Kinsources et Puck
Données ouvertes et outils libres pour l'analyse des réseaux de parenté

kinsources
archive, share, analyze and compare kinship data

Groupe « Kintip » - Traitement informatique de la Parenté (Kinship and Computing)

Arnaud Bringé (INED), Marie-Hélène Cazes (INED), Pascal Cristofoli (EHESS), Isabelle Daillant (CNRS), Laurent Gabail (Oxford University), Anne Garcia-Fernandez (CNRS), Michael Gasperoni (EHESS), Cyril Grange (CNRS), Klaus Hamberger (EHESS), Vincent Hirtzel (Univ. Paris X), Michael Houseman (EPHE), Olivier Kyburz (Univ. Paris X), Ismaël Moya (CNRS)

Puck

www.kintip.net

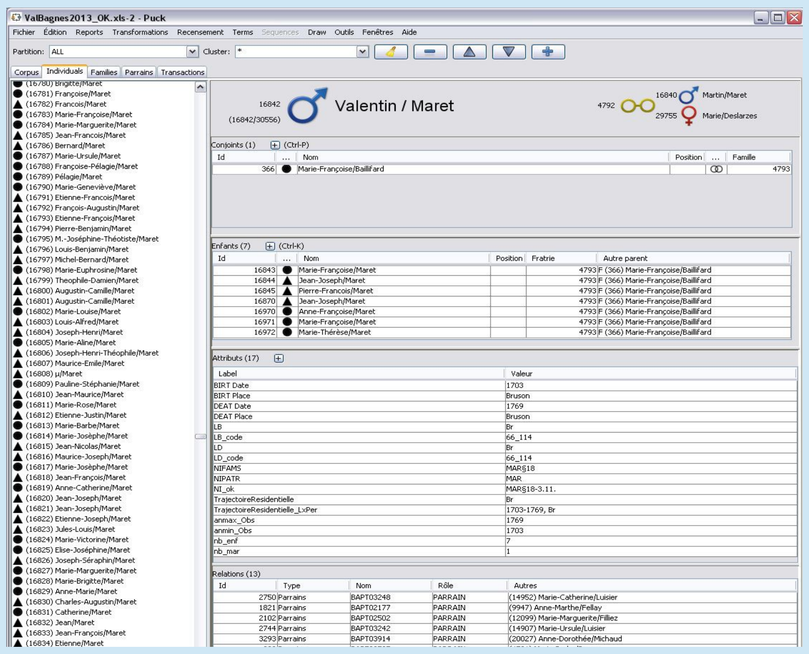
Program for the use and computation of kinship data

Puck (*) est un logiciel multi-plateforme libre, puissant et évolutif qui permet la saisie et la manipulation de données individuelles et relationnelles ainsi que leur analyse à l'aide de segmentations, statistiques, recensements relationnels et simulations.

Le logiciel Puck s'appuie sur :

- Des formats simples et ouverts permettant d'assurer la conservation des données : formats textuels tabulaires et format natif Puck (XML).
- Une interopérabilité avec les outils standards de traitement des données : Open office, Microsoft Excel, Gedcom (logiciels de généalogie), Pajek...
- La possibilité de coder et de travailler de manière coordonnée sur des données de nature différente : informations nominatives, données de parenté, réseaux sociaux, et, à venir, les terminologies de parenté et données à référence spatiale.

Créer et manipuler des données individuelles et relationnelles :

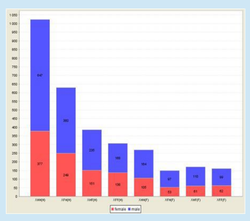


Créer
- Saisir
- Coder
- Importer
- Exporter
- Transformer
- Mettre à jour
- Fusionner

Manipuler
- Rechercher
- Naviguer
- Explorer
- Sélectionner
- Visualiser

Analyser un corpus :

Produire des statistiques générales sur un corpus
- Diagnostic d'un corpus (erreurs de cohérence, biais éventuels)
- Bilan statistique sur la structure et les propriétés d'un corpus



Segmenter un corpus
- Créer des sous-populations à partir de critères multiples
- Réduire, étendre et extraire des segments
- Agréger les données, construire des réseaux d'alliance et des réseaux de flux

Établir des recensements relationnels :

