KinOath Kinship Archiver Version 1.1

Developed by Peter Withers at

The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

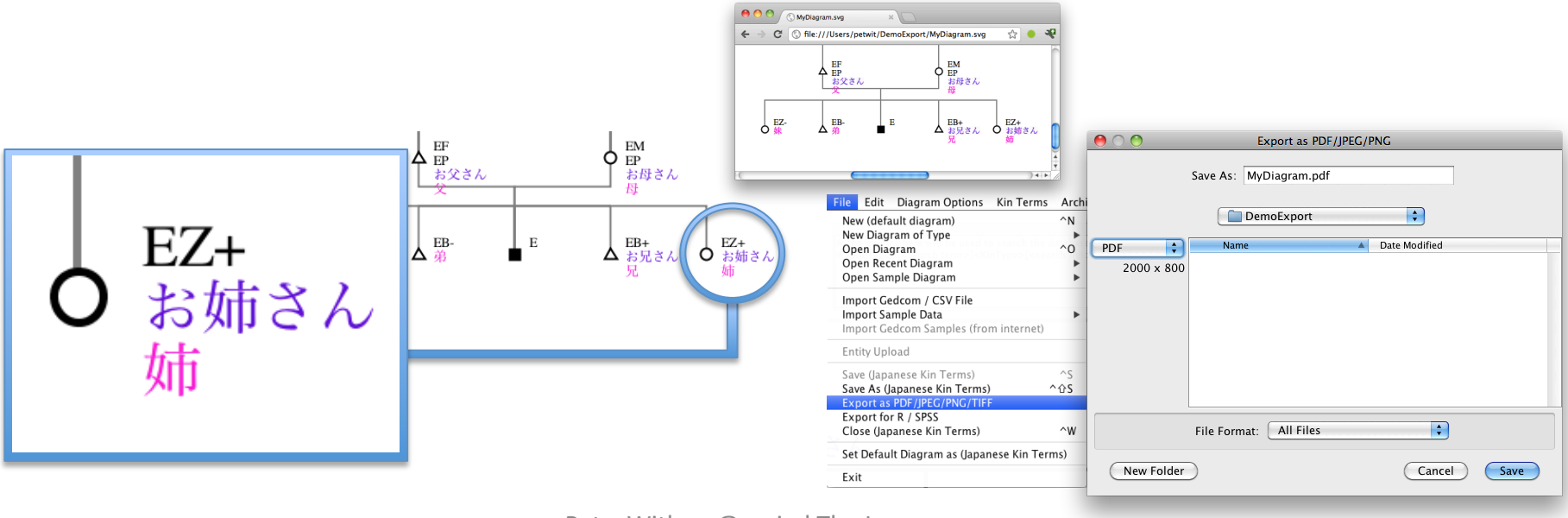
Second Stable Release in Testing

- The stable version 1.0 has been field tested.
- Formal testing has begun on version 1.1.
- Plugins have been introduced in version 1.1.
- The manual is now available on line via the website.
- The manual is also built into in the application.
- This testing version is also available for download, if you want to preview the new features.
- The change list is available on the download page.
- The todo list is also available on the download page.
- http://tla.mpi.nl/tools/tla-tools/kinoath/

Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Publishable Diagrams

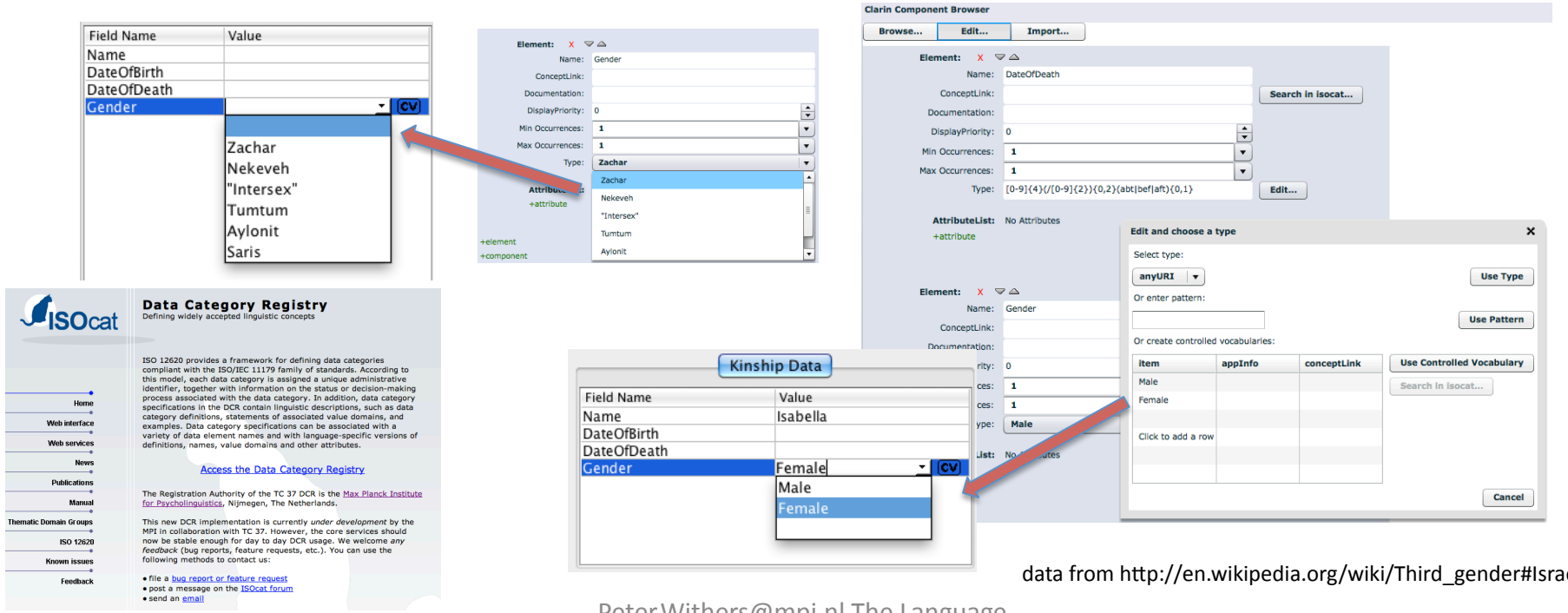
- All the diagrams produced are in a scalable vector format of publishing quality
- All diagrams can be exported into PDF format
- The working files are in SVG format that can be opened in graphics applications or viewed in a web browser



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Customisable Kin Data

- The kin data is defined in the Clarin component registry.
- Each data field can be defined in the ISOcat data category registry.
- http://catalog.clarin.eu/ds/ComponentRegistry
- http://www.isocat.org/
- This means that you can specify the data fields that you need for your project.

