

***Archiver et partager ses données.
Kinsources, une plateforme collaborative de
diffusion de données de parenté***

Anne Garcia-Fernandez

CNRS - Laboratoire d'Anthropologie Sociale, (UMR7130 Collège de France-CNRS-EHESS), Paris

Pascal Cristofoli

EHESS - Laboratoire de démographie et d'histoire sociale (LaDéHiS-CRH, UMR 8558 EHESS-CNRS), Paris

***Collecter et produire des données pour la recherche en SHS
Fréjus, 15-18 novembre 2016***

- Site web
<http://kinsources.net>
- Données de parenté
- Publication, partage, archivage
- Un modèle collaboratif
- Alimenté par les chercheurs





kinsources
archive, share, analyze and compare kinship data

Kinsources.net - Archiver et diffuser les données de la recherche en SHS

© P. Cristofoli & A. Garcia-Fernandez (Groupe Kintip)
V. 1 - Septembre 2016 - cf. contact@kinsources.net



Huma-Num

Kinsources.net est propulsée par **KIWA** (Kinsources Web-Application) – logiciel libre programmé en Java (License CeCILL v2.1).

Conception et réalisation :





Pour une vie numérique sous **MON** contrôle

PRIORITÉ AU LOGICIEL LIBRE ! 

<https://gitlab.com/kinsources/kiwa>

Partenaires :





L'ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES

université Paris Ovest
Nanterre La Défense

COLLÈGE DE FRANCE
1530

Paris SORBONNE UNIVERSITÉ

Inria
Institut National de Recherche en Informatique et en Automatique

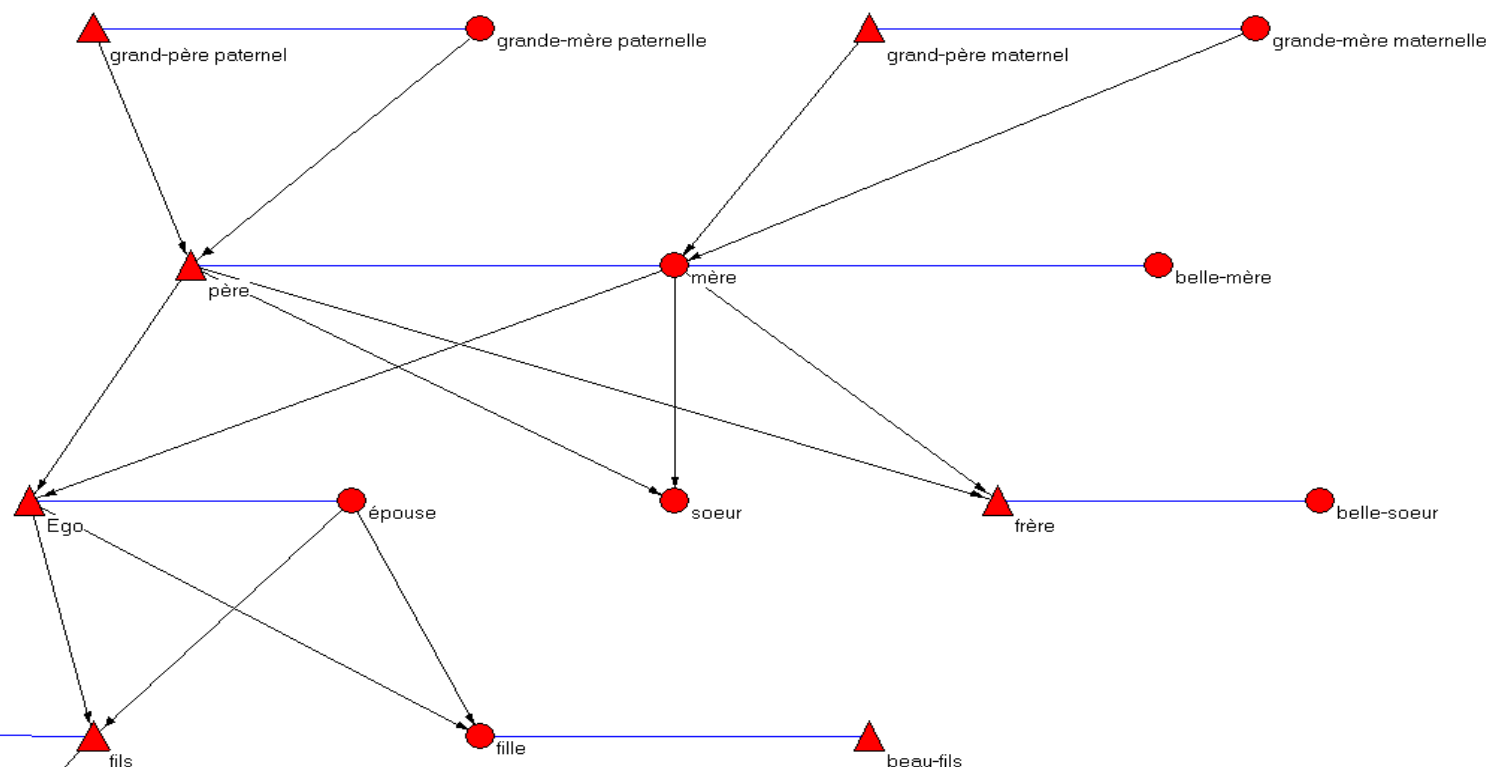
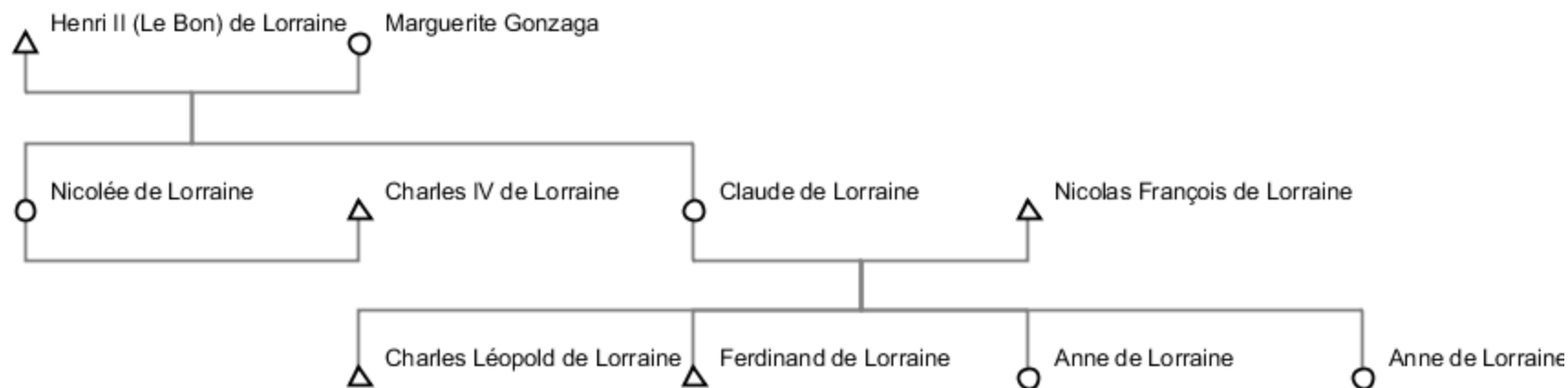
University of Kent
The UK's European university