Recensements matrimoniaux

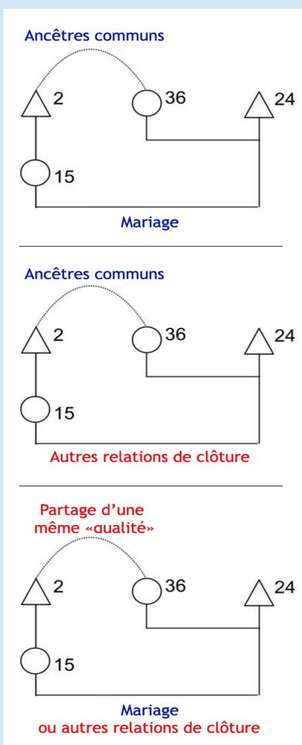
Dénombrement des circuits matrimoniaux et analyse de leur configurations

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2	96 circuits (maximal depths = [2])											
3	4 circuit types (average frequency = 24.0)											
4	90 couples concerned (4.11 % of 2190 couples examined)											
5	175 individuals concerned (2.97 % of 5890 individuals examined)											
6												
7												
8	96 circuits of order 1 (maximal depth = 2) (100.0 %)											
9	4 circuit types (100.0 %) (average frequency = 24.0)											
10	90 couples concerned (100.0 %), 4.11% of all couples											
11	175 individuals concerned (100.0 %), 2.97% of all individuals											
12												
13	Nr	Standard	Positional	Vector	#Circuits	% Circuits	# Couples	% Couples	% All Couples	# Individuals	% Individuals	% All Individuals
14	1	FBD	HH()HF	[-7, -8]	24	25.0	24	26.67	1.1	48	27.43	0.81
15	2	FZD	HH()FF	[-7, -10]	23	23.96	23	25.56	1.05	46	26.29	0.78
16	3	MBD	HF()HF	[-9, -8]	30	31.25	30	33.33	1.37	60	34.29	1.02
17	4	MZD	HF()FF	[-9, -10]	19	19.79	19	21.11	0.87	38	21.71	0.65
18												

	A	B	C	D	E	F	G
1	Sortable list						
2							
3	Nr.	Standard	Positional	Vector	Ego	Alter	Pivots	Chain
4	1	FBD	HH()HF	[-7, -8]	192	204	Charles / IV / d'Espagne = Marie Louise / de Parme	192 152 () 148 204
5	1	FBD	HH()HF	[-7, -8]	320	319	Joseph / du Portugal = Bénédicte / du Portugal	320 213 () 311 319
6	1	FBD	HH()HF	[-7, -8]	337	339	Philippe / II (Le Regent) / d'Orléans = Françoise Marie / de Blois	337 25 () 58 339
7	1	FBD	HH()HF	[-7, -8]	566	567	Ernest Louis / I / de Saxe-Meiningen = Dorothée / de Saxe-Gotha-Altenbourg	566 850 () 541 567

Recensements relationnels généralisés

La relation de clôture n'est plus un mariage:
relations d'équivalence (même propriété), relations exogènes, chaînes ouvertes



Recensements semi-classificatoires et circuits matrimoniaux mixtes

Circuits matrimoniaux ou relationnels intégrant une relation non généalogique entre les « ancêtres »

Création de relations à partir des recensements

Simuler des corpus de parenté :

Simulations paramétrées, simulations par transformation (« Reshuffling ») ou par extraction (« Virtual Fieldwork »), et simulation inverse (« Machine Learning »)

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

Références : Hamberger et Daillant (2008), « L'analyse de réseaux de parenté. Concepts et outils », *Annales de Démographie Historique* 116, 13-52 ; Hamberger, Houseman et White (2011), « Kinship Network Analysis », In P. Carrington et J. Scott (dir.) *The Sage Handbook of Social Network Analysis*, London, Sage Publications, 533-549 ; Hamberger, Houseman et Grange (2014), « Scanning for patterns of relationship: analyzing kinship and marriage networks with Puck 2.0 », *History of the Family*

Kinsources

www.kinsources.net

Archive, share, analyse and compare kinship data

Kinsources.net est une plateforme web ouverte et interactive dédiée à l'archivage, au partage et à la comparaison des données de parenté (généalogiques, terminologiques, résidentielles et relationnelles).

La plateforme Kinsources propose :
- Un environnement et des fonctionnalités simples permettant aux chercheurs de publier et diffuser eux-mêmes leurs données tout en valorisant leur travail et en protégeant leurs droits.
- Une opportunité de participer au mouvement d'ouverture des données en SHS en favorisant leur archivage, leur cumul et en assurant les conditions d'une reproduction des analyses scientifiques.

De nombreux corpus d'anthropologues, d'historiens, et de démographes sont déjà publiés sur le site Kinsources.net.

