



cirad

LA RECHERCHE AGRONOMIQUE
POUR LE DÉVELOPPEMENT

INRA table ronde

séminaire de clôture data-partage
22 Novembre 2016

Partage des données d'enquêtes:
contexte enjeux besoins au SUD

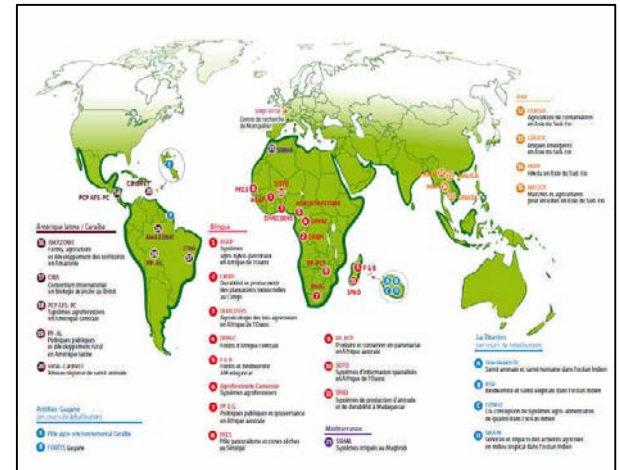
Pascal Bonnet (CIRAD département
E&S - PNS Patrimoine Numérique
Scientifique)



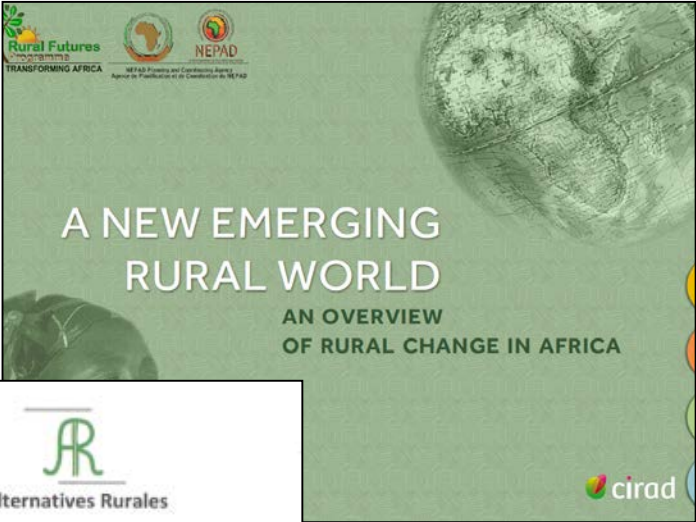
CIRAD Contexte: enquêtes au Sud, 3 départements BIOS

PERSYST ES concernés

- Données acquises en **partenariat** Nord - Sud et dans un contexte Sud
 - Déficiences de la **statistique publique** au Sud
 - Importance de mettre à disposition de la **donnée alternative**
 - par ex: d'observatoires des ménages ruraux
 - Déficiences ou incertitudes sur **règles « CNIL » locales**
 - se porter garant des bonnes pratiques
- Utilité d'une **conservation** et **ouverture** de certaines Données d'enquêtes
 - Rétrospectif** : **évaluer** reprise de certaines **données historiques** (apprécier diversité des enquêtes)
 - Prospectif** : **préparer** les **nouvelles collections scientifiques** numériques et la réutilisation sur des **enjeux thématiques majeurs**
 - Contemporain** : **cataloguer** données vivantes, en garantir une compréhension du **périmètre de réutilisation (validité)**