UCLA



ined
INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES

Réseaux de parenté : liens de filiation et d'alliance entre individus



Corpus de données de parenté

1  Louis XVI de France Individu

1  Louis de France
3  Marie Joséphine de Saxe

Conjoints (1) + (Ctrl-P)

Id	Nom	Pos...	Famille
4	Marie Antoinette d'Autriche		2

Relations de parenté :
liens entre individus

Enfants (4) + (Ctrl-K)

Id	Nom	Pos...	Fratrie	Autre parent
107	Marie Thérèse de Marnes			2F (4) Marie Antoinette d'Autriche
108	Louis de France			2F (4) Marie Antoinette d'Autriche
109	Louis de France			2F (4) Marie Antoinette d'Autriche

Attributs (6) + ++ C S

Label	Valeur	Informations
BIRT Date	23/8/1754	temporelles,
BIRT Place	Versailles	spatiales...
COMPL	complete	
DEAT Date	21/1/1793	
DEAT Place	Paris	

Il s'agit d'un **ensemble** de données concernant **des individus ayant des liens entre eux**, collectées au cours d'enquêtes ethnographiques ou à partir de sources documentaires. Cette population, définie dans le cadre d'une recherche scientifique, est enregistrée sous forme de **données numériques** caractérisant aussi bien les **individus** eux-mêmes que les **relations** qu'ils entretiennent, et comprenant des **informations spatiales et temporelles** qui peuvent leur être associées.

Extension aux données relationnelles

	A	B	C	D	E	
1	Transactions					
2	Id	name	TRANS1	TRANS2	#rel_annee	#trans
3	1	TR000297	12075	18560	1743	c
4	2	TR000299	22462	21979	1743	g
5	3	TR000300	23202	26882	1743	d
6	4	TR000305	21979	6720	1743	v
7	5	TR000306	21979	6724	1743	v
8	6	TR000308	353	21979	1743	ec
9	7	TR000310	2380	26629	1743	g
10	8	TR000311	24564	2466	1743	v
11	9	TR000312	3967	12227	1743	c

	A	B	C	D	E	F
1	Parrains					
2	Id	name	EGO	PARRAIN	#rel_annee	#Rel_def
3	1	BAPT00001	36	4574	1753	Ego-Parrain
4	2	BAPT00002	52	8361	1754	Ego-Parrain
5	3	BAPT00003	53	26423	1757	Ego-Parrain
6	4	BAPT00005	55	26467	1761	Ego-Parrain
7	5	BAPT00007	57	8430	1764	Ego-Parrain
8	6	BAPT00011	44	50	1865	Ego-Parrain
9	7	BAPT00012	45	4642	1868	Ego-Parrain
10	8	BAPT00013	47	4642	1870	Ego-Parrain
11	9	BAPT00014	67	23825	1749	Ego-Parrain

Si les relations envisagées dans les corpus de parenté privilégiaient initialement les liens de descendance et de mariage, **désormais n'importe quelle relation est susceptible d'être modélisée sous forme de « nœud relationnel »** [Hamberger & al., 2014]. La parenté spirituelle, la relation de confiage ou tout autre type de « réseaux sociaux » peuvent être codées en complément. C'est aussi le cas des événements biographiques, des séquences d'événements, des affiliations sociales multiples ou encore des terminologies de parenté.

Données au croisement de traditions disciplinaires et scientifiques

Depuis des décennies, de nombreuses recherches en anthropologie, en histoire ou en démographie se fondent sur la production et la manipulation de ce type de données

- Étude des systèmes de parenté
- Analyse informatisée de la parenté
- Analyse démographique, Démographie historique et reconstitution des familles
- Analyse des réseaux sociaux, analyse des données relationnelles issues de sources documentaires

Différents types d'échantillonnages : corpus généalogiques « complets » de communautés, monographies villageoises ou régionales ; généalogies tronquées patrilinéaires ou matrilinéaires ; prosopographies réalisées par les historiens ou encore réseaux ego-centrés produits par certaines enquêtes socio-démographiques.

La connaissance de la nature de l'échantillonnage et de ses caractéristiques (profondeur temporelle, envergure spatiale, systématisme du recueil des informations,...) est évidemment indispensable pour procéder à l'analyse des données et à leur interprétation.

