

# Panel probabilistes Internet en Europe : principaux avantages, défis et nouvelles opportunités

*Paris, 15 Juin 2023*

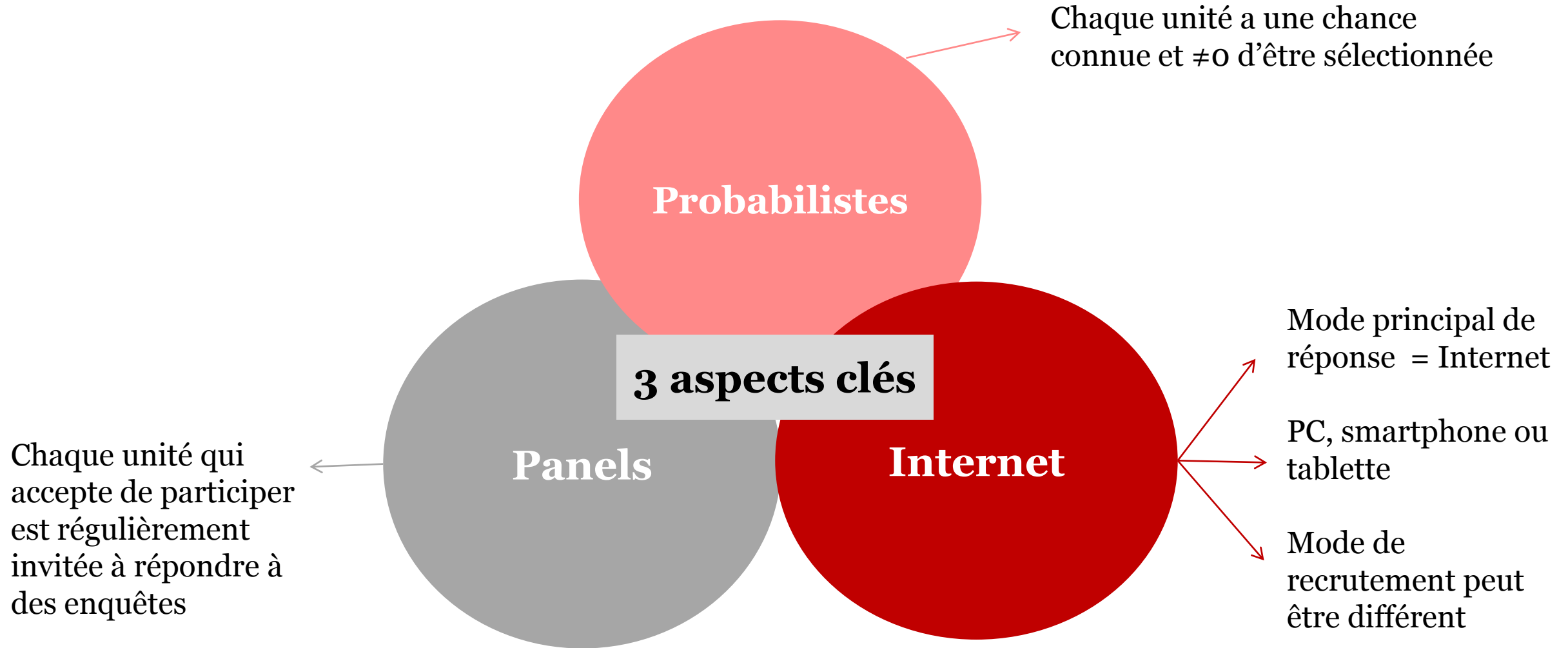
**Melanie Revilla | IBEI**

*Remerciements : Le projet WEB DATA OPP a reçu un financement du Conseil Européen de la Recherche (ERC) dans le cadre du programme de recherche et d'innovation Horizon 2020 de l'Union européenne (accord de subvention n° 849165).*



