <https://geobretagne.fr/datahub/dataset/5446b728-a13e-4b43-a0eb-9f37acd3ff77>

La vidéo présente ensuite trois enjeux de la politique d'ouverture des données publiques :

- pour **renforcer la transparence et la confiance** des citoyen·nes envers les institutions et les élu·élues. Cette exigence découle directement du principe défini à l'article 15 de la Déclaration des droits de l'homme et du citoyen : *La société a le droit de demander compte à tout agent public de son administration*. L'ouverture des données publiques et leur accès en open data répondent directement à cet enjeu.



C'est par exemple la publication rendue obligatoire par la Loi pour une République numérique, pour les collectivités territoriales de plus de 3500 habitants et employant plus de 50 agents, des documents qu'elles communiquent comme les arrêtés, délibérations, ou procès-verbaux des séances du conseil municipal.

- pour **améliorer les services publics** : l'ouverture des données publiques permet aux citoyens/citoyennes mais aussi aux entreprises et aux acteurs privés d'utiliser cette matière première pour créer de nouveaux services ou améliorer ceux existants. C'est par exemple le cas de l'ouverture des données épidémiologiques ayant permis la création de services et applications nées d'initiatives citoyennes comme :
 - **CovidTracker**, un outil de visualisation des données et statistiques liées à l'évolution de l'épidémie de Covid-19 en France et dans le monde,
 - **OpenCOVID**, un service de collecte et d'agrégation de données publiées par les autorités publiques ayant contribué à l'accélération de l'ouverture des données sur l'épidémie.
- pour **simplifier le service rendu** aux citoyens/citoyennes et usagers des services publics. C'est par exemple la concrétisation du principe « dites-le nous une fois » qui facilite le partage des données entre administrations pour éviter aux usagers de services publics de fournir leurs pièces à chaque nouvelle démarche. C'est le cas du département du Puy-de-Dôme : chaque année, les familles doivent fournir leur quotient familial au conseil départemental qui calcule le tarif applicable pour les services de cantines.
- Grâce à une interface, le conseil départemental utilise les données transmises par les particuliers à d'autres administrations pour calculer les tarifs de la cantine applicables aux familles (pour en savoir plus : <https://api.gouv.fr/guides/portail-famille-tarification>).

Bien évidemment, les usagers doivent donner leur accord pour que cet échange ait lieu, et seules les structures habilitées sont autorisées à consulter la base de données.



Vidéo - Les données personnelles sont-elles publiques ?

Cette vidéo présente les grands principes liés au traitement de nos données personnelles et de nos données sensibles.

Objectif 1 : permettre d'identifier et de différencier nos données personnelles et nos données sensibles.

La vidéo donne plusieurs exemples de ces types de données pour aider à prendre conscience de leur utilisation et de leur partage dans nos vies quotidiennes.

La vidéo introduit différentes notions relatives à la réglementation et à la protection de nos données personnelles et sensibles :

- le règlement général à la protection des données (RGPD), un texte qui encadre le traitement des données personnelles dans les pays de l'Union européenne et renforce le contrôle des citoyens sur l'utilisation de leurs données,
- la Commission nationale informatique et libertés (CNIL), une autorité administrative indépendante chargée de veiller à la protection des données personnelles contenues dans les fichiers et traitements informatique ou papier, aussi bien publics que privés. Elle s'assure également de l'application et de la bonne connaissance du RGPD.

Objectif 2 : rassurer et ouvrir le champ des possibles dans le partage des données, y compris à caractère personnel.

En effet, nos données personnelles font l'objet d'une protection renforcée, tout particulièrement quand les données sont identifiées



comme sensibles. Ce ne sont pas des données publiques, elles ne peuvent donc pas être rendues publiques, comme sur la plateforme data.gouv.fr, du moins pas avant d'avoir réalisé certains traitements (voir ci-après).

La vidéo présente aussi les cas où la collecte et l'utilisation des données personnelles sont autorisées. Il est en effet possible de diffuser des données à caractère personnel dans trois exceptions, prévues par la loi :

- si une disposition législative contraire autorise cette publication,
- si les personnes concernées par la publication ont donné leur accord,
- si les documents concernés appartiennent aux catégories identifiées par le Code des relations entre le public et l'administration permettant de les rendre publics sans processus d'anonymisation préalable (ex : autorisations individuelles d'urbanisme).

Rappel : depuis la loi 3DS du 21 février 2022, les administrations peuvent partager des données ou des informations entre elles lorsqu'elles sont strictement nécessaires et présentées par un usager pour améliorer le traitement de demandes (ex : transmission de données entre le service des impôts et les services des bourses de l'Éducation nationale).

