



**HAL**  
open science

## Données et recherches participatives

Kenneth Maussang, Nicolas Larousse, Thomas Jouneau, Hélène Jouguet,  
Jean-François Martin

► **To cite this version:**

Kenneth Maussang, Nicolas Larousse, Thomas Jouneau, Hélène Jouguet, Jean-François Martin. Données et recherches participatives. Printemps de la Donnée 2024, INRAE; Université Haute-Alsace; Université de Strasbourg; INSA; PNDB; AgroParisTech; Université de Lille; Sorbonne Université; Data Terra, Mar 2024, Montpellier, France. hal-04681965

**HAL Id: hal-04681965**

**<https://hal.inrae.fr/hal-04681965v1>**

Submitted on 30 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Données et recherches participatives

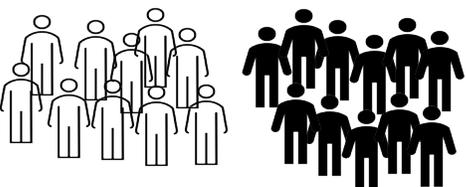
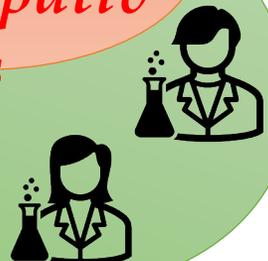
*Enjeux et recommandations issues d'exemples de projets de recherches participatives*

Kenneth Maussang, Nicolas Larrousse

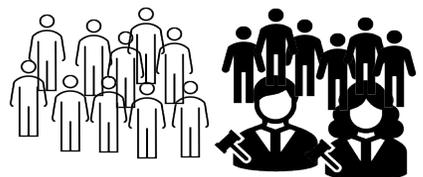
Printemps de la donnée – 19 mars 2024

# Recherches participatives

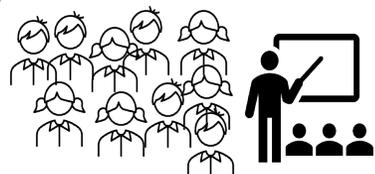
Recherches Participatives



Associations



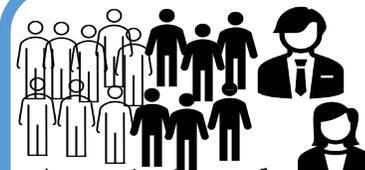
Associations de victimes



Scolaires



Associations de malades



Associations de professionnels



Individus



R&D  
Innovation ouverte

*"(...) des formes de production de connaissances scientifiques auxquelles des acteurs non-scientifiques professionnels – qu'il s'agisse d'individus ou de groupes – participent de façon active et délibérée."*

François Houllier, *Sciences Participatives en France. Etats des lieux, bonnes pratiques et recommandations*, (Février 2016) <https://hal.inrae.fr/hal-02801940>.

S. R. Arnstein, *A Ladder of Citizen Participation*, JAIP, Vol. 35, 4, July 1969, pp. 216-224.

# Une recherche scientifique « hors les murs »

Forme de science ouverte où des non-professionnels sont actifs dans le processus de recherche, où **la donnée est la matière première de l'échange avec les chercheurs.**

Implication de non-professionnels dans le processus de recherche : **motivations ?**

Impact sur la connaissance produite : **fiabilité ? crédibilité ?**



**OS** Ouvrir  
la science !

Comité pour la Science Ouverte  
*collège Données de la recherche*

Recommandations générales pour  
aborder un projet de recherches  
participatives

# Méthodologie

## Une enquête nationale en ligne + série d'entretiens ciblés

### Prise en compte de l'hétérogénéité :

- des pratiques ;
- des disciplines ;
- des métiers ;
- des opinions.

Le tout sous l'angle exclusif des **données de la recherche**.

### Chiffres clés

- ✓ 5 personnes
  - Hélène Jouguet (IE CNRS – Huma-Num)
  - Thomas Jouneau (Cons. bib. – Univ. Lorraine)
  - Jean-François Martin (PR Institut Agro)
  - Kenneth Maussang (MdC Univ. de Montpellier)
  - Nicolas Larrousse (IR CNRS – Huma-Num)
- ✓ **Enquête nationale** (Déc. 21 – Jan. 22)  
359 réponses complètes (enquête)
- ✓ **9 entretiens ciblés** (Mai 22 – Jan. 23)

# Données, informations et connaissances

## **Compétence :**

*« mobilisation ou l'activation de plusieurs savoirs, dans une situation et un contexte données »* Guy le Boterf