## INTRODUCTION

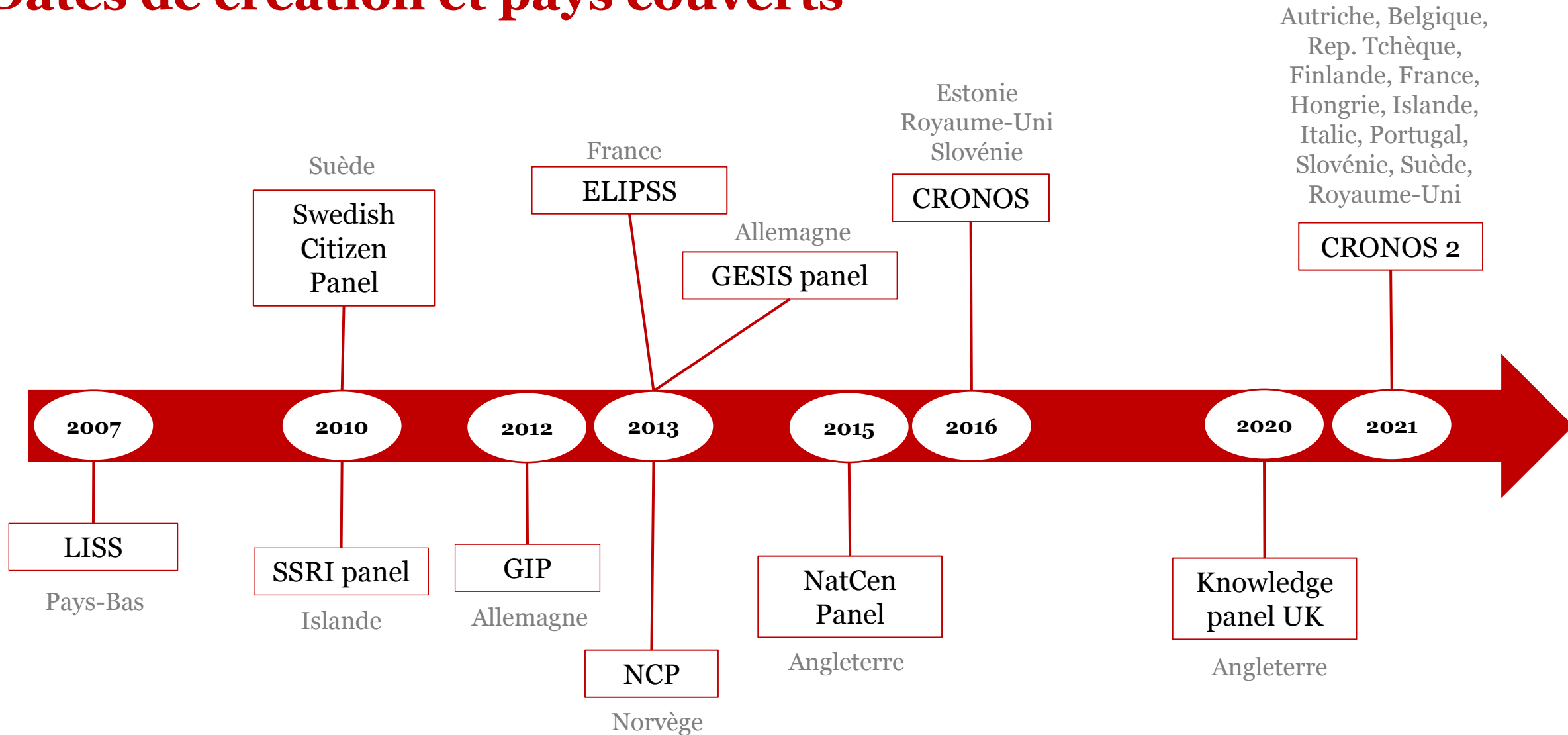
# Panel probabilistes Internet



# Principaux panels probabilistes Internet en Europe

## PRINCIPAUX PANELS

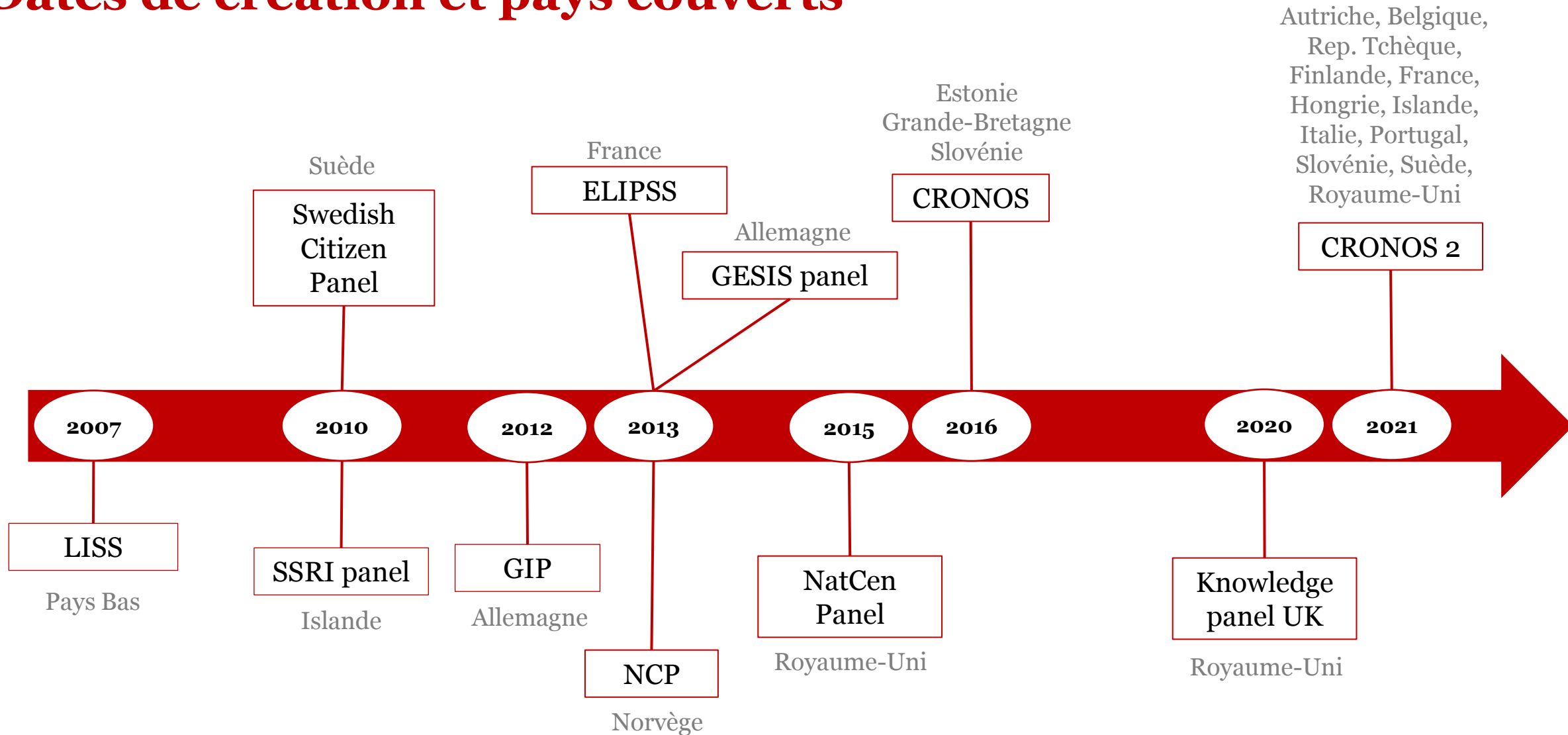
# Dates de création et pays couverts



➤ Peu de panels

## PRINCIPAUX PANELS

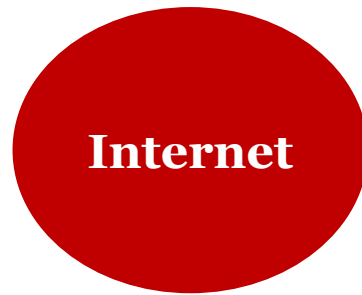
# Dates de création et pays couverts



➤ Peu d'occasions de les utiliser (taille réduite, fréquence enquêtes limitée)