Enfin, la vidéo aborde les notions d'anonymisation et de pseudonymisation, des traitements rendant impossible l'identification des propriétaires de données personnelles et permettant ainsi de les publier, les partager, les réutiliser :

- l'anonymisation est un traitement de données irréversible qui rend impossible l'identification de tout individu dans une base de données soit à partir d'une donnée, de groupes de données reliés entre eux ou encore à partir de la déduction d'une information à partir de données éparses mais concernant le même individu. Si le traitement par anonymisation est correctement réalisé, la législation relative à la protection des données ne s'applique plus,

- la pseudonymisation est un procédé réversible, par lequel on peut remplacer des données reliées à une personne permettant son identification (nom, prénom, etc.) par des données ne permettant son identification que de manière indirecte (alias, pseudo, etc.). Le traitement de données à caractère personnel existe cependant toujours et la législation relative à la protection des données continue de s'appliquer.

ASTUCE : distinguer les données à caractère personnel des données sensibles : quelques exemples.

Données à caractère personnel : il s'agit de données permettant d'identifier directement ou indirectement une personne physique. C'est par exemple l'adresse, le numéro de téléphone, le numéro de sécurité sociale, une photo, etc.

Données sensibles : il s'agit d'une catégorie de données à caractère personnel. C'est par exemple l'origine ethnique, les opinions politiques, convictions religieuses, les données concernant l'orientation sexuelle, etc. Le RGPD interdit leur collecte et leur traitement, sauf dans des cas très précis : si la personne a donné son accord explicite, si les données ont été rendues publiques par la personne concernée, si ces données sont nécessaires à la sauvegarde de la vie humaine, si leur utilisation est justifiée par l'intérêt public et validé par la CNIL, etc.

Pour consulter la liste complète, rendez-vous sur le site de la CNIL :

- <https://www.cnil.fr/fr/definition/donnee-sensible>
- <https://www.cnil.fr/fr/cnil-direct/question/une-donnee-caractere-personnel-cest-quoi>



Vidéo - Collectivités territoriales et open data, une histoire de transparence

Cette troisième vidéo illustre l'open data comme un enjeu de renforcement de la transparence de l'action publique mais également comme vecteur de l'amélioration des services publics.

La vidéo ne revient pas sur la définition de ces enjeux et des principes liés à l'open data. Elle présente un exemple concret d'ouverture des données des collectivités locales et d'accès pour les citoyens/citoyennes.

L'exemple identifié est celui de Mégalis, un syndicat mixte de coopération territoriale qui a mis à disposition des collectivités territoriales et des organismes publics des outils pour les accompagner dans la publication de certaines de leurs données (obligatoires).

Objectif 1 : incarner les principes de transparence de l'action publique et de confiance des citoyens/citoyennes, par la donnée.

Là encore, on suit Stéphane, un secrétaire de mairie, dans une de ses tâches quotidiennes. Il se sert des outils mis à disposition par Mégalis pour publier certaines données de sa commune : les délibérations du conseil municipal, les budgets.

La vidéo introduit un concept, celui de la marque blanche : c'est un service ou un produit créé par une entreprise ou un organisme et mis à disposition d'autres administrations ou utilisateurs pour leur propre compte. Par exemple, une même fabrique de parfums qui fournit des produits à différentes marques produit des parfums *en marque blanche*. La marque blanche proposée par Mégalis prend la forme d'un moteur de recherche des documents publiés en open data intégré librement sur le site des communes de la région.

Objectif n°2 : montrer les effets de l'ouverture des données publiques.

L'exemple de la Bretagne montre que les collectivités, mêmes celles non concernées par l'obligation réglementaire (communes de +3500 habitants uniquement), choisissent de partager en open data leurs documents et certaines de leurs données.

Les données sont rendues publiques pour les citoyens/citoyennes des communes concernées (sites internet) et de l'ensemble du territoire, sur la plateforme data.gouv.fr.

Accès aux délibérations des organismes adhérents de Mégalis Bretagne sur data.gouv :
<https://www.data.gouv.fr/fr/organizations/megalis-bretagne>.

Enfin, tous les développements du projet Mégalis sont disponibles en open source sur une communauté Github, une plateforme notamment ouverte aux projets de développement de logiciels libres. Sur Github, certaines collectivités font même des demandes d'évolution pour adapter l'outil proposé à leur situation. C'est le gage d'une transparence complète, jusque dans le déploiement du projet.

Depuis la création de la vidéo, l'interface de la marque-blanc a évolué pour prendre en compte l'extension de publication de tous les actes produits par les collectivités, et non plus seulement les délibérations, rendue obligatoire au 1^{er} juillet 2022.