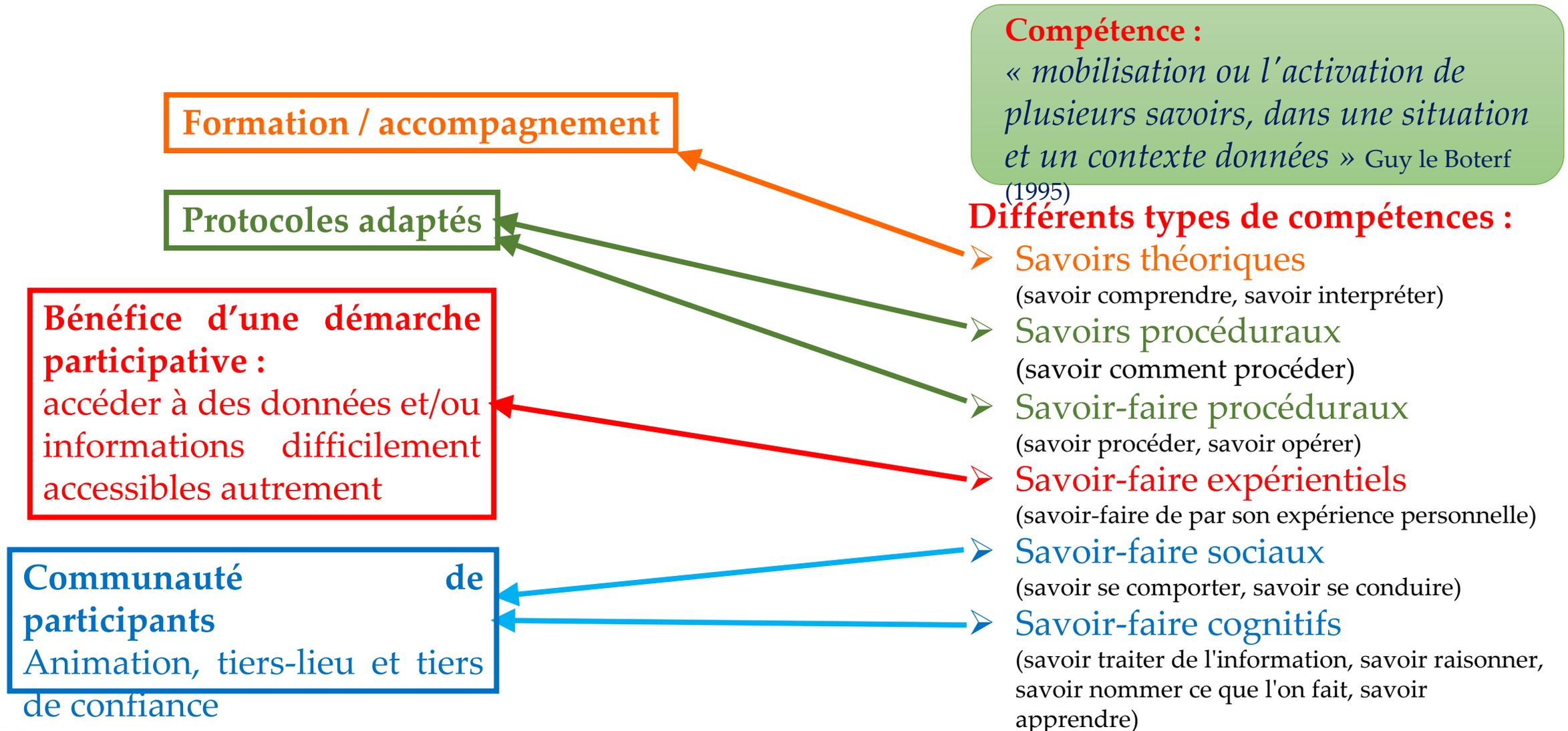
(1995)

## **Différents types de compétences :**

- **Savoirs théoriques**  
(savoir comprendre, savoir interpréter)
- **Savoirs procéduraux**  
(savoir comment procéder)
- **Savoir-faire procéduraux**  
(savoir procéder, savoir opérer)
- **Savoir-faire expérimentiels**  
(savoir-faire de par son expérience personnelle)
- **Savoir-faire sociaux**  
(savoir se comporter, savoir se conduire)
- **Savoir-faire cognitifs**  
(savoir traiter de l'information, savoir raisonner, savoir nommer ce que l'on fait, savoir apprendre)

Guy le Boterf (1995)