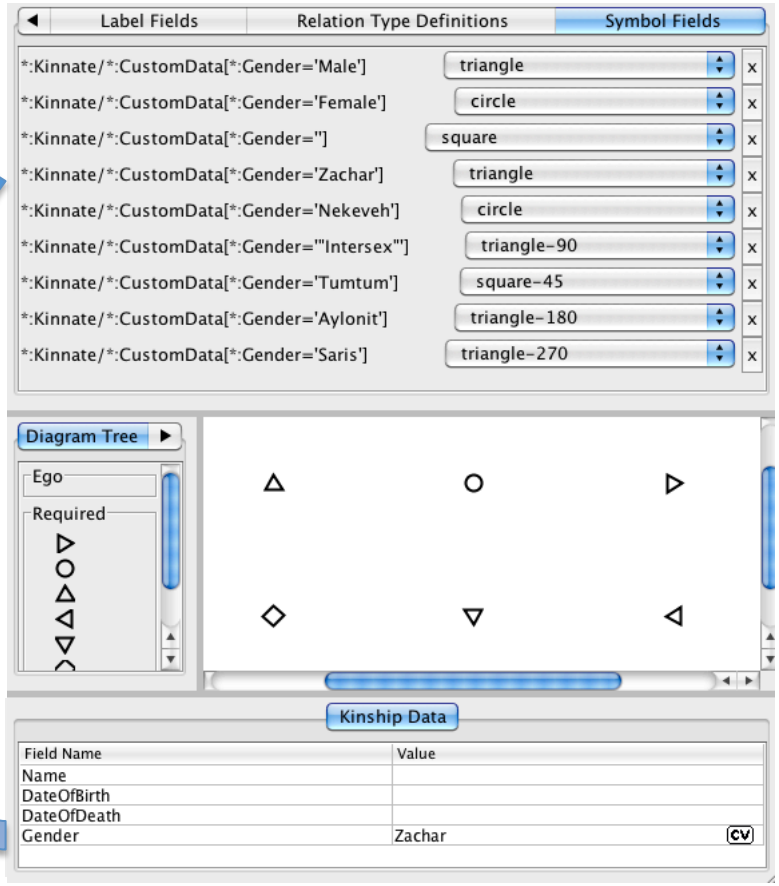


data from http://en.wikipedia.org/wiki/Third_gender#Israel

Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Custom Symbols

- A symbol can be any SVG definition.
- Custom symbols can only be manually inserted at this sage.
- A number of symbols are already included in the default diagram.
- These symbols can be associated with any kin data.
- The kin data field can be dragged to the symbol definition list.
- Then a symbol can be selected from the diagram definitions.



Gender data from http://en.wikipedia.org/wiki/Third_gender#Israel

Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Custom Kin Types

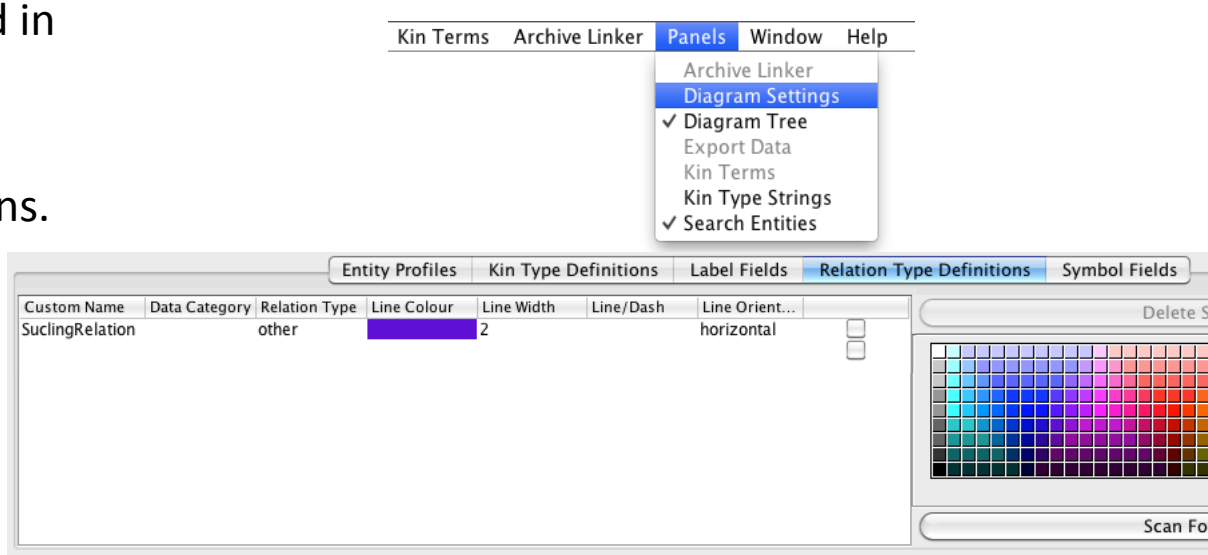
- The kin types used in the application can be customized.
- Each kin type can use any string and any symbol.
- These custom kin types are stored in the diagram file.
- They can also be stored in the default diagram of the application.