Réutilisation de données

- En *entrée* :
 - Corpus issus de carnets de terrain, publications... : numérisation et codage pour dépôt sur la plateforme
 - Formats
- En *sortie* :
 - Publication de données pour permettre leur réutilisation
 - Formats
 - Méta-données

Formats d'entrée et sortie gérés par la plateforme

- Les principaux formats utilisés dans le domaine de la parenté et des graphes :
 - txt (Texte)
 - xls (Excel)
 - ged (Gedcom)
 - paj (Pajek)
 - ...
 - et xml/puc (Puck)

Arelation2.txt

225 M 449 448

Parrainage

Id	Name	PARRAIN	FILLEUL	#LIEU	#STATUS	#TYPE
1	alpha	10	20	Paris	Civil	Mariage
2	bravo	11	21	Massy	Autre	Baptême
3	charlie	12,14	22			
4	delta	13	23,15			

CA

Id

Name

MEMBRE

PRESIDENT

#SALLE

#SECRETAIRE

Coller

Aligner

mise en forme conditionnelle

Styles

Actions

Thèmes

J16

ID	Name	Gender	Father	Mother	Spouses							
1	H 1	H	277	278	2	324	99					
2	F 2	F	263	262	1	179	108					
3	H 3	H	1	2								
4	H 4	H	1	2								
5	H 5	H	223	324	6							
6	F 6	F	128	127	5	179						
7	H 7	H	5	6								
8	F 8	F	5	6								
9	H 9	H	5	6								
10	H 10	H	352	353	11	408	20					
11	F 11	F	356	357	10	362	405					
12	F 12	F	1	324								
13	F 13	F	15	312	15							
14	F 14	F	15	312								
15	H 15	H	362	11	312	16	13					
16	F 16	F	308	247	15							
17	H 17	H	10	11	18							
18	F 18	F	281	205	17							
19	F 19	F	17	18								
20	F 20	F	17	18	10							
21	H 21	H	32	33	22							
22	F 22	F	307	247	21	188						
23	H											
24	H											
25	F											
26	H											
27	H											
28	F											
29	H											
30	F											

Ebrei.ged - WordPad

Fichier Edition Affichage Insertion Format ?

0 HEAD

1 SOUR Généalogie

2 VERS V1.0

2 CORP DefCompany

1 DEST Généalogie

1 DATE 26 MAR 2009

1 FILE Ebrei

1 GEDC

2 VERS 5.01

1 CHAR IBM WINDOWS

0 @I1@ INDI

1 NAME Sabato /Cagli/

1 SEX M

MATE-SHS La convergence vers un format de données XML (PUC)

```
- <individuals size="4735">
  - <individual id="1">
    <name>Sabato /Cagli/</name>
    <gender>MALE</gender>
  - <attributes size="6">
    - <attribute>
      <label>BIRT_PLAC</label>
      <value>Cagli</value>
    </attribute>
    - <attribute>
      <label>BIRT_DATE</label>
      <value>1610</value>
    </attribute>
    - <attribute>
      <label>OCCU</label>
      <value>Banchiere</value>
    </attribute>
    - <attribute>
      <label>RESI</label>
      <value>San Marino</value>
    </attribute>
    - <attribute>
      <label>CHAN_DATE</label>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<corpus ... >

  <individuals size="20"> ... </individuals>

  <families size="10"> ... </families>

  <RelationModels size="5"> ... </RelationModels>

  <Relations size="30"> ... </Relations>

</corpus>
```

Collaborative web platform for kinship data share

Archive, share, analyse and compare kinship data.

We are pleased to welcome you to this web site dedicated to archive, analyze and compare kinship data. Join us and share your kinship data.

Enjoy the resources of a scientific community.

[JOIN US](#)
[MORE ABOUT KINSOURCES](#)


Trois usages pour une plateforme

Interface publique de Kinsources.net

- # Accéder aux corpus de parenté publiés
- 👁️ Naviguer parmi les corpus de données via l'interface du site
- ➡️ Rechercher des corpus selon différents critères
 - 🔍 Rechercher sur les métadonnées des corpus
 - 📊 Rechercher selon les statistiques des corpus

La fiche d'un corpus public sur Kinsources.net

Description & référencement	Identification	Diffusion des Données
Métadonnées	-> Nom du corpus -> Contributeur -> Auteur(s) -> etc.	✓ Intégrité
Lien pérenne		📊 Statistiques
Date publication		⬇️ Téléchargement
Documentation		© License libre de diffusion
Références scientifiques		🏆 Validation scientifique



Espace de stockage des corpus publics

Un espace public

Un accès à la communauté



Un espace privé

Interface privée des contributeurs Kinsources

Devenir "contributeur"

Créer son espace personnel (login - mot de passe)



Espace de stockage de A



La fiche d'un corpus public sur Kinsources.net

Description & référencement

Métadonnées



Lien pérenne



Date publication



Documentation



Références
scientifiques



Identification

- > Nom du corpus
- > Contributeur
- > Auteur(s)
- > etc.



Diffusion des Données

Intégrité



Statistiques



Téléchargement



License libre
de diffusion



Validation
scientifique



Author: Public Domain | **Publication date:** 17/05/2010

Status: 

License: <http://creativecommons.org/licenses/by-nc-nd/3.0>

Short description: -

General

Name: Ancien Régime
Simplified name: ancien_regime
Author: Public Domain
Coder: Klaus Hamberger
Contributor: Klaus Hamberger
Contact: -
Creation date: 20/11/2013
Edition date: 20/11/2013

Place

Location: Western Europe
Country: -
Region: Western Europe
Continent: Europe
Coordinate: -
Radius from center (km): -

Culture

Ethnic or cultural group: European high nobility
Atlas code: -
Bibliography: -

Time

Period (from/to): 16th-18th centuries
Coverage: -
Period note: -

Data publication

Permanent link: <https://www.kinsources.net/kidarep/dataset-53-ancien-regime.xhtml>