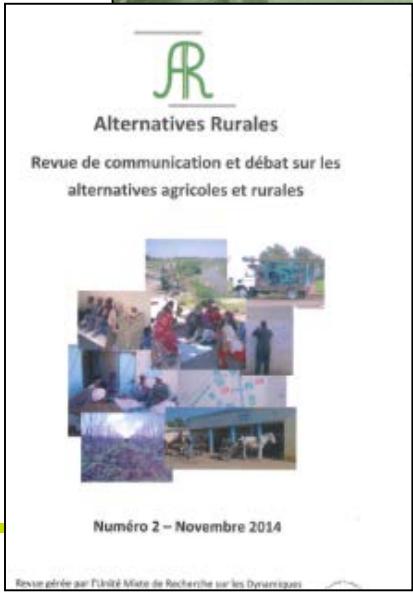


CIRAD Enjeux et ambition scientifique = créer et ouvrir des *collections thématiques* de données *réutilisables* sur de grands thèmes: « ex transformation structurelle des exploitations », comparaisons ici ailleurs , avant après => obj: utilisation donnée publique et données de recherche (ex: programme Ruralstruct et portail de données BM)



A NEW EMERGING RURAL WORLD
AN OVERVIEW
OF RURAL CHANGE IN AFRICA

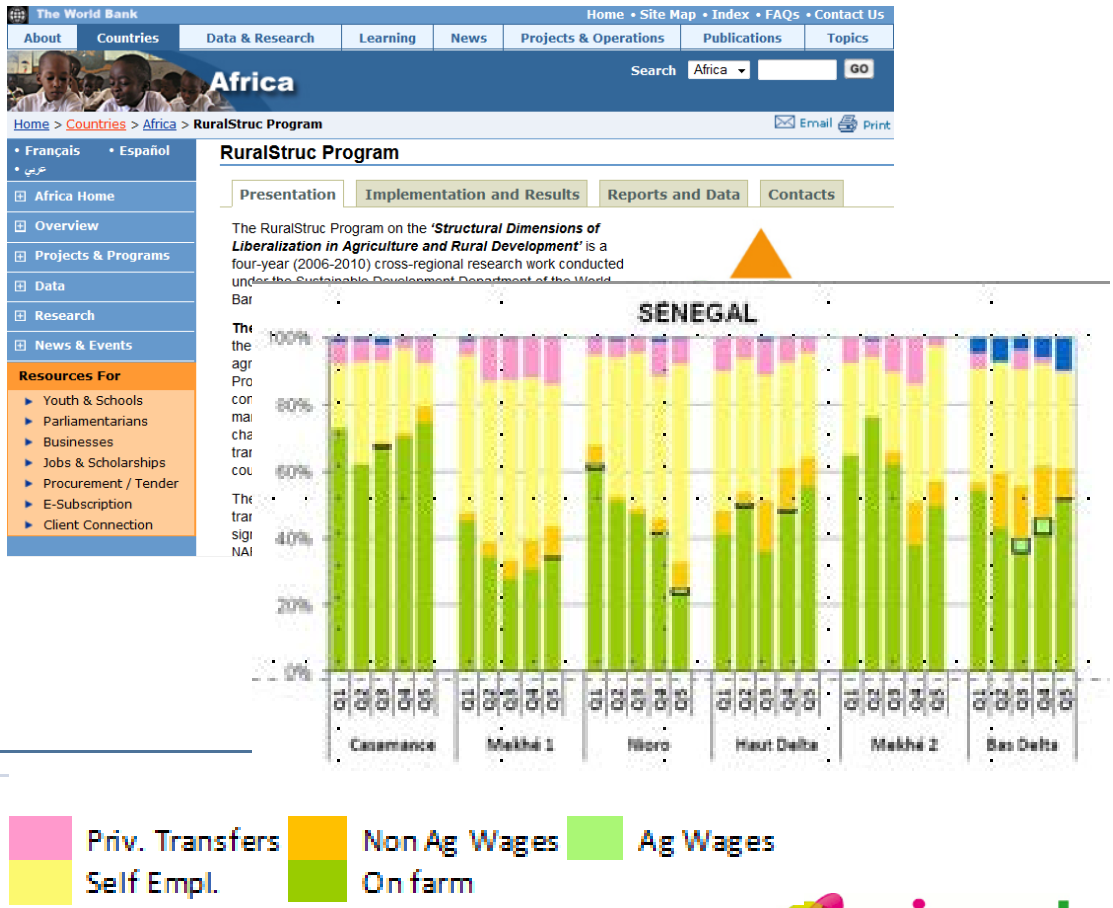
cirad



Alternatives Rurales
Revue de communication et débat sur les
alternatives agricoles et rurales