# Données, informations et connaissances



Guy le Boterf (1995)

# Données, informations et connaissances

Non-professionnels scientifiques, **le niveau de compétence des participants doit être évalué** :  
Scores de confiance dynamique / stratégies qualité

**Formation / accompagnement**

**Protocoles adaptés**

**Bénéfice d'une démarche participative :**  
accéder à des données et/ou informations difficilement accessibles autrement

**Différents types de compétences :**

- **Savoirs théoriques**  
(savoir comprendre, savoir interpréter)
- **Savoirs procéduraux**  
(savoir comment procéder)
- **Savoir-faire procéduraux**  
(savoir procéder, savoir opérer)
- **Savoir-faire expérientiels**  
(savoir-faire de par son expérience personnelle)
- **Savoir-faire sociaux**  
(savoir se comporter, savoir se conduire)
- **Savoir-faire cognitifs**  
(savoir traiter de l'information, savoir raisonner, savoir nommer ce que l'on fait, savoir apprendre)

**Communauté de participants**  
Animation, tiers-lieu et tiers de confiance

Guy le Boterf (1995)

# Données, informations et connaissances

Formation / accompagnement

Protocoles adaptés

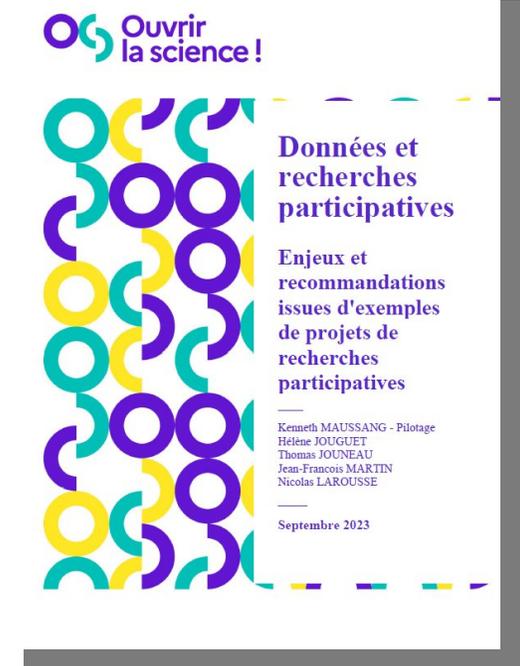
Bénéfice d'une démarche participative : accéder à des données et/ou informations difficilement accessibles autrement

Communauté de participants  
Animation, tiers-lieu et tiers de confiance

Non-professionnels scientifiques, le niveau de compétence des participants doit être évalué : Scores de confiance dynamique / stratégies qualité



<https://dx.doi.org/10.52949/33>



<https://dx.doi.org/10.52949/38>

<https://dx.doi.org/10.52949/39>



# Principales conclusions

Première question à se poser : **quel intérêt d'une démarche participative ?**

**Relation de confiance et/ou tiers de confiance et/ou tiers lieu**  
Asymétrie d'information/compétence entre participant/chercheur.

**Le temps**

**La communauté**

**La donnée**

**Qualité de la donnée ≠ Fiabilité de la donnée ≠ Crédibilité de la donnée.**

**L'outil**

**Importance de l'ergonomie.**

# Stratégie qualité

**Nécessite de maintenir le lien entre une donnée et**

- son (ses) contributeur(s) – donc leur identité (dans le respect du RGPD) ;**
- les traitements appliqués et les codes correspondants ;**
- l'échantillon ou l'objet d'étude associé ;**
- le protocole associé ;**
- ~~la documentation du hardware (ou références si commercial).~~**

# Ouverture du processus de recherche, ouverture des données

**Le partage des données (ouverture) permet d'élever la qualité**

Les données étant immédiatement partagées par les pairs, il y a une forme d'exigence mutuelle de la part des producteurs de données qui se met en place.

**L'ouverture des données est un facteur de renforcement de la motivation des participants (donner à « voir » leur contribution).**

# Valeur ajoutée - Valorisation des données

## Exploitation des données / des outils développés :

- participants ;
- tiers partenaires ;
- politiques publiques ;
- outils de gestion de crise.