Conditions of use: <http://creativecommons.org/licenses/by-nc-nd/3.0>

Statistics

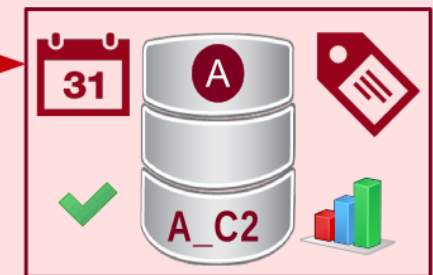
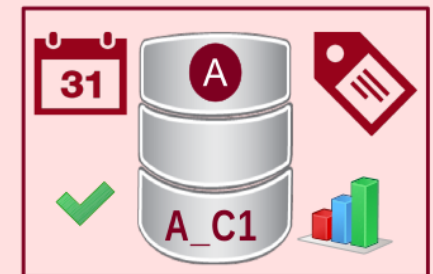
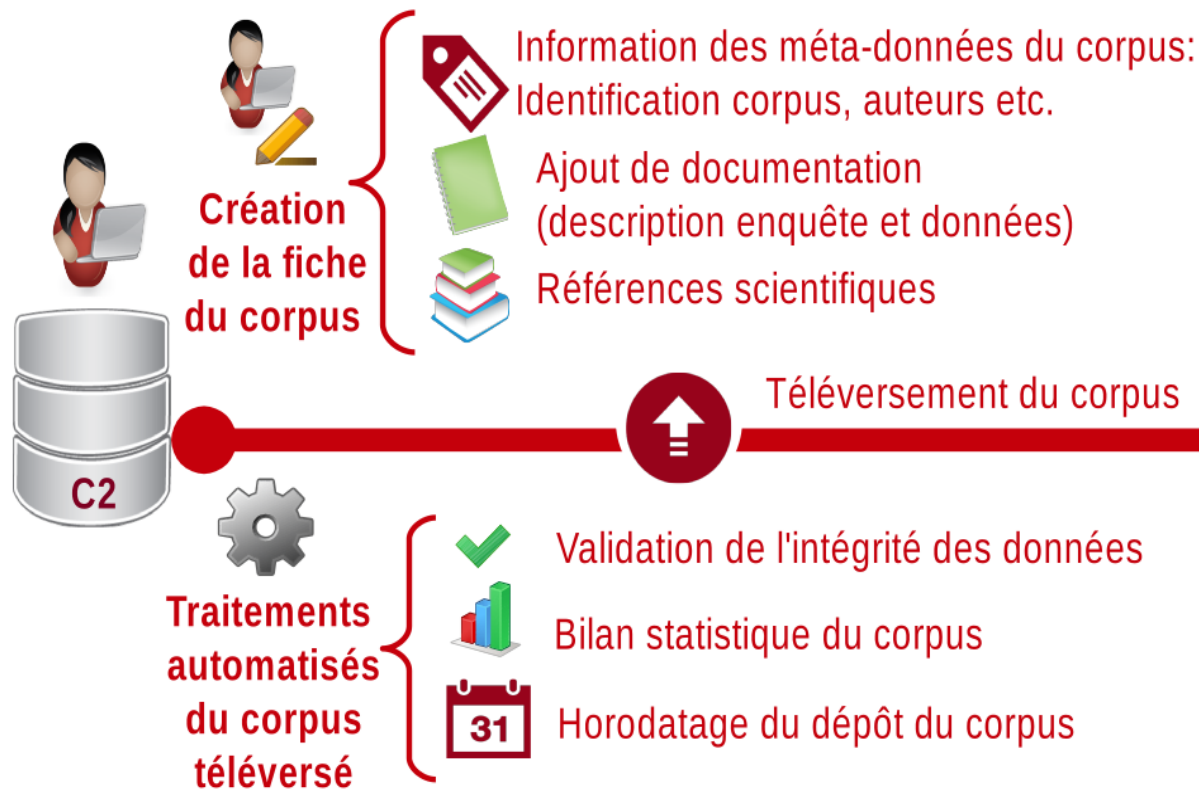
Individuals 5891 Unions 2310 Relations 0 Generations 11

Individuals	5891		Unions	2310		Depth:	11
Men	2889	49.04%	Marriages	2189		Depth mean:	2.36
Women	2970	50.42%	Marriages density		0.63%	Agnatic fratty size mean:	3.81
Unknown	32	0.54%	Fertile marriages	1185	54.13%	Uterine fratty size mean:	3.21
Non single men	1824		Parent child ties	8373			
Non single women	2054		Filiation density		2.41%	Components:	5
Mean spouse of men	1.2		First cousin marriages	90	4.11%	Mean components share (agnatic):	0.06%
Mean spouse of women	1.07		Cross (first) cousin marriages	48	2.19%	without singleton:	0.23%
Co-husband_relations	146		Parallel (first) cousin marriages	42	1.92%	Mean components share (uterine):	0.06%
Co-wife_relations	437		Niece/Nephew marriages	26	1.19%	without singleton:	0.14%
			Levirate marriages	14	0.64%	Max components (agnatic):	4.28%
			Sororate marriages	22	1.01%	Max components (uterine):	3.85%
			Double or exchange marriages	77	3.52%		
			Double marriages	33	1.51%		
			Exchange marriages	50	2.28%		
			Mean children per fertile couple	3.03			

REBUILD STATISTICS

WATCH RAW STATISTICS

Archiver un corpus sur la plateforme



MATE-SHS Publier un corpus : préparation

- Préparation de la version publique du corpus :
 - Filtrage des données
 - Anonymisation
 - Choix des attributs