Numéro 2 – Novembre 2014

Revue créée par l'Institut Méditerranéen de Recherches sur les Dynamiques Rurales

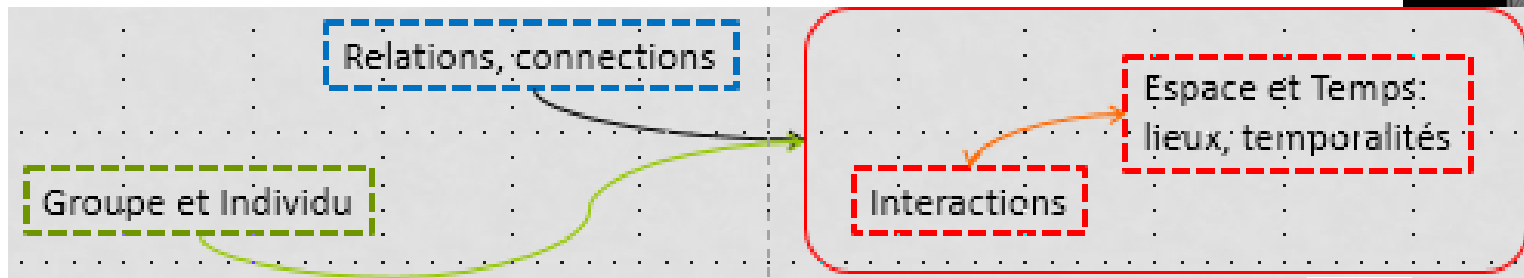
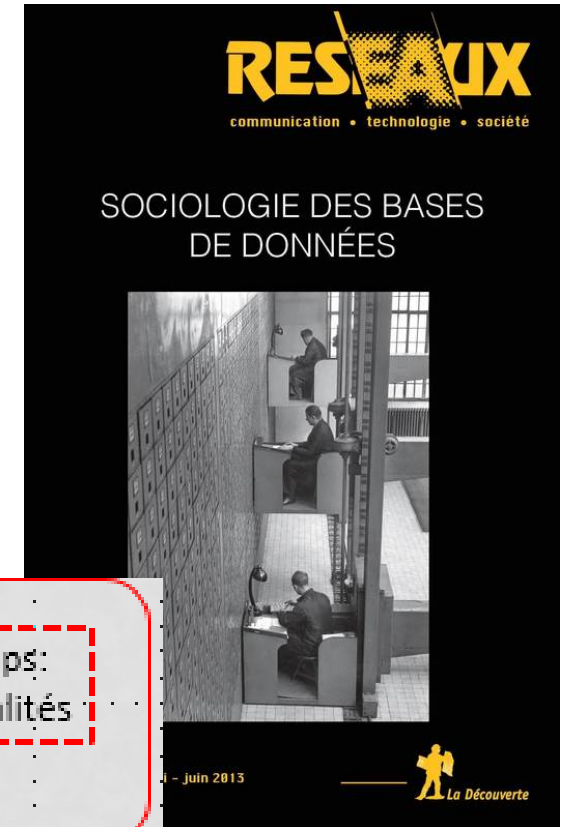


<http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/AFRICAEXT/0,,contentMDK:21079721~pagePK:146736~piPK:146830~theSitePK:258644,00.html>



CIRAD Enjeux d'accompagnement: mieux comprendre le **fonctionnement** et le **périmètre des « communautés de recherche Nord & Sud »**, avant de les accompagner vers un **partage raisonné** des données scientifiques

- Construction et diffusion de collections => savoir faire à développer et savoir être
- On navigue entre « *sociologie des BDD* » et « *BDD en sociologie anthropologie ethnologie économie agronomie* »
 - Connaitre:
 - Valeurs, Règles et Conventions sous jacentes aux communautés vis-à-vis des données, connaître leurs **espaces réels ou virtuels** d'interaction (entrepôts, portails...)...
 - Dynamiques de communautés intra établissement et externes (alliances..)
 - Avant de les faire évoluer..