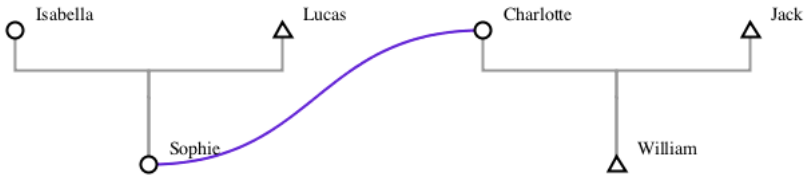
Kin Type String	Relation Type	Symbol Type	Display Name
Ef	none	circle	Ego Female
Em	none	triangle	Ego Male
Fa	ancestor	triangle	Father
Mo	ancestor	circle	Mother
Br	sibling	triangle	Brother
Si	sibling	circle	Sister
So	descendant	triangle	Son
Da	descendant	circle	Daughter
Hu	union	triangle	Husband
Wi	union	circle	Wife
Pa	ancestor	square	Parent
Sb	sibling	square	Sibling
Sp	union	square	Spouse
Ch	descendant	square	Child
F	ancestor	triangle	Father
M	ancestor	circle	Mother
B	sibling	triangle	Brother
Z	sibling	circle	Sister
S	descendant	triangle	Son
D	descendant	circle	Daughter
H	union	triangle	Husband
W	union	circle	Wife
P	ancestor	square	Parent
G	sibling	square	Sibling
E	none	square	Ego
C	descendant	square	Child
m	none	triangle	Male
f	none	circle	Female
x	none	square	Undefined
*			Any Relation

F	ancestor	triangle	Father
M	ancestor	circle	Mother
B	sibling	triangle	Brother
Z	sibling	circle	Sister
S	descendant	triangle	Son
D	descendant	circle	Daughter
H	union	triangle	Husband
W	union	circle	Wife
P	ancestor	square	Parent
G	sibling	square	Sibling
E	none	square	Ego
C	descendant	square	Child

Custom Relation Types

- Custom relation types can be defined in a diagram.
- This can be done in the diagram settings under relation type definitions.
- The custom name, type and display style can be entered.
- The new relation type will then be available on the selected entities.
- These relations can be created by dragging the relation handles.