Attribute filters edition

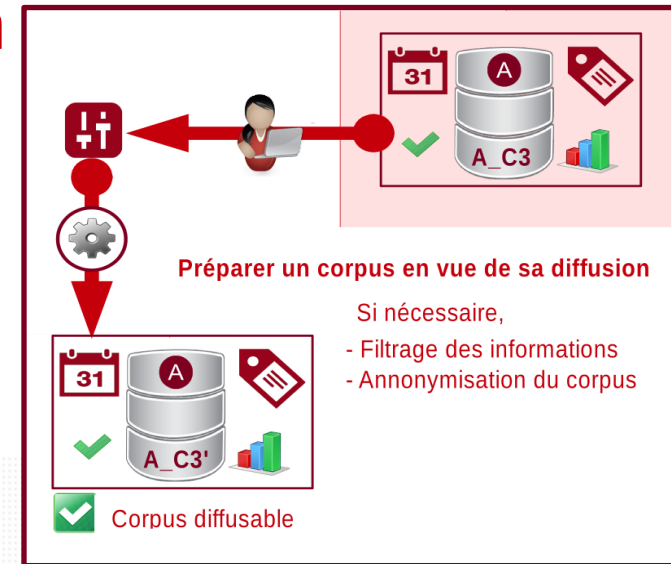
In this step, you define attribute filters for the public version of your dataset file. This allows you to secure your file in Kinsources while not publishing confidential information about people.

For legal reason, this step is very important. For example, do not publish personal information about living people.

Individual names

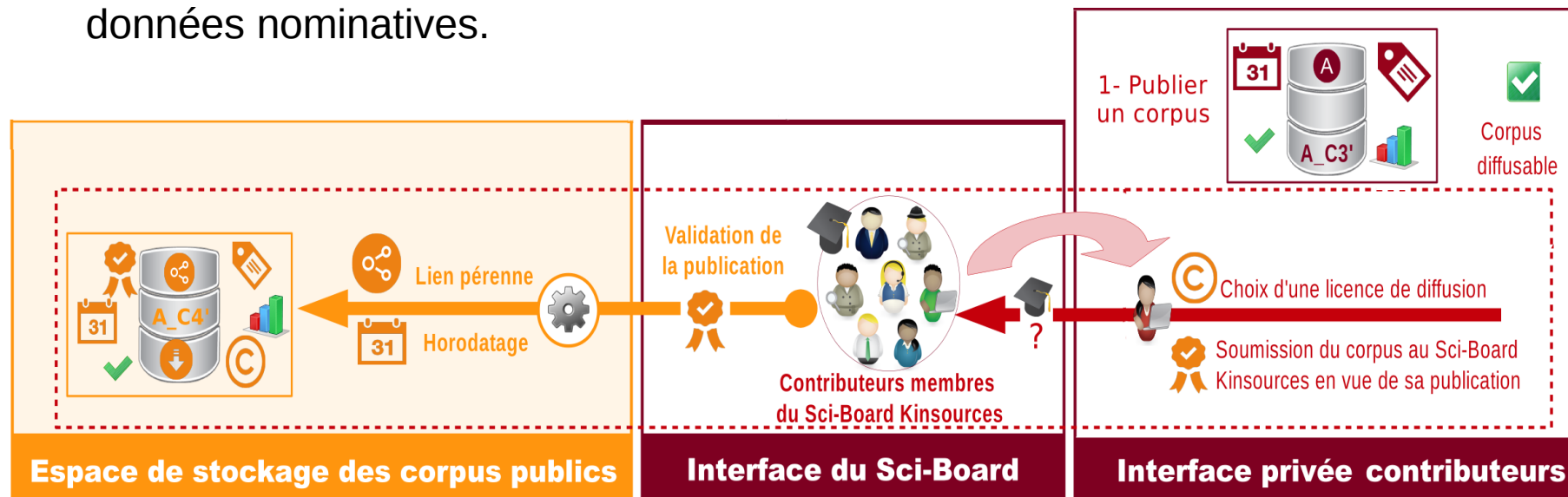
- ☐ No filtering
- ☒ Anonymize by gender and id (default settings, check if this is useful to you)
- ☐ Anonymize by first name
- ☐ Anonymize by last name

Scope	Attribute	None	Anonymize	Reduce date	Remove
Individuals	DEAT_DATE	☐	☐	☐	☒
Individuals	DYNASTY	☐	☐	☐	☒
Individuals	BIRT_DATE	☐	☐	☐	☒
Individuals	COMPL	☐	☐	☐	☒
Individuals	DEAT_PLAC	☐	☐	☐	☒
Individuals	BIRT_PLAC	☐	☐	☐	☒