Maison des Sciences de l'Homme
« *Environnement & Sciences Unis pour un @utre Développement* »

les enquêtes au SUD: DIAGNOSTIC

- A Partir de la **diversité** de leurs **enquêtes**, **quelles communautés** (SHS)?, **pratiques** et **instruments** (communautés de *pratique*) du partage, **objets** et **questions** de recherche incitant au partage, formes et **niveaux** de **collaboration** ?

- **Pratique des enquêtes , le besoin d'un diagnostic de la diversité**

- Quelles pratiques « respectives » du **partages de données de recherche**?
- Quelle culture « respective » du **quantitatif** , du **qualitatif**?
- Quels **instruments** communs mobilisés?
- Quels lieux **virtuels** communs de la collaboration / participation?
- Quelles pratiques de **réutilisation** de la stat publique ?
- Quelles spécificités des **disciplines** ?
- ...

- **Disciplines**

- Sociologie , Anthropologie..., Sciences politiques, « Sciences humaines élargies » :
Economie ... Géographie humaine..., enquêtes avec d'autres: agronomie , zootechnie , santé..



Quelle **diversité** des enquêtes d'exploitations?

Quelles **collections**? Quelles **pièces uniques**? (*exemples extraits du catalogue annuaire-Data CIRAD*)

DATA RECHERCHE

Exploitations PRASAC 2001 ; Tchad 2001 exploitations

Enquête diagnostic **agronomique**

Enquête paysanne sur le suivi des **pratiques agricoles** et mesure des **biomasses** a la récolte sur dix villages (trente **parcelles** par village) du degré carré de Niamey



Enquête crédit 2012

Base données **Exploitations Conseil** Individuel 2003

Base de données sur les **dimensions structurelles** des **exploitations** agricoles des villages d'intervention du PRASAC au Cameroun et au Tchad

DATA

RESULTS FROM THE DEMOGRAPHIC AND HEALTH SURVEYS

STAT PUBLIQUE

Burkina Faso 2010

Quelles **communautés d'instruments** et de **méthodes**, degré de **standardisation** d'enquête et d'analyse, pratiques de **suivi permanent**, de **mutualisation** ? (communauté de pratiques et de terrains ex: **modélisation Olympe**, **BDD mutualisées**, suivis et réseaux de fermes..)

Agriculture de conservation; **exploitations familiales**;
techniques agro-écologiques; gestion intégrée du sol; base
de données **géo-référencée**; développement durable;
changements d'échelle; outil de reporting; module

cartographique; **base de données mutualisée**

Suivi exploitations; filières **fruits**; agro-économie; itinéraire technique

Modélisation revenu exploitation
hévécicoles zone de Battambang

données type **Olympe** sur 70
exploitations agricoles

données type **Olympe** 5 jeux de données
'caractérisation des **exploitations**' sur **Excel** sur 230
exploitations agricoles et 45 sur Olympe

données type **Olympe** sur 35 **exploitations**
hévécicoles de la zone de Battambang
Cambodge

Modélisation revenu et **enquêtes** de caractérisation des exploitations
agricoles de la zone du Lac Alaotra (Madagascar)



*CIRAD: Enjeux sur les enquêtes au SUD et anticipation des besoins : **DOUBLE TRANSITION***