# Principaux avantages et défis pour ces panels

# Avantages et défis : Internet (vs autres modes)



**Absence d'enquêteur**

**Stimuli visuel**

**Nécessite équipement/accès**



---

Participation faible

Biais de sélection

Efforts plus élevés

Qualité

Plus grande flexibilité

Coût réduit

## Avantages et défis : Probabiliste (vs non probabiliste)



Nécessite une **liste complète des unités** pour que toutes aient une chance connue ( $\neq 0$ ) d'être sélectionnées



Recrutement difficile

Coût plus élevé

Taille de l'échantillon réduite

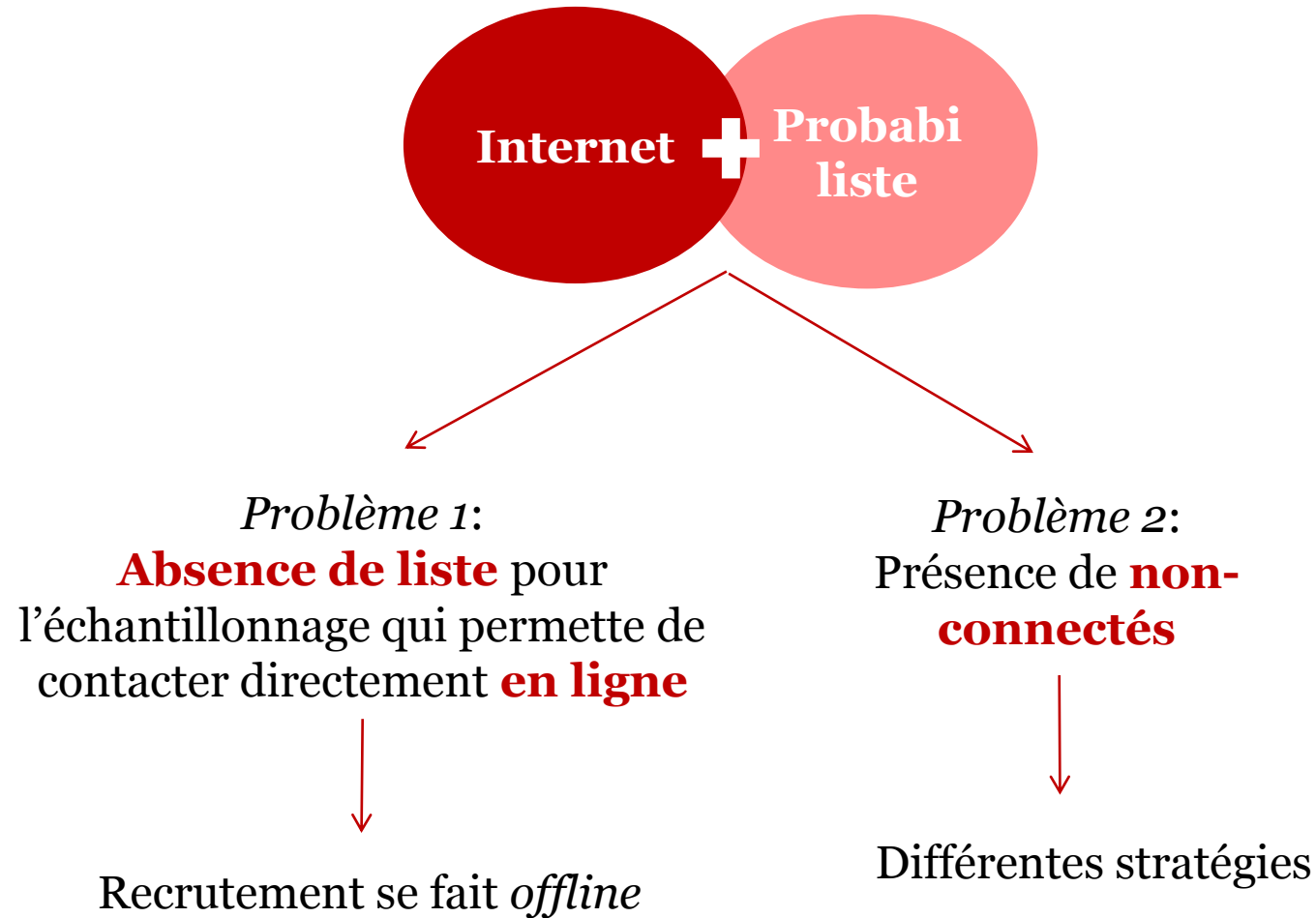
Qualité

Représentativité de la population cible

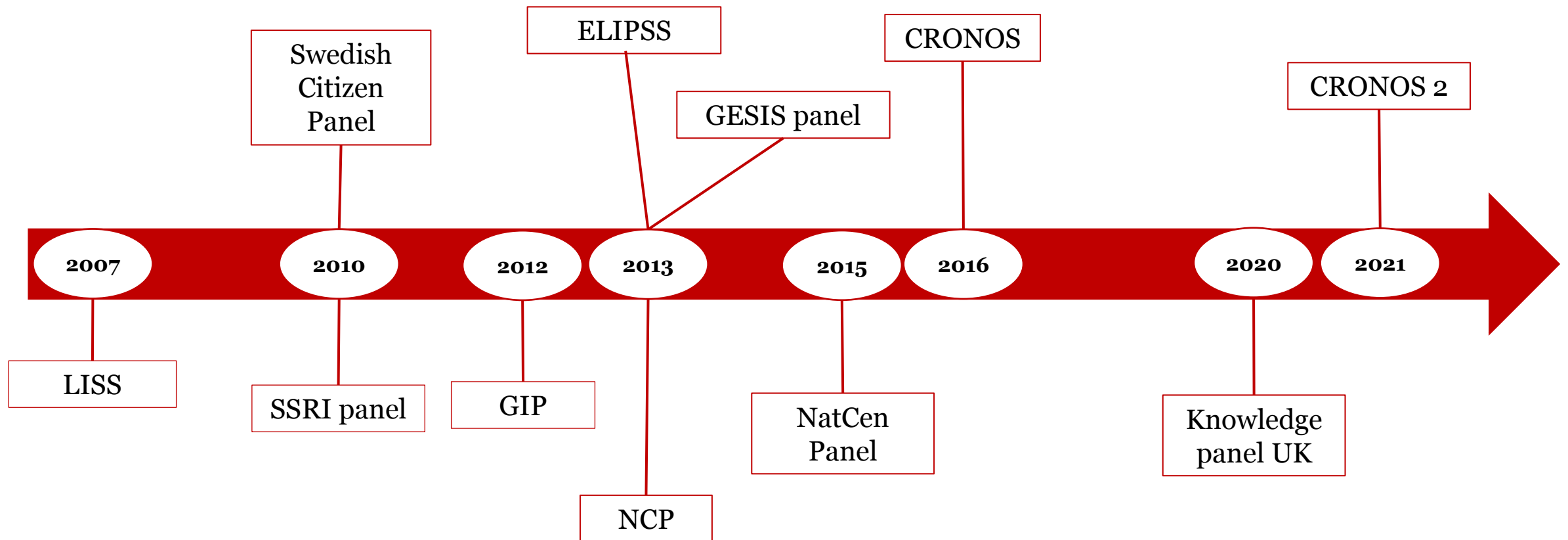
Meilleur contrôle de l'échantillon



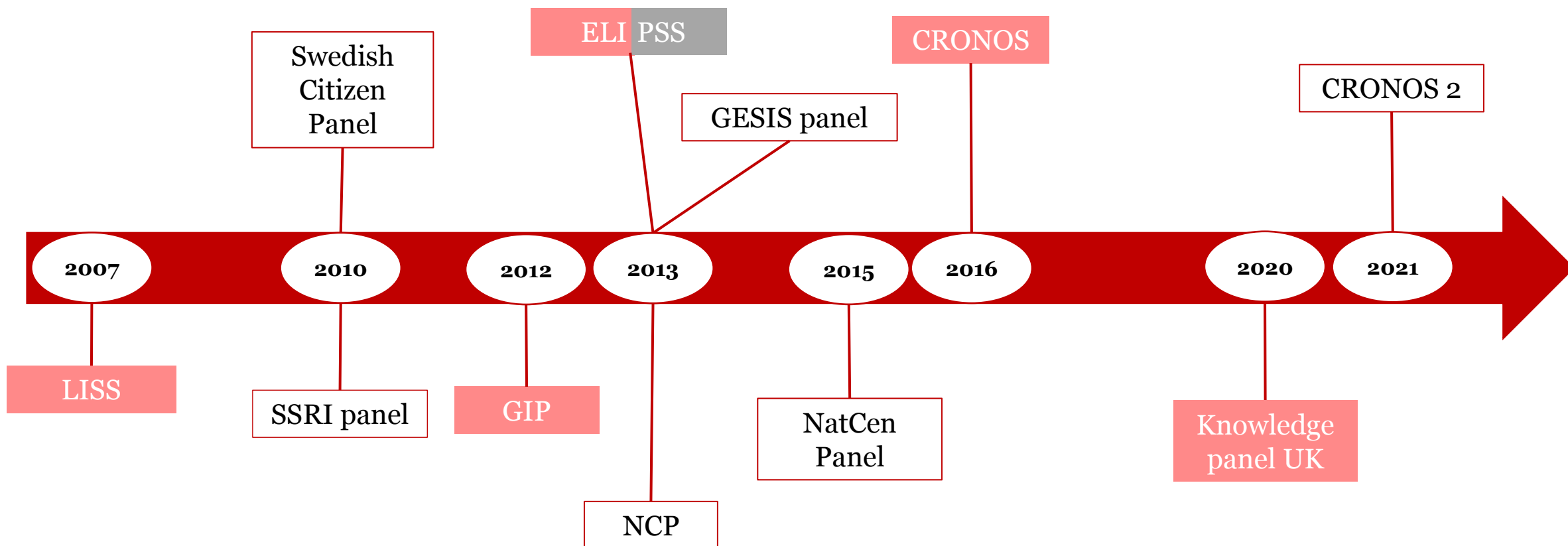
# Défis : combinaison Internet + probabiliste