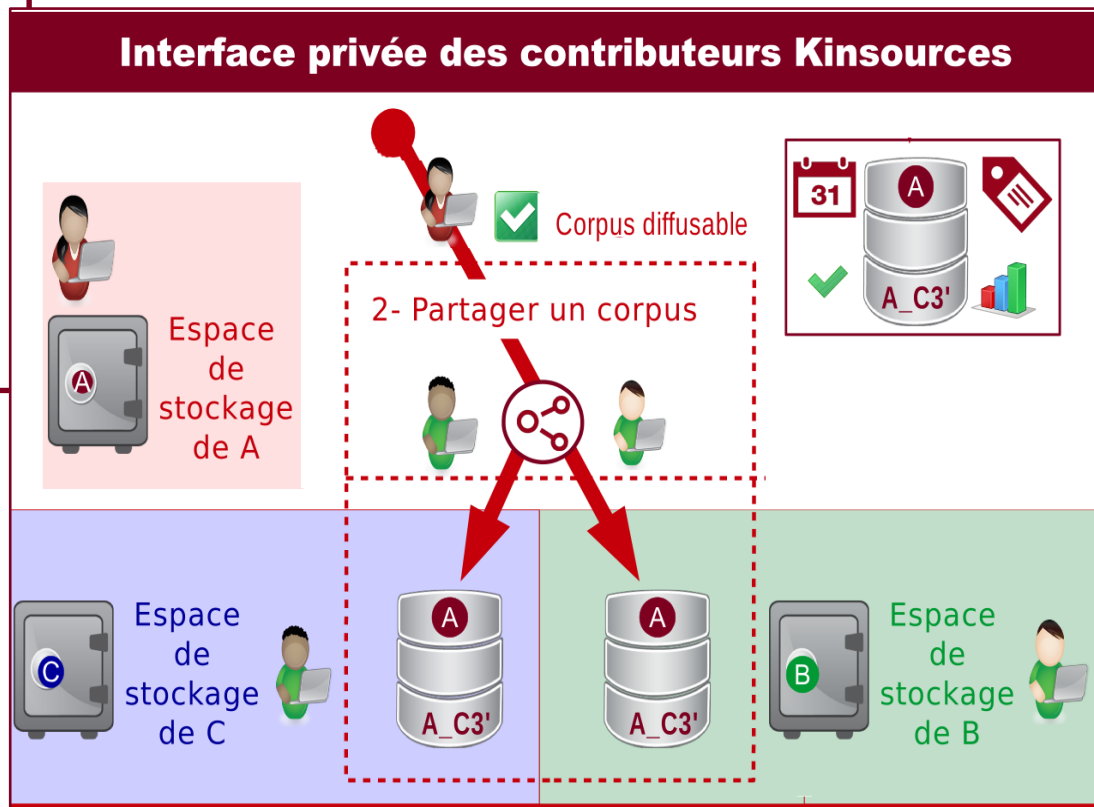
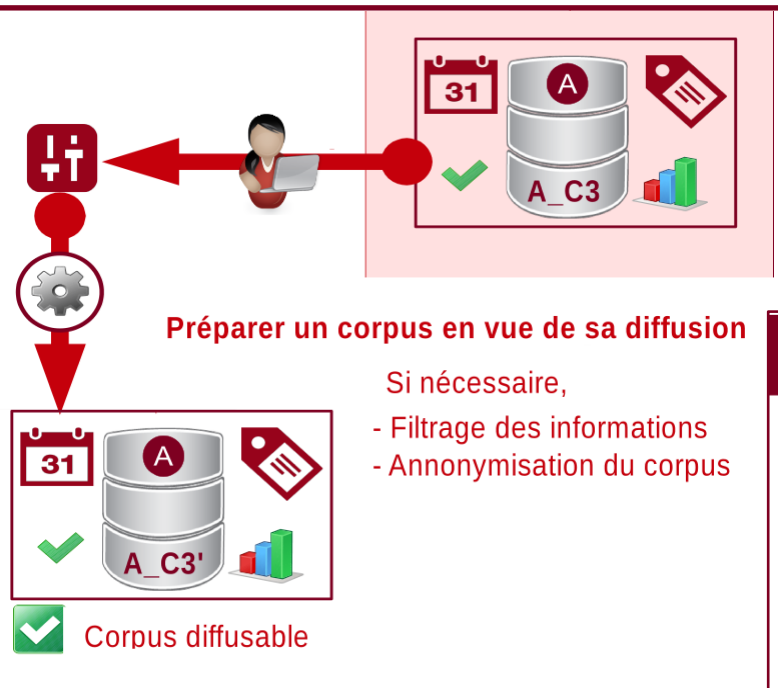
Validation par le conseil scientifique de la plateforme qui ne juge pas du contenu du corpus proprement dit – celui-ci reste de l'entière responsabilité du contributeur –, mais qui vise à s'assurer que **le corpus respecte les critères indispensables à sa diffusion** :

- Une **information suffisante des métadonnées du corpus**, assurant sa contextualisation, notamment : le contexte scientifique de la création des données et le cas échéant les références bibliographiques associées ; la description de l'enquête, des sources mobilisées et du mode d'échantillonnage ; la description des données et des codes utilisés.
- la **conformité du corpus aux obligations légales** en matière de diffusion de données nominatives.

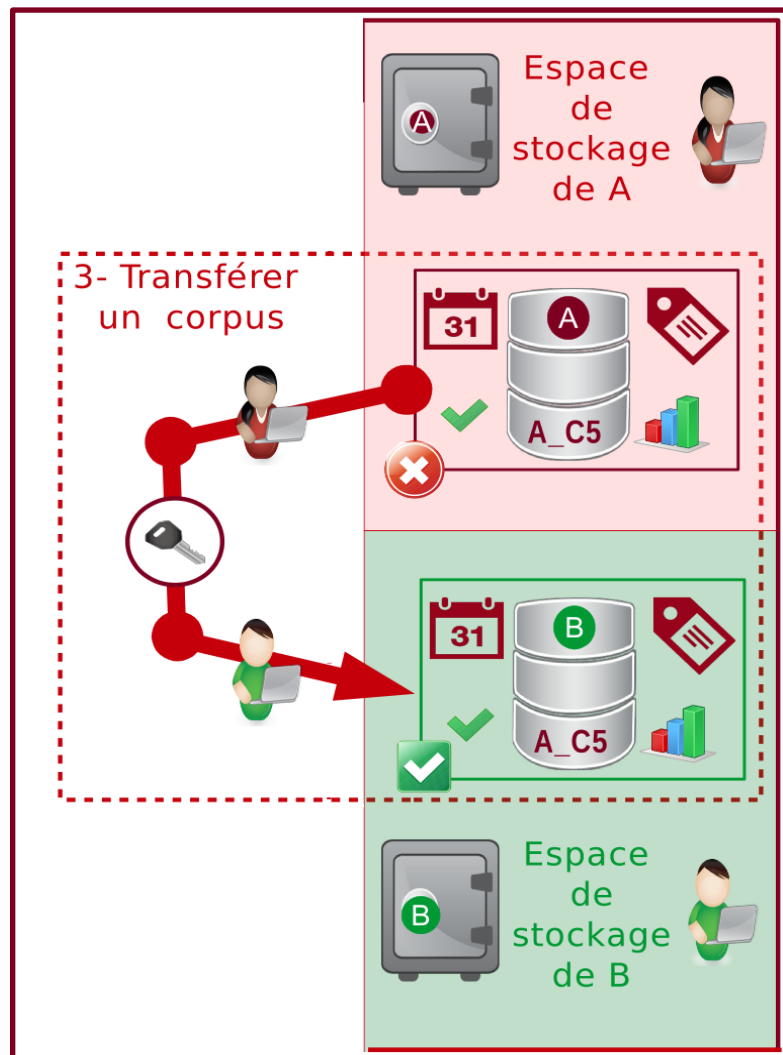


Partage privé d'un corpus

- Partage privé
 - Avec des collaborateurs ou des tiers
 - Choix des informations à partager
 - propre à chaque partage