- **Intégrer les transitions** en cours dans les sociétés objets d'étude

- **Transition 1 numérique de la société au SUD**

- secteur agricole, alimentation , environnement..
- Évolution universelle vers de nouvelles sources de données +- ouvertes
 - **Web**, échanges d'information entre acteurs au cœur des **nouveaux réseaux** (sociaux) ..., **participation on line**, qui s'ajoute à l'existant (**enquêtes, entretiens structurés..**)
 - **BDD** des organisations (de producteurs ..)
- Évolution numérique permise dans les formes de collecte, d'accès, d'échanges : nouveaux workflow
 - Nouveaux **Capteurs** : forte pénétration **TIC** au SUD, importance des groupes de téléphonie mobile, arrivée des **opérateurs** et **capteurs** de l'agriculture numérique (robotique)

*CIRAD Enjeux sur les enquêtes au SUD et anticipation des besoins : **DOUBLE TRANSITION***

- **Intégrer les transitions** en cours dans les collectifs de recherche et sur les méthodes mobilisables

– Transition 2 numérique de la recherche

- Transformation de la **méthode** (inductif) avec les « **Sciences numériques** »
 - **Humanités numériques: Socio-informatique** (ex: des controverses...), TDM **fouilles de textes & data science** , à la manière de la « *bioinformatique* »
 - Évolution dans les **formes de traitement et modèles** (ex: SNA sociologie quantitative, exigence d'exhaustivité..)
 - **Outils de la collecte et workflow** de constitution de **nouveaux Corpus**: TIC tablettes.., capteurs, enregistrements audio & vidéos, SI FrontlineSMS, portails..
- Interactions avec les **Sciences du numérique** (INRIA) , sciences de l'information, informatique, systèmes - réseau – interfaces
 - nouvelles méta-communautés (RDA) et formes de collaborations
- Vers plus **d'Interdisciplinarité** et plus **d'interopérabilité** des vocabulaires, des données, ontologies

CIRAD Besoins génériques:

Équiper le cycle de vie: Rééquilibrer la double vie des données, pendant le projet et **après** projet (réutilisation)



Besoins: « équiper » le cycle de vie des données
composantes à combiner: hard, soft, compétences, contenus,
protocoles et procédures ... nouveaux services et produits

Besoins spécifiques: aller vers de nouvelles productions (Data Paper)

Accepter la **complexité** des contenus d'enquêtes

Développer les pratiques de **résumés, de documentation adaptée (Métadonnées MD..), expérimenter..**

30 exploitations, 1 année (2008 ou 2009), + d'une **centaine de variables**

Mots clés: Coton; exploitation paysanne; agriculture familiale; main-d'œuvre; terre; OGM; variétés; semences; coût de production; contrôle des ravageurs;

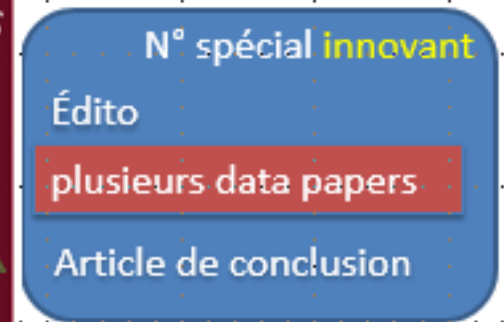
Entre MOTS CLES , Métadonnées et
Quasi résumé.. vers des Data Paper

"La base de données contient l'ensemble des informations recueillies sur le terrain lors d'un travail de recherche sur la caractérisation des performances agronomiques des systèmes de culture riz-maraîchage de bas-fonds à partir d'un échantillon de bas-fonds rizicoles situés au Sud-Est du Mali (2 bas-fonds, 45 exploitations agricoles couvrant 94 parcelles), au Sud-Ouest du Bénin (5 bas-fonds, 49 exploitations et 160 parcelles) et au Sud de la Sierra Leone (4 bas-fonds, 1 collecté agricole structuré (main d'activité culturelle ménage dans le espèce (comme enquête des sols des bas-fonds référencée par un track fortement ébranlé la Si

collecté agricole structuré (main d'activité culturelle ménage dans le espèce (comme enquête des sols des bas-fonds