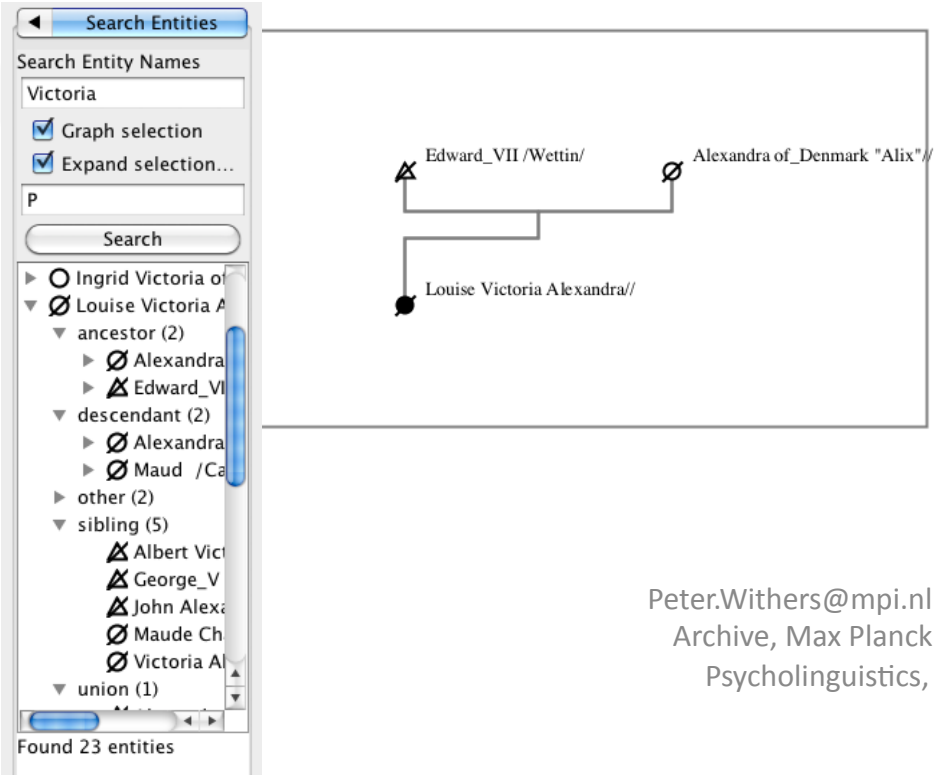




Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Search Tree

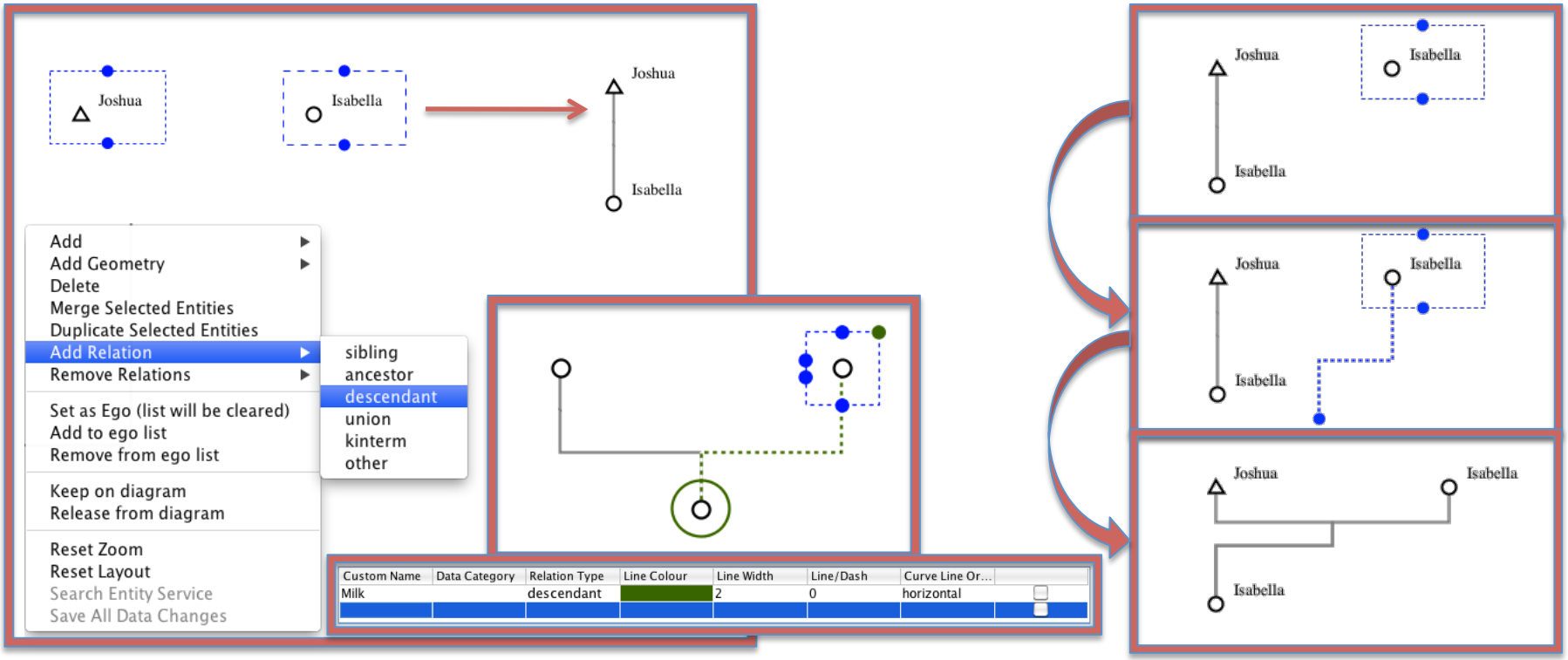
- Free text can be searched with fuzzy matching.
- The selected results can be inspected on the graph.
- The selection can be expanded on the graph by kintype string.
- As the selection changes the graph is updated by kintype string.
- The tree can also be browsed for relations.