Vidéo - La base adresse nationale : quand la donnée vous donne l'adresse

Cette vidéo, à travers les exemples des bases adresses locales et de la base adresse nationale, illustre les enjeux de l'ouverture et de la mise à jour continue des données (ici, les adresses). Il est ici question des adresses référencées sur l'ensemble du territoire et de leur traduction en données géolocalisables.

Le cas d'usage par lequel s'ouvre la vidéo permet de mesurer immédiatement l'intérêt du bon référencement des adresses dans une base nationale consolidée et à jour : grâce au bon référencement des adresses dans la base adresse locale, tout particulièrement des numéros attribués aux lieux-dits d'une commune, les services de secours peuvent intervenir rapidement, sans erreur de géolocalisation de la zone d'accident.

La base adresse nationale (BAN) est une base de données nationale alimentée à partir des bases adresses locales dans lesquelles les communes référencent leurs adresses. C'est une des 9 bases de données du service public de référence.

Site internet pour consulter la BAN : <https://adresse.data.gouv.fr/>.

Pour en savoir plus sur les bases de données du service public de référence : <https://www.data.gouv.fr/fr/pages/spd/reference/>.

Il s'agit d'un service numérique partagé et sur lequel s'appuient de nombreuses politiques publiques. C'est par exemple le cas du déploiement de la fibre optique qui s'appuie sur la BAN en tant que référentiel d'adresse unique.

Le fonctionnement de cette base de donnée illustre tout particulièrement des enjeux de l'open data puisqu'il s'agit d'une base de données :

- ouverte : accessible à tous et toutes et dont le code source est lui-même ouvert,
- fiable : elle est mise à jour en continu par les communes via leurs bases adresses locales, par d'autres bases de données nationales et, bientôt, par les citoyens/citoyennes qui pourront signaler une évolution souhaitée sur une adresse.

Le saviez-vous ? Les communes possèdent depuis toujours la compétence du nommage. La loi 3DS leur impose d'inscrire ces noms (lieux-dits, rue, etc.) dans une base de données unique, la base adresse locale, qui alimente la BAN.

Loi n°2022-217 du 21 février 2022 relative à la différenciation, la décentralisation, la déconcentration et portant diverses mesures de simplification de l'action publique locale : article 169.



Vidéo - La base Géo'DAE : quand l'open data sauve des vies

Cette vidéo illustre un exemple de partage et de réutilisation de données dans le domaine de la santé : la localisation des défibrillateurs automatiques externes (DAE) sur le territoire (établissements recevant du public ou ERP).

Le cas d'usage présenté dans la vidéo illustre comment la donnée, renseignée et ouverte sur Géo'DAE, peut contribuer à sauver des vies : en orientant rapidement les personnes témoins d'un arrêt cardiaque vers le DAE le plus proche, elles peuvent dispenser les gestes de premier secours en attendant l'arrivée des services de secours sur place.

La base nationale Géo'DAE assure la mise en œuvre de la loi du 28 juin 2018 relative au défibrillateur cardiaque rendant obligatoire la déclaration des DAE et leur maintenance sur une base unique et accessible en open data.



Cet exemple permet d'illustrer comment la donnée accompagne et facilite la mise en œuvre des politiques publiques. Dans ce cas, il s'agit d'accompagner la mise en œuvre du plan national de santé publique prévoyant de former 80 % de la population aux gestes de premiers secours et d'augmenter le recours à un DAE en cas d'arrêt cardiaque.

Tous les établissements disposant d'un DAE doivent le déclarer sur Géo'DAE et garantir sa maintenance. La loi fixe également un format standard pour la déclaration des données, afin de fiabiliser le schéma de transmission des données et de faciliter la création de services. Grâce à l'ouverture et au partage des données, il est ici possible de repérer un DAE dans l'espace public, de connaître le taux d'équipement sur le territoire, de signaler des équipements défectueux et ainsi améliorer les chances de survie en cas d'arrêt cardiaque.

Les données sont disponibles en open data, notamment sur le site data.gouv.fr et via des supports de recensement de DAE, parmi lesquels des applications mobiles et des sites internet (partage de données ouvertes = création de services, appli, etc.).

La Métropole de Toulouse propose également une extraction du jeu de données de la base Géo'DAE sur le périmètre métropolitain, avec une visualisation sur une carte en ligne : <https://data.toulouse-metropole.fr/explore/dataset/defibrillateurs/map/?location=11,43.65829,1.39891&basemap=jawg.streets>.



Vidéo - Data et IA : des outils au service des agents publics (en 2 parties)

Cette vidéo aborde les enjeux des liens entre la data et l'intelligence artificielle. Elle définit les termes du sujet, ses enjeux et donne à voir les cas d'usage de l'IA dans les services publics et l'administration,

en se plaçant du point de vue des agents. Elle est disponible en 2 parties, visionnables l'une après l'autre, afin de limiter la quantité d'informations transmise au public (durée initiale de la vidéo : 3'45 minutes).