# Traitement des non-connectés dans les principaux panels

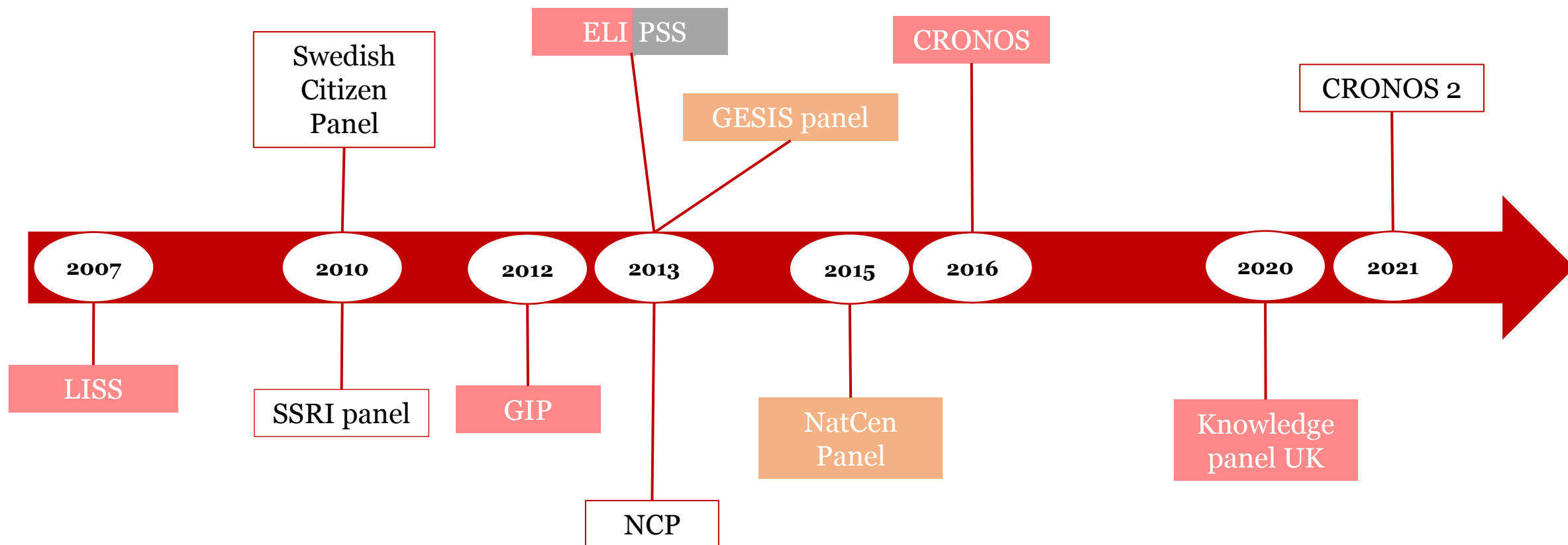


# Traitement des non-connectés dans les principaux panels



Facilite accès à Internet (si la personne invitée ne l'a pas déjà)

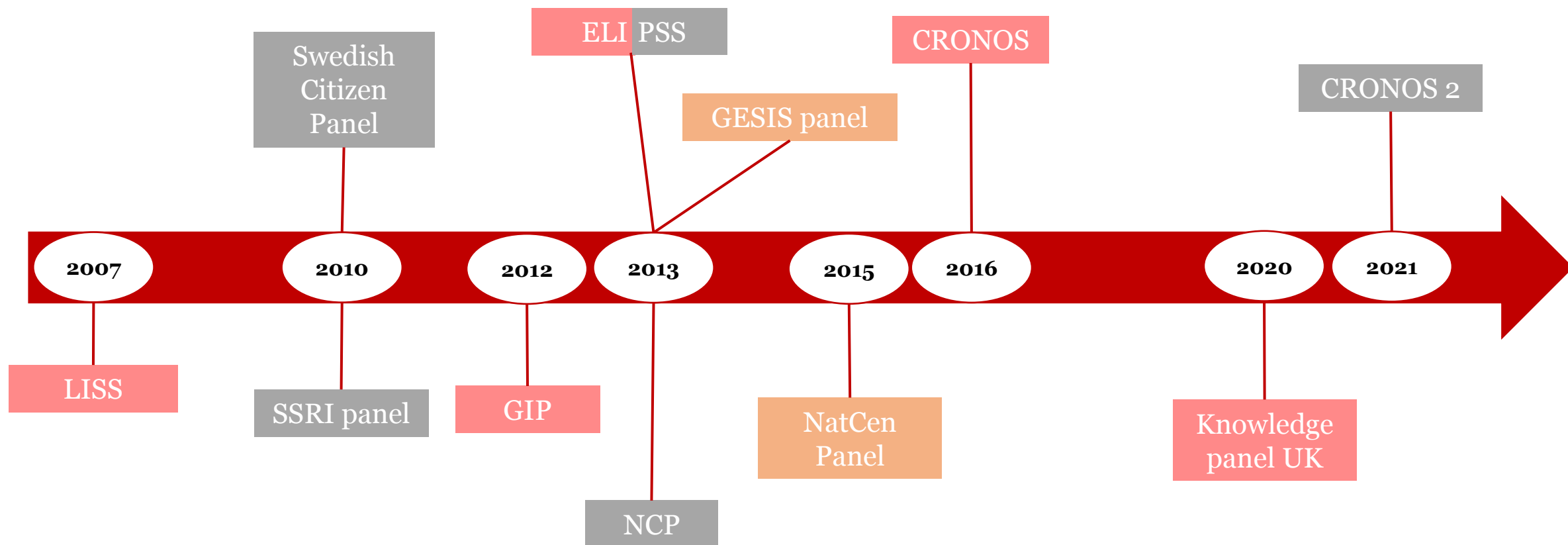
# Traitement des non-connectés dans les principaux panels



Facilite accès à Internet (si la personne invitée ne l'a pas déjà)