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Creating Relations

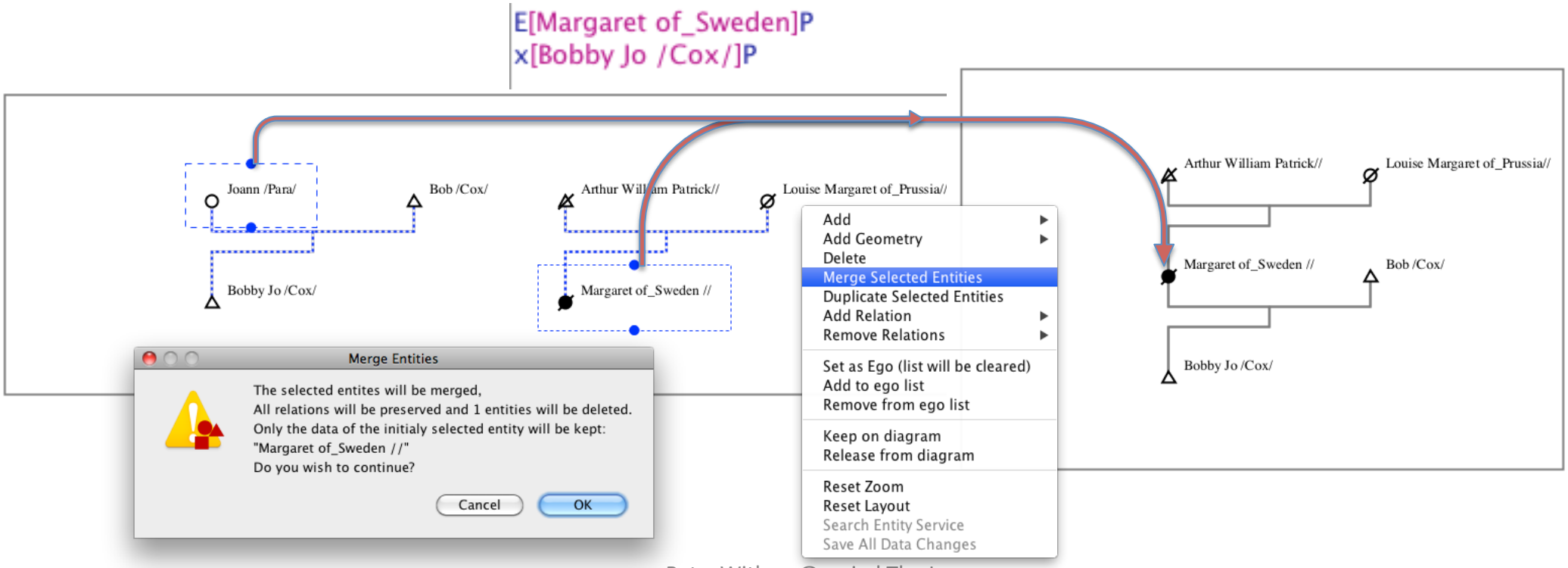
- Relations can be added via the context menu.
- Or via the drag handles of the selection.
- Custom relations are also available via the drag handles.



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Merge Individuals

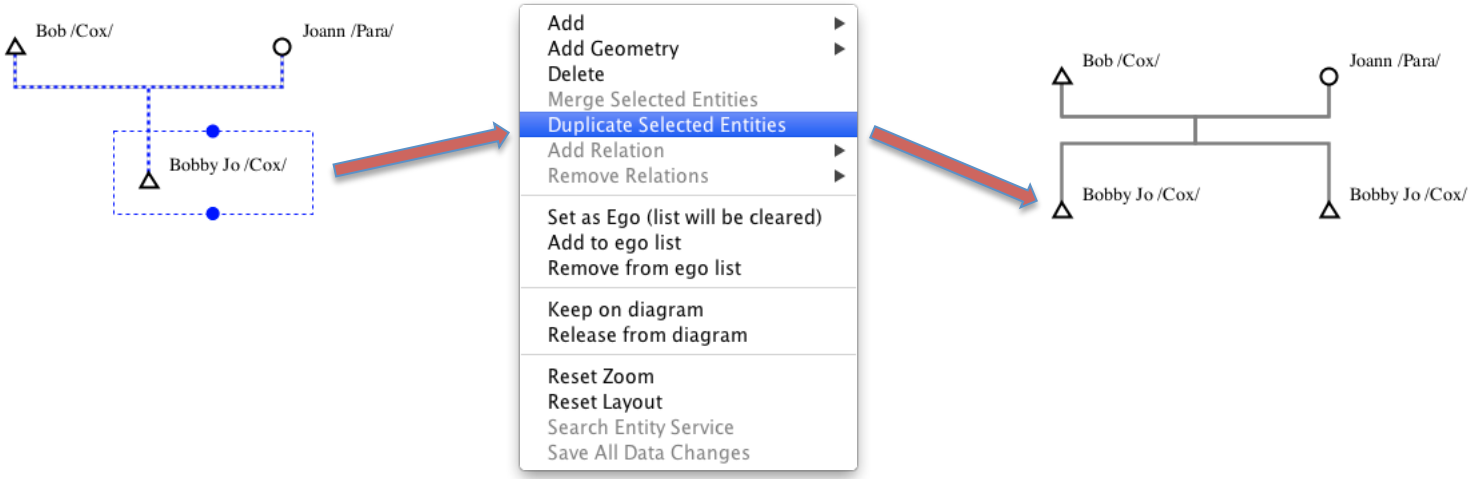
- When duplicate entries are found individuals can be merged, keeping the relations of both originals.