Objectif n°1 de la vidéo : faire prendre conscience que l'IA ne date pas d'hier et qu'elle est déjà présente partout (ou presque) dans nos vies quotidiennes.

La vidéo l'illustre par quelques exemples de la vie quotidienne : la reconnaissance faciale pour déverrouiller nos téléphones (existe depuis 2017 sur l'iPhone), les algorithmes de suggestion de lecture, playlists ou de publicités ciblées sur les sites marchands, réseaux sociaux ou applications musicales, etc.

La vidéo ne le mentionne pas mais l'IA révolutionne déjà depuis quelques années le monde médical. En radiologie, les algorithmes d'IA analysent des données massives sur des radios pour détecter plus rapidement et de façon plus fiable les anomalies. En oncologie, l'IA peut aider à croiser plusieurs sources de données, comme les antécédents médicaux, pour poser un diagnostic précis et améliorer la prise en charge. Des systèmes d'IA sont aussi utilisés en médecine généraliste pour retranscrire les comptes rendus (reconnaissance vocale) ou aider au diagnostic et aux prescriptions de médicaments.

Vous pouvez trouver plus d'informations sur le sujet ici :

<https://www.inserm.fr/dossier/intelligence-artificielle-et-sante/>

Objectif n°2 : expliquer ce qu'est l'IA de façon claire et simple.

La vidéo ne donnera pas une vision très technique de ce qu'est l'IA. Pour cela, des ressources complémentaires sont proposées à la fin de ce guide. En revanche, elle permet d'expliquer simplement comment certains types d'IA fonctionnent, notamment par apprentissage supervisé. Elle explique aussi pourquoi on entend beaucoup plus parler d'IA depuis les années 2000 et 2020 : production et analyse de quantités de données de plus en plus massives (big data), capacités de calcul décuplées, etc.



Exemple d'une ressource complémentaire accessible pour comprendre toutes les formes d'IA et leur fonctionnement :

<https://cafeia.org/index.php/qu-est-ce-que-l-ia/>

Objectif n°3 : illustrer les usages de l'IA déjà en cours dans les services publics et les démythifier pour les agents publics

La vidéo illustre différents usages de l'IA déjà mis en œuvre dans les administrations et les services publics :

- **transports** : l'IA pour informer sur l'affluence à bord des transports en commun ; on peut utiliser différentes sources de données comme le pointage des cartes de transport, l'achat de tickets à la borne, le poids à l'essieu, les informations transmises par les passagers en direct pour ensuite prédire les périodes d'affluence à bord et donner des indications aux voyageurs et voyageuses,
- **connaissance du territoire** (occupation du sol à grande échelle) : l'IA peut aider à modéliser, à cartographier avec plus de précision le territoire, en s'entraînant à distinguer les différents types d'objets : bâti, végétation, types de sols, etc. Elle peut ensuite les cataloguer précisément et croiser les résultats avec d'autres sources de données pour améliorer la connaissance de tout le territoire,
- **automatisation des tâches répétitives, réduction des opérations chronophages, aide à la décision** : l'IA est utilisée pour synthétiser, résumer des documents et proposer des brouillons de réponses, comme supports d'aide à la décision. L'IA peut ainsi permettre aux agents de dégager du temps pour se concentrer sur d'autres tâches : conseil, négociation, travail collaboratif, etc.

Sur les transports en commun, voir l'article suivant :

<https://www.keolis.com/ces-reseaux-qui-predisent-l-affluence-a-bord/>

Sur la connaissance du territoire, consulter l'Atlas 2024 de l'IGN :

<https://www.ign.fr/atlas-ign-des-cartes-de-lanthropocene-2024-intelligence-artificielle>

Sur l'automatisation des tâches répétitives pour les agents : voir par exemple l'utilisation d'Albert dans les France Services : <https://www.france-services.gouv.fr/actualites/experimentation-dun-modele-dassistance-france-services-IA>

IA et chatbot, c'est pareil ?

Les chatbots, ou *agent conversationnel*, ça ne vous dit rien ? Mais si ! Vous en utilisez ou en croisez dans votre navigation sur internet presque tous les jours : sur le site de La Poste ou de la SNCF. Un chatbot est un programme avec lequel vous pouvez interagir dans une fenêtre de dialogue. Il est programmé pour apporter des réponses à des requêtes, sur des règles paramétrées et préétablies. L'utilisateur ou utilisatrice interagit directement avec le chatbot dans un espace conversationnel. Certains chatbots peuvent intégrer des systèmes d'apprentissage autonomes, qui s'apparentent à de l'IA, et vous donner des réponses personnalisées. Chat GPT est un chatbot intégrant de l'IA avec des mécanismes d'apprentissage autonome. L'outil va ensuite enrichir sa base de données au fur et à mesure des requêtes (ou prompts) formulées. D'autres chatbots n'intègrent pas nécessairement des modèles d'IA et sont paramétrés, mis à jour et améliorés par les agents directement.