Mixed mode

# Traitement des non-connectés dans les principaux panels

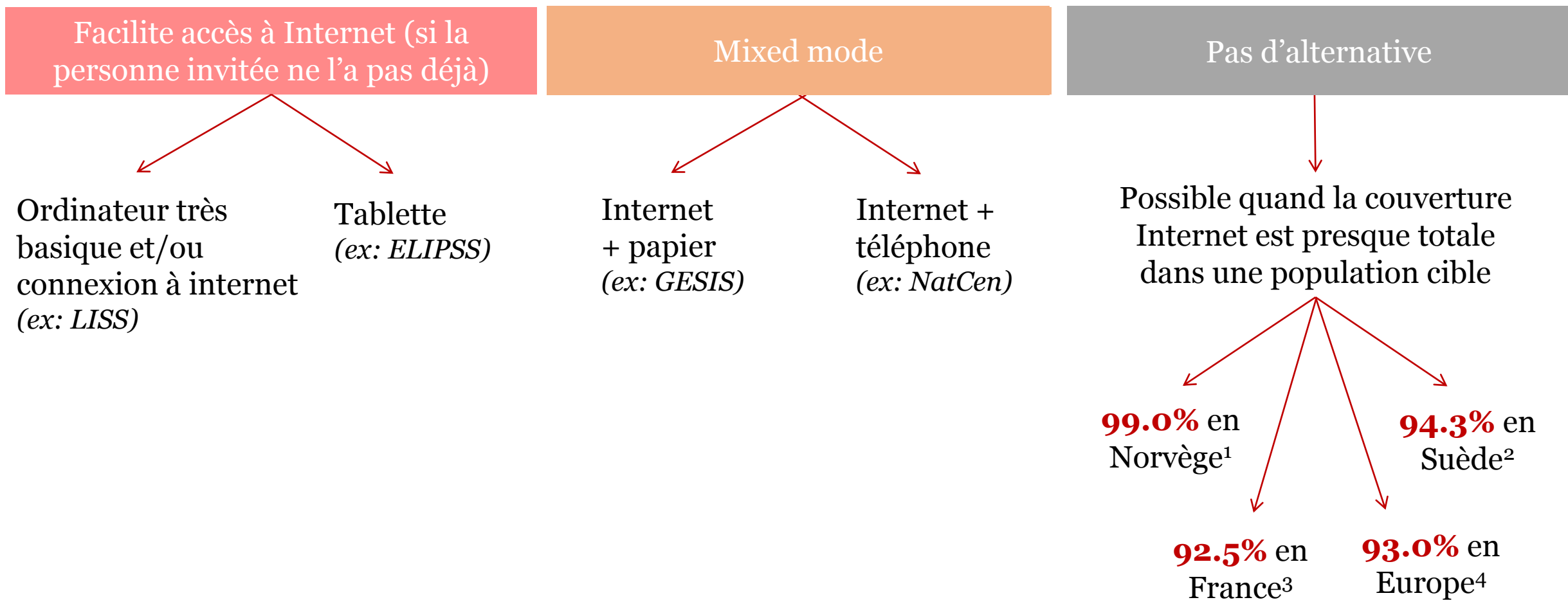


Facilite accès à Internet (si la personne invitée ne l'a pas déjà)

Mixed mode

Pas d'alternative

# Différentes stratégies pour les non-connectés



<sup>1</sup> <https://www.statista.com/statistics/377770/household-internet-access-in-norway/>

<sup>2</sup> <https://www.statista.com/statistics/377767/household-internet-access-in-sweden/>

<sup>3</sup> [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital\\_economy\\_and\\_society\\_statistics\\_-\\_households\\_and\\_individuals#Internet\\_access](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals#Internet_access)

<sup>4</sup> <https://www.statista.com/statistics/377711/household-internet-access-in-france/>

# Avantages et défis spécifiques à chaque stratégie

Facilite accès à Internet (si la personne invitée ne l'a pas déjà)

Mixed mode

Pas d'alternative



Coût élevé  
Difficultés pratiques



Représentativité (+)  
Uniformité



Coût élevé  
Difficultés pratiques  
Comparabilité des résultats



Représentativité (++)