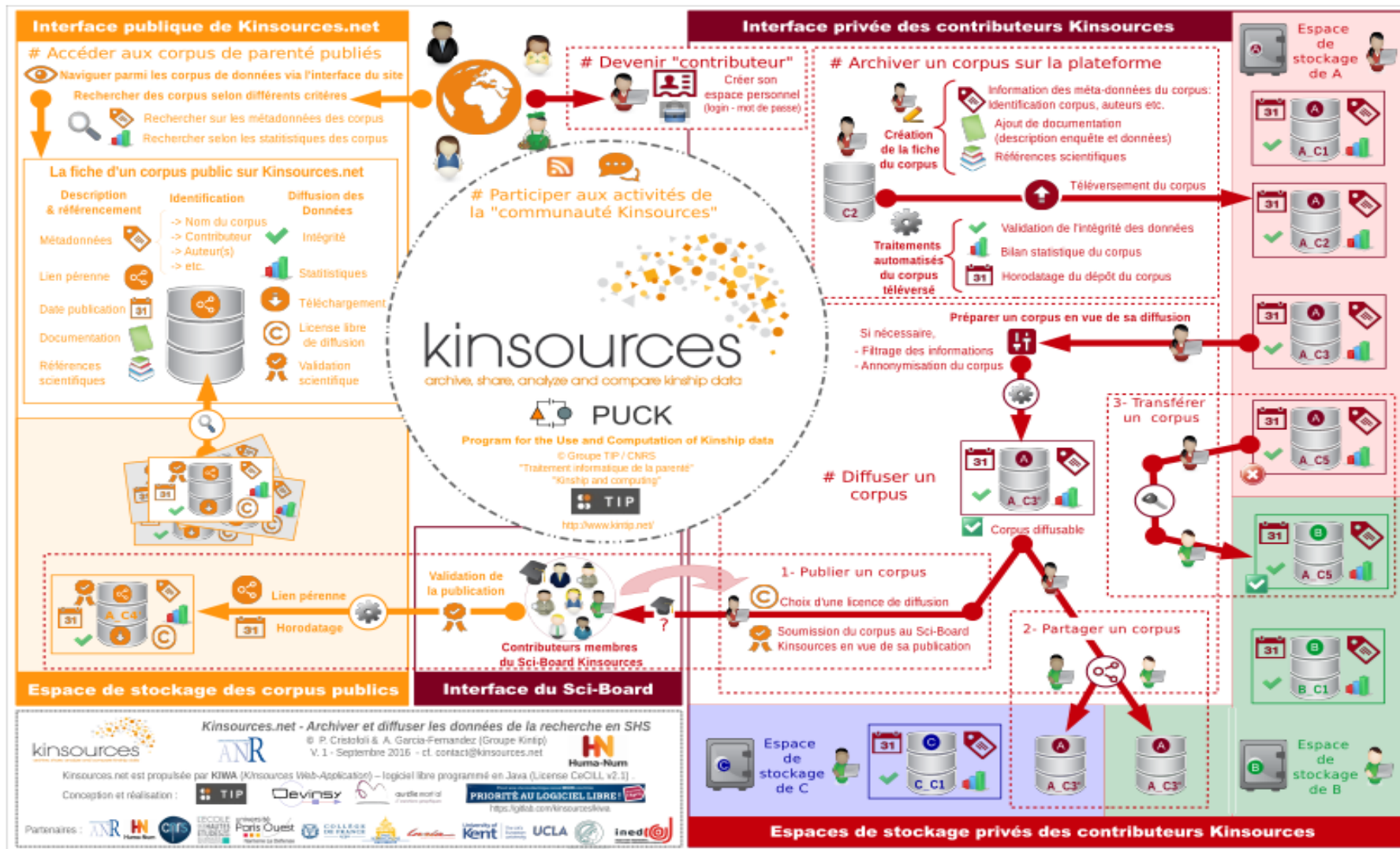


Transférer un corpus



- Partage patrimonial
- Le corpus change de propriétaire

Bref... pour faire simple





- Développement d'outils : logiciel Puck
 - Interface de saisie et de gestion
 - Traitement et analyse de corpus
 - Visualisation généalogique, spatiale...
 - Terminologies de parenté
- Échanges, accompagnement, formations...
 - Documentation, manuels
 - Séminaires, ateliers
 - Forum, liste de diffusion



- Les auto-censures des chercheurs :
 - Quoi et quand publier ?
 - Dimension affective...
- Créer/coder/diffuser un corpus :
 - Un travail ingrat...
 - Comment le valoriser ?
- Pluridisciplinarité :
 - Source d'inspiration
 - Contraintes propres à chaque disciplines
- Peu de recul sur l'utilisation de la plateforme
 - Réutilisation des données
 - Cumul des recherches, des terrains
- Législation et éthique vis à vis des populations étudiées
 - Équilibre entre la pratique et l'intérêt d'une recherche et les contraintes juridico-ethico-déontologique