Exemples de chatbots utilisés par des administrations :

- Rebecca, assistante virtuelle pour aider les gestionnaires RH à accéder à la réglementation RH adaptée. Le référentiel de Rebecca est mis à jour tous les mois par les équipes du CISIRH qui la développent pour intégrer de nouvelles règles de gestion à ses réponses. En savoir plus : <https://www.bercynumerique.finances.gouv.fr/le-chatbot-du-cisirh-est-desormais-accessible-tous-sur-internet>,
- Chatbot disponible sur le site <https://www.masecurite.interieur.gouv.fr> pour échanger anonymement avec des agents des forces de l'ordre dans le cadre de vos démarches en ligne.

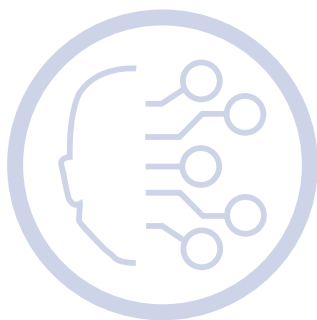


Objectif n°4 : pas d'IA sans data

Pourquoi produire une vidéo sur l'IA dans le dispositif DATAcculturation ? Compléter le dispositif par cette vidéo émane d'un réel besoin afin de répondre aux questions des agents publics sur les différents usages des données. L'IA est un sujet d'actualité pour les agents, aussi bien dans leurs vies professionnelles que personnelles. Il a dès lors semblé important d'aborder le sujet de l'IA sous l'angle des liens indéfectibles qu'elle entretient avec les données.

À la base de tout système d'IA, il y a des données : ce sont elles qui lui permettent de s'entraîner et de réaliser les tâches pour lesquelles elle est programmée. Sans données : pas d'IA. Du moins, pas d'IA fiable, qui limite les risques d'erreur, etc. Connaître les données, ce qu'elles sont, comment elles fonctionnent, leur production et leur partage par les administrations, est essentiel pour comprendre comment l'IA fonctionne et ce qu'elle peut nous apporter.

Pour les services publics, maîtriser les données publiques produites par les administrations, leur qualité, leur fiabilité et leur accessibilité, c'est s'assurer de mieux maîtriser les systèmes des IA souverains.



D'AUTRES RESSOURCES DISPONIBLES POUR ALLER (ENCORE) PLUS LOIN

Le jeu sérieux *Data City*

Ce jeu sérieux en ligne permet d'améliorer ses connaissances sur l'open data en s'amusant. Il propose d'incarner un personnage : le ou la maire d'une petite ville. Vous allez ensuite devoir répondre aux questions que les habitants/habitantes de votre commune se posent en matière d'open data. En y répondant, vous débloquentes des *datapoints* et contribuez au développement de la ville. Ce jeu a été créé par les élèves de l'école de design *ActivDesign*. Il a été conçu pour acculturer les élus/élues des collectivités territoriales aux enjeux de l'open data. Il est disponible gratuitement : <https://datacity.sib.fr/>

ASTUCE : organisez une séance de découverte sur les collectivités territoriales et l'open data. Vous pouvez par exemple conseiller à vos collègues de visualiser la vidéo sur l'open data des collectivités et le projet de Mégalis Bretagne et les inviter ensuite à prendre en main le jeu *Data City*. Pour plus de convivialité, testez le jeu à plusieurs et comparez le développement de vos communes !

Le jeu sérieux *Aïe Aïe IA*, pour démystifier l'intelligence artificielle

Ce jeu sérieux, co-conçu par la Métropole de Lyon, permet de découvrir en s'amusant les thématiques liées à l'usage de l'intelligence artificielle. En environ 1 heure de jeu, vous participerez à une véritable enquête et jeu de rôle grandeur nature, en immersion dans une entreprise fictive.





Cerise sur le gâteau : tous les supports du jeu sont disponibles en open data en téléchargement libre et gratuit.

Il suffit de bien manier votre paire de ciseaux et d'imprimer les cartes et guides d'animation !

Pour consulter toutes les ressources : <https://lesbases.anct.gouv.fr/ressources/aie-aie-ia-le-jeu-pour-demystifier-l-intelligence-artificielle>.

Attention : 8 joueurs ou joueuses maximum par boîte.