**Importance de la reproductibilité pour la crédibilité institutionnelle.**

**Accéder à des données difficilement accessibles autrement / obtenir des résultats difficilement réalisables dans un contexte « classique ».**

**Outils de diffusion de la culture scientifique, mais surtout d'acculturation à la démarche scientifique.**

**Appropriation de l'objet de recherche par « la société », devenant un objet culturel.**



*Festival international de Théâtre de rue d'Aurillac (ECLAT) 2023*

# Pour aller plus loin...



<https://dx.doi.org/10.52949/33>



<https://dx.doi.org/10.52949/38>  
<https://dx.doi.org/10.52949/39>



***Birds of a Feather (BoF) à la 21<sup>ème</sup> plénière RDA***

*Collecting and Managing Data from Citizen Science Research Projects*

# Répliquabilité / reproductibilité

## **Evaluation de la répliquabilité des protocoles par chaque participant**

- ✓ **hétérogénéité** : hardware, niveau de compétence, implication
- ✓ **accessibilité et robustesse des protocoles** : bêta-tests, phasage de « recrutement » avec pool de rodage,  
co-construction des protocoles avec les participants (e.g. Inserm – Gram)
- ✓ **diversité des supports** : formation, protocoles, outils de communication (compétences numériques)
- ✓ **vocabulaire commun/partagé**, discours au niveau de connaissances et compétences des participants.

## **Evaluation de la reproductibilité de la production des données**

Capacité à **produire des données similaires par plusieurs participants dans des conditions similaires.**

## **Evaluation de la reproductibilité de l'analyse des données**

Capacité à **produire des résultats identiques ou cohérents quand des données similaires sont analysées par le même participant.**

## **Evaluation de la qualité de la réalisation de la tâche confiée au participant**

Capacité à **produire des résultats ou des données permettant de répondre à la question de recherche.**

## **Evaluation de la tolérance à l'erreur / aux incertitudes expérimentales**

# Stratégie qualité

## ✓ Contrôle statistique et/ou moyennage

- compenser un taux d'erreur plus élevé par un moyennage statistique (volume de donnée) ;
- contrôle aléatoire par un professionnel ;
- échantillonnage biaisé (réseau CiSStats – méthodes statistiques pour les données de recherches participatives, <https://cisstats.mathnum.inrae.fr>)

## ✓ Scores de confiance dynamique

- jeu de données calibré pour évaluer le participant ;
- distribution d'un jeu de donnée sur  $N$  participants différents ;
- transcriptions : comparaison des résultats sur des segments communs du corpus ;
- identification des participants défaillants (modalités d'exclusions), des données suspectes (*fake data*, trolls)

# Stratégie qualité

## ✓ Plan de gestion de données

à prévoir dès le début du projet (pas de spécificités aux projets de recherches participatives *a priori*)

## ✓ Importance de la formation, de l'accompagnement

(e.g. accréditations/certification pour l'accès à certaines données)

## ✓ Importance des protocoles et de leur robustesse

- Protocoles fermés, les plus bornés possibles.

*Inconvénient* : le participant risque de n'être qu'un « exécutant ».

- Equilibre entre la rigueur scientifique et créativité / intelligence collective des participants

## ✓ Plan de gestion de données

Outil de réflexion et d'anticipation parfaitement adapté.

## ✓ Historicisation et traçabilité

- **tout conserver : données brutes, traitements appliqués ;**

- y compris en cas d'**erreur** (mais le signaler) – qualification de la donnée (normale, anormale, erreur,...) ;

- possibilité de commenter les données ;

- possibilité d'indiquer « *je ne sais pas* » (qui est une information en soit)

# Outils

## ✓ Importance des outils

- **automatiser le plus possible le renseignement des métadonnées** (quand c'est possible) ;
- **varier les outils** : plateforme web, application smartphone, API,... Mitigation du risque (dysfonctionnement, obsolescence, défaut d'adhésion) ;
- Anticiper la **fracture numérique** (version papiers) - inclusivité (y compris les publics empêchés) ;
- **utiliser des standards interopérables** ;
- **volumétrie et flux de données - calibration des infrastructures (stockage et archivage)** ;

## ✓ Ergonomie des outils

## ✓ Données partagées, données ouvertes, transparence : favoriser la confiance

- outils open source ;
- **fiabilité de la mesure mais crédibilité du résultat** ;
- **ouverture des codes et des traitements appliqués aux données** ;
- mise à disposition de la donnée brute ;
- **différents niveaux de partage et d'ouverture des données.**

## ✓ Cas des capteurs (smartphone ou matériel spécifique)

CASPA - CApteurs et Sciences PArticipatives ([caspa.fr](http://caspa.fr))

Données et recherches participatives - Enjeux et recommandations issues d'exemples de projets de recherches participatives © 2024 by Kenneth Maussang is licensed under CC BY 4.0