référencée par un track fortement ébranlé la Si

en 2014, les données ne portent que sur une seule saison culturale et sont incomplètes. Toutes les données ont été initialement saisies et vérifiées sous Excel, puis stockées dans la base de données Access en effectuant une importation des fichiers Excel. Les données ont été saisies en respectant la structure des questionnaires d'enquêtes et des formulaires d'observations agronomiques. La base de données a fait l'objet de vérification, de correction et de validation de la qualité des données. L'application ""SYS MARIZ"" (base de données spatio-temporelle) qui a été produite facilite l'extraction des données de terrain pour rechercher la mise en évidence des relations éventuelles existant entre les observations agronomiques, les éléments déterminants des systèmes de



NUMEROS SPECIAUX

Besoins spécifiques: former sur les étapes et savoir faire encore mal connus pour diffuser / réutiliser les enquêtes

Diffusion de microdonnées 12 Q? – anonymisation

- 1. Qu'est-ce que les **microdonnées** ?
- 2. Qu'est-ce que les métadonnées ?
- 3. Quels sont les arguments en faveur de la diffusion des microdonnées ?
- 4. Quels sont les coûts et les risques liés à la **diffusion de microdonnées** et comment peuvent-ils être maîtrisés ?
- 5. **A qui les microdonnées sont-elles destinées** ?
- 6. Quelles sont les conditions régissant la diffusion des microdonnées ?
- 7. Qu'entend-on par « **anonymisation** » des microdonnées ?
- 8. L'accès doit-il être payant ou gratuit ?
- 9. A quel **moment du cycle de diffusion** les microdonnées doivent-elles être rendues publiques ?
- 10. Quelles sont les exigences à remplir en termes **d'infrastructure technique** ?
- 11. Quelles sont les exigences **institutionnelles** relatives à la diffusion de microdonnées ?
- 12. Comment **promouvoir l'utilisation** de microdonnées ?

IHSN



RÉSEAU INTERNATIONAL POUR LES ENQUÊTES AUPRÈS DES MÉNAGES

Diffusion des fichiers de microdonnées

Principes, procédures et pratiques

Olivier Dupriez and



Anonymisation:
managing data
protection risk
code of practice

2005
2010

cirad

Conclusion (1)

- **Partage des données d'enquêtes**
- Accompagner la transition et répondre aux **besoins exprimés ou non encore exprimés** (en anticipation) des chercheurs et communautés de recherche
- Cela concerne *plusieurs domaines à la fois* :
 - Les **structures** : physiques, économiques, institutionnelles et organisationnelles
 - La **culture** : valeurs collectives, normes, représentations, visions d'avenir
 - Les **pratiques** : routines productives, comportements, manière d'être et réflexivité...

Conclusion (2)

- ❖ Favoriser la **pratique du partage** (intra: CIRAD et partenaires, et externe) et la **réutilisation** de données (internes et externes):
- ❖ Envisager le partage des enquêtes comme partie d'un **chemin d'impact**, comme un appui au développement (pour les bénéficiaires, paysans et sociétés au Sud)
 - ✓ Données « **utiles utilisables utilisées** »:
 - ✓ Enjeux de mieux **caractériser** et **faire connaître** les besoins des agriculteurs du SUD
 - Advocacy : partager = **faire connaître** les spécificités des ménages ruraux au sud
 - La recherche a un rôle dans les **nouvelles collections de référence** sur des transformations en cours (cf Rural struct)
 - Au sud admettre la **rareté** d'éléments de description & caractérisation de processus en cours
 - Ne pas remplacer stat publique (FAOstat), mais connaître ses limites
 - ✓ Traduire positivement les questions de **déontologie et d'éthique** sur les données personnelles
 - Développer une capacité pour **sécuriser** les données personnelles (anonymiser)
 - ✓ s'engager dans la **réutilisation** exige **l'organisation** du partage
 - Instances de **gouvernance** intra et inter établissements

Merci