ASTUCE : organisez un atelier ludique autour de la donnée et de l'IA !

Si vous avez environ 2 heures, vous pouvez :

- organiser un café de la data sur la thématique data et IA (visionnage des vidéos correspondantes) – environ 45 min,
- faire découvrir le jeu Aïe Aïe IA à l'issue du café de la data pour approfondir de manière ludique la découverte du sujet.



Comprendre les liens entre IA et données et IA et intelligence humaine grâce à la pop culture.

La série *Rematch*, disponible sur Arte, raconte l'histoire du système d'intelligence artificielle *Deep Blue* mis au point par IBM pour affronter le champion international d'échecs, Garry Kasparov. Elle illustre bien les liens indéfectibles qui existent entre les données massives et tout système d'IA mais aussi entre l'intelligence humaine, la force de travail des humains, et la machine. Si IBM n'avait pas dégagé des moyens humains, financiers et techniques colossaux pour améliorer *Deep Blue* en continu, peut-être n'aurait-elle pas gagné face à Kasparov en 1997.

2024, 1 saison, 6 épisodes, produite par Grégoire Auger

BIBLIOGRAPHIE THÉMATIQUE

Ressources complémentaires sur les techniques d'animation innovantes et de facilitation d'ateliers

Si vous voulez enrichir les modes d'animation des cafés de la data ou apprendre d'autres techniques d'animation en intelligence collective, voici une liste non-exhaustive de ressources disponibles en libre accès :

- **guide Oser l'innovation RH** pour concevoir et animer un atelier de travail, disponible sur le portail de la fonction publique. Pour télécharger le livret : rendez-vous sur le site : fonction-publique.gouv.fr,
- **guide Utilo**, un recueil co-construit par la communauté du Tilab, laboratoire d'innovation publique de l'État en Bretagne. Pour télécharger le guide, rendez-vous sur le site : modernisation.gouv.fr.

Ressources complémentaires sur les vidéos

La donnée et l'open data : la donnée par et pour le service public.

- *Pour une politique publique de la donnée*, rapport de la mission Bothorel, remis au Gouvernement en décembre 2020, disponible en consultation et en téléchargement.
- Fiches ressources Opendata France *Fiches pédagogiques : pour Comprendre*, 2019 : <https://www.opendatafrance.net/ressources/documents-opendata-france>.
- Données, algorithmes et codes sources : les 15 feuilles de route ministérielles (2021).
- La transparence de l'action publique, webinaire *Le mois de la data* par OpendataFrance.
- Ouverture des données : la CNIL publie son cahier air 2021, article du 11 mars 2022 : <https://www.cnil.fr/fr/ouverture-des-donnees-la-cnil-publie-son-cahier-air2021>.



- Espace collaboratif *Atelier données et territoires*, créé par Open-dataFrance et le département numérique de la Fédération nationale des collectivités concédantes et régies : <https://programme-tnt.gitbook.io/donnees-territoires/>.
- *L'impact des données ouvertes* : 6 articles publiés sur data.gouv pour comprendre les impacts, dans différents secteurs ou périmètres de la société, de l'ouverture et du partage des données : <https://www.data.gouv.fr/fr/posts/impact-des-donnees-ouvertes-1-6>.
- *Data et territoires* rapport de la mission Data et territoires, remis au ministre de la Transformation et de la fonction publique en septembre 2023, consultable ici : https://www.transformation.gouv.fr/files/ressource/Rapport_Mission_Data_Territoires.pdf
- Les vidéos partagées sur la page Youtube de Numérique en commun[s] (NEC) : vous y trouverez des contenus variés et très riches sur les sujets de la donnée et du numérique en général : <https://www.youtube.com/@numeriqueencommuns2389/videos>.
- Sur les données massives (big data) : lettre de la DAJ du Ministère de l'économie et des finances sur l'Explosion des données numériques (2023) à lire ici : <https://www.economie.gouv.fr/daj/lettre-de-la-daj-explosion-des-donnees-numeriques>.

Les données personnelles sont-elles publiques ?

- Article 9 du Règlement général sur la protection des données.
- Rectificatif au règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), JOUE L119 du 4 mai 2016.
- Article « RGPD : de quoi parle-t-on ? » disponible sur le site de la CNIL : <https://www.cnil.fr/fr/rgpd-de-quoi-parle-t-on>.
- Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

- Guide pratique de la publication en ligne et de la réutilisation des données publiques (open data), CNIL et Cada, 2019.
- Article *Échanges de données : ce que change la loi 3DS* : <https://www.lagazettedescommunes.com/792424/echange-de-donnees-et-api-ce-que-change-la-loi-3ds/?abo=1>.
- L'anonymisation de données personnelles, article CNIL : <https://www.cnil.fr/fr/lanonymisation-de-donnees-personnelles>.
- Le rôle du délégué ou de la déléguée à la protection des données : <https://www.cnil.fr/fr/designer-un-pilote>.