Logiciels

- **Kinsource.net**

- Plateforme web développée dans le cadre du projet ANR KINSOURCES par l'équipe Kintip, Christian Momon (Devinsy , architecture et développement), Anne Garcia-Fernandez (CNRS, développement et webmastering), et Aurélie Martial (design).
- La plateforme est propulsée par KIWA (KInsources Web-Application), un programme libre (License CeCILL v2.1) écrit en Java.
- Informations sur le projet : contact@kinsources.net
Suggestions et commentaires sur le site : webmaster@kinsources.net

- **Puck**

- Program for the Use and Computation of Kinship data - © Groupe TIP / CNRS - Créé en 2007, publié sous Licence CeCILL 2 - Écrit par Klaus Hamberger, Christian Momon, Edoardo Savoia et Telmo Menezes (en Java 1.6)
- Téléchargement sur www.kintip.net
- Forum des utilisateurs : www.kinsources.net/agora/forums.xhtml

Webographie

- La plateforme Kinsources : <https://www.kinsources.net>
- Le groupe Kintip : <http://kintip.net/>
- Zoom sur... Kinsources et Puck, des outils libres pour l'analyse de données de parenté, Lettre de l'InSHS n° 43, pp.16-26 : http://www.cnrs.fr/inshs/Lettres-information-INSHS/lettre_infoinshs43hd.pdf

Groupe Kintip

- **Kinship and Computing (Traitement informatique de la Parenté)**

- www.kintip.net
- Arnaud Bringé (INED), Marie-Hélène Cazes (INED), Pascal Cristofoli (EHESS), Isabelle Daillant (CNRS), Pietro Fornasetti (IMAF), Laurent Gabail (Univ. Toulouse II), Anne Garcia-Fernandez (CNRS), Michael Gasperoni (CNRS), Cyril Grange (CNRS), Klaus Hamberger (EHESS), Vincent Hirtzel (CNRS), Michael Houseman (EPHE), Olivier Kyburz (Univ. Paris Ouest Nanterre), Ismaël Moya (CNRS).

- Barry L., 1998, « **Les modes de composition de l'alliance. Le "mariage arabe"** », L'Homme 147, pp. 17-50.
- Barry L. et Gasperoni M., 2008, « **L'oubli des origines. Amnésie et information généalogiques en histoire et en ethnologie** », Annales de démographie historique 116, pp. 53-104.
- Cristofoli P. et Hamberger K., 2014, « **Kinsources & Puck – Open data and open tools for kinship network analysis** », poster présenté à la 1st European Conference on Social Networks, Université de Barcelone, 3 juillet 2014. Posters en version anglaise, française et espagnole.
- Groupe Kintip, 2016, « **Open Data and Free Tools to Archive, Share, Analyse, and Compare Kinship Data** », poster présenté à la 2nd European Conference on Social Networks, Sciences Po Paris, 14-17 juin 2016. Posters en version anglaise.
- Gilbert J. P. et Hammel E. A., 1966, « Computer simulation and analysis of problems of kinship and social structure », American Anthropologist 68 (1), pp. 71-93.
- Goldberg H., 1967, « FBD marriage and demography among Tripolitanian Jews in Israel », Southwestern Journal of Anthropology 23 (2), pp. 177-191.
- Hamberger K. et Daillant I. 2008, « **L'analyse de réseaux de parenté. Concepts et outils** », Annales de démographie historique 116, pp. 13-52.
- Hamberger K. et Gargiulo F., 2014, « **Virtual Fieldwork : Modeling Observer Bias in Kinship and Marriage Alliance Networks** », Journal of Artificial Societies and Social Simulation, 17 (3)2, DOI: 10.18564/jasss.2495
- Hamberger K., Grange C., Houseman M. et Momon C., 2014, « Scanning for patterns of relationship: analyzing kinship and marriage networks with Puck 2.0 », The History of the Family 19 (4), pp. 564-596 (DOI: 10.1080/1081602X.2014.892436).
- Hamberger K., Houseman M. et Grange C., 2009, « **La parenté radiographiée. Un nouveau logiciel pour l'analyse des réseaux matrimoniaux** », L'Homme 191, pp. 107-137.
- Hamberger K., Houseman M. et White D. R., 2011, « Kinship Network Analysis », in Carrington P. et Scott J. (dir.) The Sage Handbook of Social Network Analysis, Sage Publications, Londres, pp. 533-549.
- Houseman M. et White D. R., 1996, « **Structures réticulaires de la pratique matrimoniale** », L'Homme 139, pp. 59-85.
- Houseman M. et White D. R., 1998, « Taking Sides : Marriage Networks and Dravidian Kinship in Lowland South America », in Godelier M., Trautmann T. R. et Tjon Sie Fat F. E. (ed.) Transformations of Kinship, Smithsonian Institution Press, Washington, pp. 214-243.
- Kasakoff A. B. et Adams J. W., 1977, « Spatial Location and Social Organisation : an Analysis of Tikopian patterns », Man (n.s.) 12 (1), pp. 48-64.
- Kunstadter R., Buhler R., Stephan F. F. et Westoff Ch. F., 1963, « Demography and preferential marriage », American Journal of Physical Anthropology 21, pp. 511-519.
- Menezes T. et Roth C., 2013, « **Automatic discovery of Agent-Based Models : An application to Social Anthropology** », Advances in Complex Systems, vol. 16 (7): 1350027. DOI : <http://dx.doi.org/10.1142/S0219525913500276>
- Roth C., Gargiulo F., Bringé A. et Hamberger K., 2013, « Random alliance network », Social Networks, vol. 35 (3), pp. 394-405.
- Selz M., 1994, « **Traitement informatique de données généalogiques : le logiciel "GEN-PAR"** », L'Homme 130, pp. 129-136.
- Surrallés i Calonge A., 1999, Au cœur du sens. Objectivation et subjectivation chez les Candoshi de l'Amazonie péruvienne, thèse de doctorat, EHESS, Paris.
- White D. R. et Jorion P., 1992, « Representing and Analysing Kinship : A Network Approach », Current Anthropology 33, pp. 454-463.

Merci de votre attention !