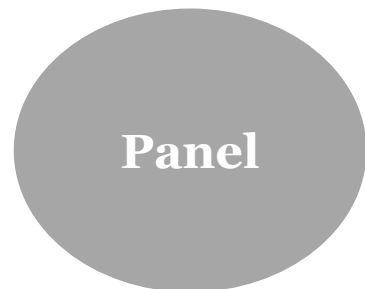


Représentativité



Coût réduit  
Implémentation simplifiée

## Avantages et défis : panel (vs enquête unique)



Participation **récurrente**

Possible dimension **longitudinale**



---

Attrition

Professionnalisation

Maintenance

Qualité

Rentabiliser le recrutement

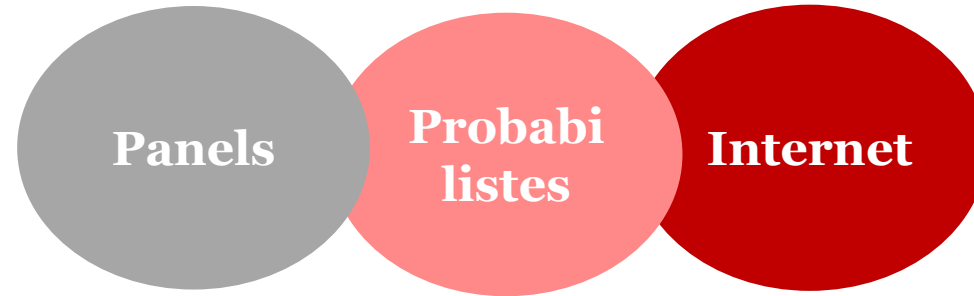
Relation de confiance

Etude du changement

Taux de réponse élevé



## Avantages et défis : conclusions



### ➤ D'importants avantages

Représentativité

Relation de confiance

Dimension longitudinale

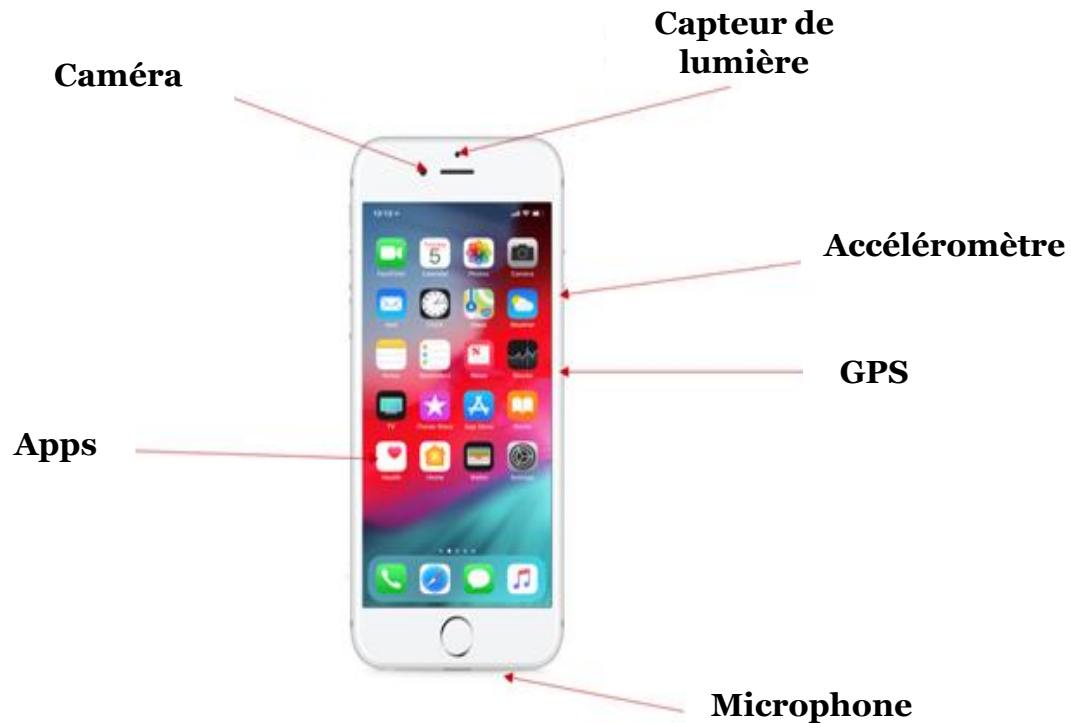
### ➤ Mais aussi d'importants défis

Possible biais → représentativité pas garantie

Problèmes de recrutement, maintenance, attrition

Nouvelles opportunités ?

## Possibilité de collecter différents types de données



Nombreux types de données,  
chacun ayant ses avantages/limites

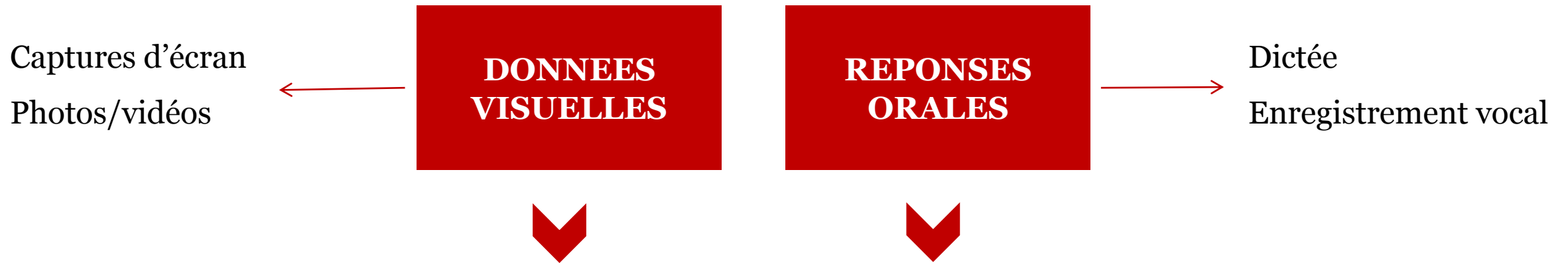


2 options principales:

→ Permettre **différents types de réponses**

→ **Combiner** avec d'autres données

# Répondre en partageant des données visuelles ou orales



## Exemple LISS (Ilic et al., 2022)

Question sur le système de chauffage

- Plus d'information dans les images
- Demander des images = réalisable + haute qualité mais participation réduite

## Exemple LISS (Meitinger et al., 2022)

2 questions (économie + minorités)

- La plupart des participants choisissent de répondre par écrit quand on leur laisse le choix
- ≈ 80% acceptent de répondre oralement quand aucune alternative n'est offerte

## Bénéfices et défis attendus

**DONNEES  
VISUELLES**

**REPONSES  
ORALES**



Participation plus faible

Problèmes techniques

Autres sources d'erreurs

Considérations éthiques

Qualité

Réduction de certaines erreurs de mesure

Information plus détaillée / spontanée

Information nouvelle / différente

Participation plus agréable

NOUVELLES OPPORTUNITES : COMBINER AVEC D'AUTRES DONNÉES

## Données collectées de façon passive

Coordonnées géographiques

**GÉOLOCA-  
LISATION**

**DONNÉES DE  
NAVIGATION**

Principalement, URL des pages web visitées



### LISS panel (Scherpenzeel, 2017)

Etude utilisant le GPS des smartphones pour collecter des données de façon continue sur les habitudes de déplacement des participants

- Données enrichies
- Mais aussi problèmes de différents types

### Exemple GESIS

GESIS is expanding its data collection services with a **web tracking panel** and a smartphone app for measuring media consumption and user behavior. In addition, GESIS is developing methods for collecting data from the Internet and for linking them with survey data.

## Bénéfices et défis attendus

GÉOLOCA-  
LISATION

DONNÉES DE  
NAVIGATION



---

Participation plus faible

Problèmes techniques

Autres sources d'erreurs

Considérations éthiques

Pas d'information subjective

Qualité

Effort quasiment nul

Information en temps réel / continue

Grande quantité d'informations

Information objective

## Bien plus de recherches nécessaires

- Apprendre plus sur les **erreurs** de ces données
  - Types d'erreurs
  - Taille
  - Comment elles affectent les résultats dans différents contextes
  
- Comprendre mieux **quand** et **comment** utiliser ces données
  - Comprendre les mécanismes
  - Remplacer ou combiner ?