Collectivités territoriales et open data, une histoire de transparence.

- Le projet expliqué sur le site de Mégalis : <https://services.megalis.bretagne.bzh/service/opendata-2/>.
- Article sur le site de la CNIL, *Les collectivités territoriales et l'open data : concilier ouverture des données et protection des données personnelles*, 2019.
- Article sur le site de la CNIL, *La publication des documents des collectivités territoriales liés à l'exercice de leur pouvoir décisionnaire*, 2021.

La base adresse nationale : quand la donnée vous donne l'adresse.

- *La base adresse nationale (BAN) franchit de nouvelles étapes en poursuivant son action au sein de l'IGN*, communiqué de presse du 17 mars 2022.
- Pour en savoir plus sur l'aménagement numérique du territoire en lien avec la BAN : <https://www.aménagement-numérique.gouv.fr/fr/referentiel-dadresses-commun-base-adresse-nationale>.
- Article 169 de la Loi n° 2022-217 du 21 février 2022 relative à la différenciation, la décentralisation, la déconcentration et portant diverses mesures de simplification de l'action publique locale.

La base Géo'DAE : quand l'open data sauve des vies.

- Dossier sur les données de santé, site de la CNIL : <https://www.cnil.fr/fr/quest-ce-que-une-donnee-de-sante>.
- La base nationale Géo'DAE : <https://geodae.atlasante.fr/apropos>.

Vidéo Data et IA : des outils au service des agents publics (en 2 parties)

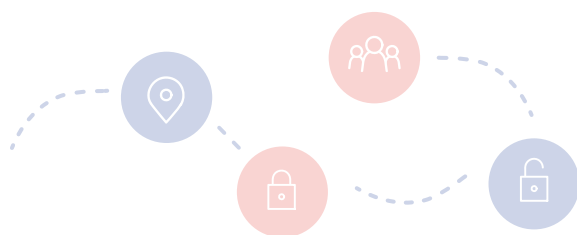
Des ressources généralistes sur l'IA

- Pour comprendre comme l'intelligence artificielle fonctionne, quelle est son histoire et trouver des exemples d'utilisation, découvrez le module pédagogique *Qu'est-ce-que l'IA ?* sur le site des Cafés IA : <https://cafeia.org/index.php/qu-est-ce-que-l-ia>.
- Plus généralement, vous pouvez consulter toutes les ressources disponibles sur le site <https://cafeia.org>.
- Vidéo *L'intelligence artificielle en 90 secondes*, Réseau Canopé : <https://www.youtube.com/watch?v=1Rsj8mfQEel>.
- Podcast *Les dames de l'algorithme*, série *Le code a changé* de France culture, janvier 2024.
- Podcast *Faut-il avoir peur de l'intelligence artificielle ?*, émission *Le débat de midi*, France inter, juillet 2024.
- *Intelligence artificielle : le cadre juridique européen de l'IA en six questions*, site vie-publique.fr, juin 2024.
- *Intelligence artificielle, de quoi parle-t-on ?*, CNIL, 2022 : <https://www.cnil.fr/fr/intelligence-artificielle>.
- *IA dans les services publics : comment mieux protéger les droits des usagers ?*, article sur vie-publique.fr, novembre 2024 : <https://www.vie-publique.fr/en-bref/296127-ia-dans-les-services-publics-et-droits-des-usagers>.
- *Comprendre les enjeux de l'intelligence artificielle au prisme de l'action publique*, entretien avec Arno Amabile, Profil Public, septembre 2024.

- Rapport *Intelligence artificielle : notre ambition pour la France*, Commission de l'intelligence artificielle, mars 2024.
- Rapport *L'intelligence artificielle dans les politiques publiques : l'exemple du ministère de l'Économie et des finances*, Cours des comptes, 2024.
- Rapport *Algorithmes, systèmes d'IA et services publics : quels droits pour les usagers ?*, La Défenseure des droits, novembre 2024.

Des exemples et cas d'usage de l'IA

- Consultez l'Atlas de l'IGN *Cartographier l'anthropocène à l'ère de l'intelligence artificielle* (septembre 2024), pour découvrir des usages de l'IA pour mieux cartographier le territoire, anticiper les risques, etc. : <https://www.ign.fr/atlas-ign-des-cartes-de-lanthropocene-2024-intelligence-artificielle>.
- Guide *Stratégie d'usage de l'intelligence artificielle en matière de gestion des ressources humaines dans la fonction publique d'État*, DGAFP, juin 2024.
- Sur l'intelligence artificielle et la transition écologique : <https://greentechinnovation.fr/les-acteurs-de-lia/> et <https://reporter-terre.net/L-insoutenable-cout-ecologique-du-boom-de-l-ia>.