N'hésitez pas à découvrir le site et à télécharger des corpus !

Archiver ses données :

Sauvegarder ses données
sur un entrepôt pérenne et sécurisé adossé à un opérateur public

Décrire ses données
- Informer les métadonnées des corpus déposés
- Associer les documentations spécifiques et les références de publications les concernant

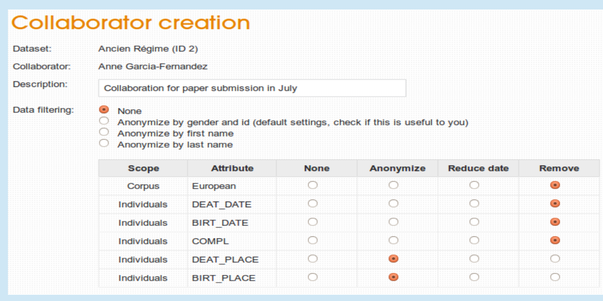
Gérer ses dépôts de données
- Chaque contributeur Kinsources contrôle son espace personnel
- A venir :
  Gestion des versions de corpus
  Gestion de collections de corpus

Partager ses données :

Partager un corpus avec des collaborateurs

Publier un corpus
- Contrôle de la cohérence des données
- Validation par un conseil scientifique
- Génération d'un lien permanent associé

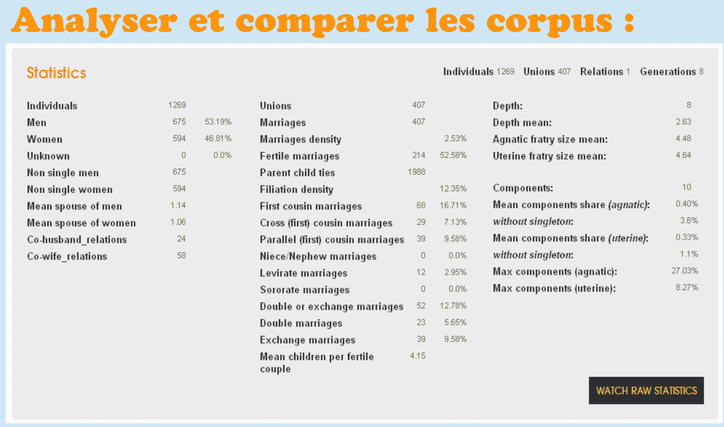
Ex : https://www.kinsources.net/kidarep/dataset-24-ayd-nl-yoruk-2005.xhtml



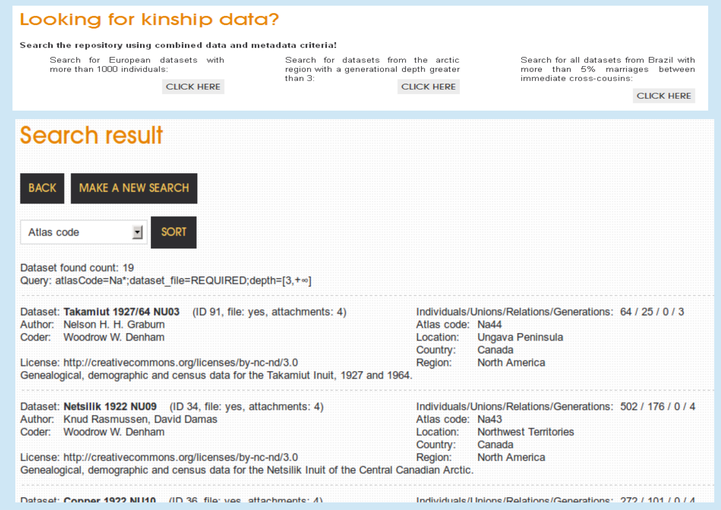
Choisir les modalités du partage et de la publication

Avec ou sans anonymisation, types d'informations diffusées, licences de diffusion, etc.

Analyser et comparer les corpus :



Analyser les corpus
- Consulter le rapport statistique associé à un corpus
- Télécharger un corpus
- Possibilité d'analyses complexes à l'aide des outils de traitement (Puck, KinOath)

Rechercher des corpus selon plusieurs critères de sélection
- Selon les caractéristiques des métadonnées des corpus (période, lieu, langage etc.)
- Selon les propriétés statistiques des corpus (profondeur générationnelle, pourcentage de mariages entre cousins, etc.)

Entrer en relation avec les contributeurs
- Proposer corrections
- Compléter des corpus
- Contribuer à des analyses collectives