# WEB DATA OPP

DONNEES  
VISUELLES

REPONSES  
ORALES

DONNÉES DE  
NAVIGATION

- Développement de *WebdataVisual*, un programme permettant de collecter les données visuelles dans le cadre d'enquêtes Internet (Revilla et al., 2022)
- Etude pour comprendre la non-participation lorsque des données visuelles sont demandées (Iglesias & Revilla, 2023)
- Guide pratique pour aider les chercheurs intéressés dans l'usage de données visuelles, en particulier en ce qui concerne l'extraction de l'information (Iglesias et al., 2022)

## WEB DATA OPP

DONNEES  
VISUELLES

REPONSES  
ORALES

DONNÉES DE  
NAVIGATION

- Développement de **WebdataVoice**, un programme permettant de collecter les données orales dans le cadre d'enquêtes Internet (Revilla et al., 2022)
- Implémentation d'enquêtes proposant aux participants de répondre en dictant ou en enregistrant leurs réponses

## WEB DATA OPP

DONNEES  
VISUELLES

REPONSES  
ORALES

DONNÉES DE  
NAVIGATION

- Développement du TEM = **“Total error framework for metered data”** (Bosch & Revilla, 2022a)
- Etude de la validité de différentes mesures de la consommation en ligne des médias d'information basées sur des données de navigation (Bosch & Revilla, 2022b)
- Etude pour estimer la présence de “tracking undercoverage” (Bosch et al., 2022)

# Conclusions

## Une option viable mais difficile

Panels

➤ Existents dans différents pays depuis plus de 10 ans

- Une option viable
- Avantages surtout en terme de représentativité
- Le temps semble aller en leur faveur → accès Internet ne cesse d'augmenter partout dans le monde

Probabi  
listes

Internet

➤ Importants défis aussi

- Coût + difficile de les créer et les maintenir
- Explique sans doute pourquoi ils sont encore si peu nombreux

## CONCLUSIONS

# Existe-t'il une meilleure alternative ?

- Dépend bien sûr de nombreux paramètres
  - Population cible
  - Thème de l'enquête
  - Budget disponible
  - ...
  
- Pour des études de la population générale: probablement une des **meilleures options** à l'heure actuelle

## CONCLUSIONS

# Existe-t'il une meilleure alternative ?



Panels

Probabilistes

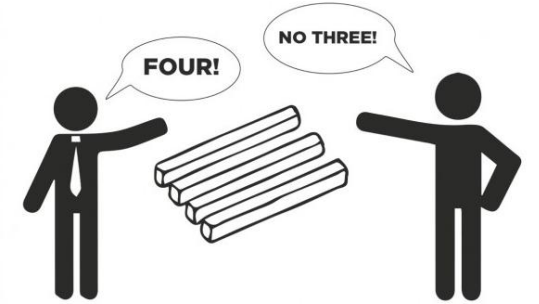
Interne  
t

Alternatives

## CONCLUSIONS

# Peut-on rendre la route un peu plus sûre ?

- Collaboration entre les différents panels existants
- Collaboration avec des entreprises privées
- Plus grande implication des gouvernements
- Permettre de nouveaux types de réponses + combiner avec d'autres données



Regarder depuis différentes perspectives enrichit nos connaissances  
→ Information différente mais complémentaire



# Merci !

## *Questions ?*

Melanie Revilla | IBEI



[mrevilla@ibei.org](mailto:mrevilla@ibei.org)



<https://www.upf.edu/web/webdataopp>



INSTITUT  
BARCELONA  
ESTUDIS  
INTERNACIONALS



## REFERENCES

- Bosch, O.J., & Revilla, M. (2022a). When survey science met web tracking: presenting an error framework for metered data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.  
<https://doi.org/10.1111/rssa.12956>
- Bosch, O.J., & Revilla, M. (2022b). Is tracking all that it takes? Exploring the validity of news media exposure measurements created with metered data. AAPOR Annual Conference, 11th-13th May 2022
- Bosch, O.J., Sturgis, P., & Kuha, J. (2022). Track me but not really: Tracking undercoverage in metered data collection. AAPOR Annual Conference, 11th-13th May 2022.
- Iglesias, P., & Revilla, M. (2023). Skills, availability, willingness, expected participation and burden of sharing visual data within the frame of web surveys. *Quality and Quantity*. Published online first.  
<https://doi.org/10.1007/s11135-023-01670-3>
- Iglesias, P., Ochoa, C., and Revilla, M. (2022). What do I do with these images? A practical guide to the classification of images sent by survey participants, RECSM Webinar, 31st March 2022. Available at:  
<https://www.youtube.com/watch?v=IQoKbO4XsZI>
- Ilic, G., Lugtig, P., Schouten, B., Streefkerk, M., Mulder, J., Kumar, P., & Höcük, S. (2022). Pictures Instead of Survey Questions: An Experimental Investigation of the Feasibility of Using Pictures in a Housing Survey, *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185(2): S437–S460, <https://doi.org/10.1111/rssa.12960>

## REFERENCES

- Meitinger, K., van der Sluis, S., & Schonlau, M. (2022). Implementing Voice-recordings in a probability-based panel: What we learnt so far. CIPHER conference. March 2022.
- Revilla, M., Iglesias, P., Ochoa, C., & Antón, D. (2022). WebdataVisual: a tool to gather visual data within the frame of web surveys. OSF. <http://doi.org/10.17605/OSF.IO/R7CAX>
- Revilla, M., Iglesias, P., Ochoa, C., & Antón, D. (2022). WebdataVoice: a tool for dictation or recording of voice answers in the frame of web surveys. OSF. <http://doi.org/10.17605/OSF.IO/B2WYZ>
- Scherpenzeel, A. (2017). Mixing online panel data collection with innovative methods. In: *Eifler S and Faulbaum F (eds) Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung*. Wiesbaden: Springer, pp. 27–49.