LEXIQUE

Licence ouverte Etalab

Il s'agit d'une licence conçue par le service Etalab (gouvernement français), élaborée avec l'ensemble des acteurs concernés, dont le but est de faciliter et d'encourager le partage et la réutilisation des données publiques mises à disposition gratuitement.

Facilitateur ou facilitatrice

Son rôle est de favoriser les échanges et le partage des idées dans un cadre de travail collaboratif bienveillant.

Open data

Définition proposée par la CNIL : « c'est le mouvement d'ouverture et de mise à disposition des données produites et collectées par les services publics. »

RGD

Pour « Règlement Général à la Protection des Données ». Il s'agit d'un texte qui encadre le traitement des données personnelles dans les pays de l'Union européenne et renforce le contrôle des citoyens de l'utilisation de leurs données.

Référence : règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).

Ice breaker

Ou en anglais littéralement « brise-glace ». Il s'agit d'une activité dont le but est de mettre en confiance les participants d'un groupe de travail ou une séquence collaborative. En brisant la glace, souvent de façon ludique et décalée, les membres du groupe apprennent à se connaître avant de travailler ensemble.

REMERCIEMENTS

Ce livret a été rédigé en 2022 par **Justine Sperandio-Martinez**, chargée de projet innovation numérique et open data au Secrétariat général pour les affaires régionales (SGAR) Auvergne-Rhône-Alpes. Elle a été accompagnée par un comité de rédaction composé de :

- **Benoît Duée**, conseiller en organisation du travail au sein de la Plateforme régionale des Ressources Humaines (PFRH) Occitanie
- **Frédérique Rehfeld**, cheffe Service interacadémique de la modernisation de l'action publique (Siamap) Auvergne-Rhône-Alpes

Le contenu du livret a fait l'objet de plusieurs séances de relecture par le comité de rédaction et avec la contribution de relecteurs/relectrices. Nous les remercions vivement pour leurs retours constructifs.

En 2024, le guide a été mis à jour et enrichi de nouveaux contenus. Nous remercions les animateurs et animatrices de cafés de la data ainsi que l'ensemble des collègues qui ont participé à la relecture, correction et à l'enrichissement du guide, parmi lequel/lesquelle : **Claire Soriano, Bertrand Giroux, Élodie Bernard, Margot Godefroi, Lisiane Verrier, etc.**

Nous remercions tout particulièrement les **groupes d'élèves de la promotion Gerty Archimede de l'IRA de Lyon** ayant testé les différents formats d'animation des cafés de la data proposés dans le livret, en septembre et octobre 2022. Le livret a ainsi été enrichi de leurs retours.

Enfin, la mise en page a été confiée aux agents du service de la communication du rectorat de l'académie de Clermont-Ferrand : **Béatrice Humbert**, cheffe de service, et **Géraldine Goetz**, technicienne d'édition et de graphisme.

Nous les remercions d'avoir fait prendre vie à ce livret !

Nous remercions également l'ensemble des partenaires ayant travaillé sur le projet depuis ses débuts. Sans vous, le projet n'aurait pas pu voir le jour !

Les partenaires

Ministère de l'Intérieur
Ministère des Armées
Ministère de l'Éducation
Préfecture de région Bretagne
Préfecture de région Île-de-France
Préfecture de région Occitanie
Académie de Clermont-Ferrand
Direction générale des finances publiques (DGFIP)
Institut régional d'administration (IRA) de Lyon
Agence régionale de santé Auvergne-Rhône-Alpes

Institut national de l'information géographique et forestière (IGN)

Mégalis Bretagne
Métropole de Lyon
Communauté d'agglomération du Pays-Basque

Conseil national du numérique



**DATAcculturation est un projet lauréat
du Fonds d'innovation RH (FIRH) 2022.**

DATAculturation

Sensibiliser à la donnée, à l'IA
et animer un café de la data

Vous avez **envie de partager votre curiosité** et vos **connaissances sur la data et/ou l'intelligence artificielle** avec vos collègues ou dans votre service ?

Les cafés de la data sont faits pour vous !

Ce livret permet de **concevoir** et d'**animer** des sessions de présentation participatives qui vous permettront d'aborder des sujets relatifs à la donnée et à l'open data avec vos collègues.

Vous y trouverez des **conseils** et des **méthodes concrètes** pour élaborer vos cafés de la data.

Vous accéderez également à des ressources complémentaires pour enrichir ces temps d'animation et vos connaissances.