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Duplicating Individuals

- An individual can be duplicated.
- This will duplicate all relations.
- All kin data is also duplicated.
- Both individuals can then be edited separately.



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Gedcom Import

sample.ged

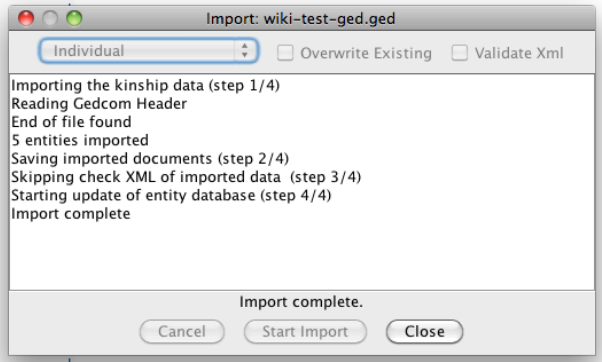
```
0 HEAD
1 SOUR Reunion
2 VERS V8.0
2 CORP Leister Productions
1 DEST Reunion
1 DATE 11 FEB 2006
1 FILE test
1 GEDC
2 VERS 5.5
1 CHAR MACINTOSH
0 @I1@ INDI
1 NAME Bob /Cox/
1 SEX M
1 FAMS @F1@
1 CHAN
2 DATE 11 FEB 2006
0 @I2@ INDI
1 NAME Joann /Para/
1 SEX F
1 FAMS @F1@
1 CHAN
2 DATE 11 FEB 2006
0 @I3@ INDI
1 NAME Bobby Jo /Cox/
1 SEX M
1 FAMC @F1@
1 CHAN
2 DATE 11 FEB 2006
0 @F1@ FAM
1 HUSB @I1@
1 WIFE @I2@
1 MARR
1 CHIL @I3@
0 TRLR
```

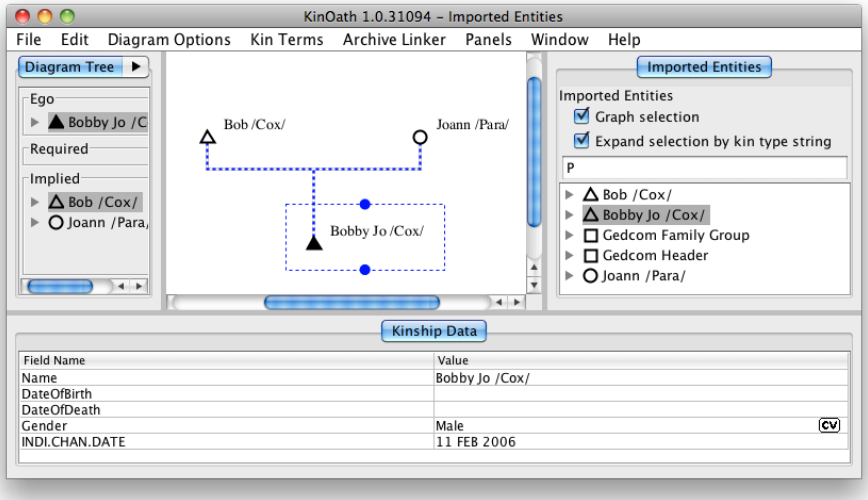
- GEDCOM is the most common kinship data format.
- These files can be imported via the file menu.
- All GEDOM fields are imported.
- Resource links are preserved.
- For testing the "GEDCOM 5.5 Torture Test Files" are used:

http://www.geditcom.com/gedcom.html

- This simple example is from the Wikipedia Gedcom page:

http://en.wikipedia.org/wiki/Gedcom





Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

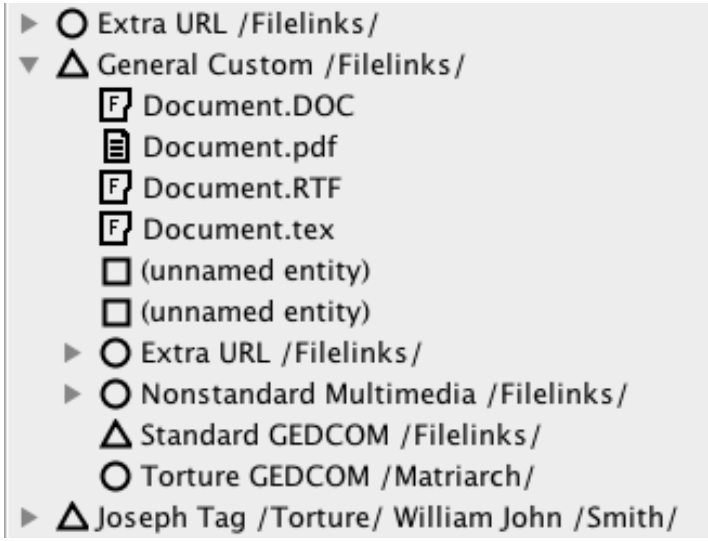
New Import Formats

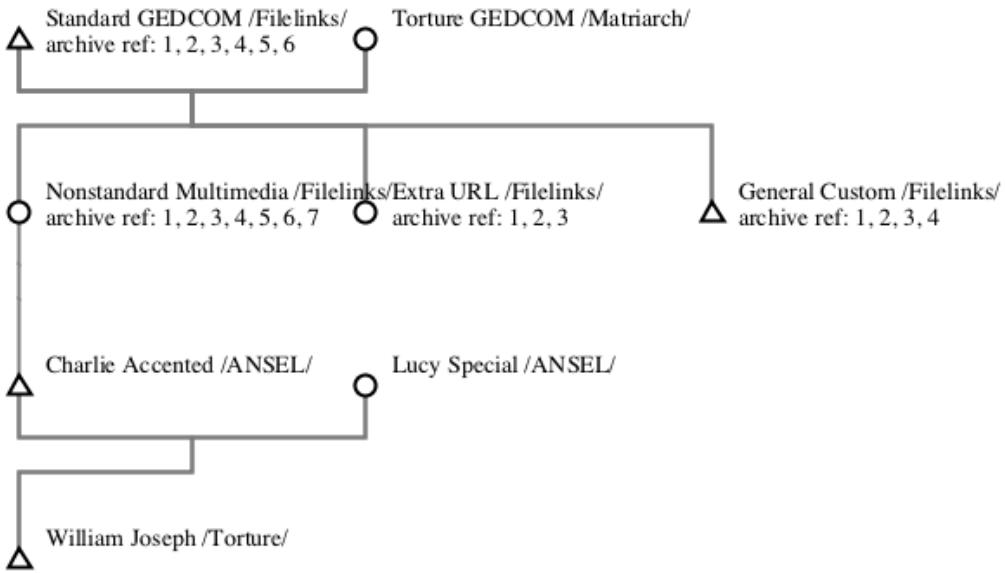
- GEDCOM import had been improved
- TIP format is now supported
- CSV import is now extended to include:
  - Optional 'ID' column.
  - If the ID column contains the same value twice, then the preceding record will be updated/appended to.
  - Lines starting with * will be treated as comments and ignored.
  - Recommended data columns are: Name, Gender, DateOfBirth, DateOfDeath
  - Relation columns must map to the values in the ID column.
  - Recognised relation columns are: Spouse, Union, Parent, Father, Mother
  - Any other columns will be added to the kindata but not automatically used in the subsequent diagrams.
  - If the ID column exists and any row contains the string `ID' then the current column names are replaced with current row values (for PUCK txt files).
  - If the first record does not start at zero then all relations referencing ID of zero will be ignored (for PUCK txt files).

Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Imported Resource Files

- Like manually created resource links, imported GEDCOM resources files are available both on the diagram and via the diagram tree.



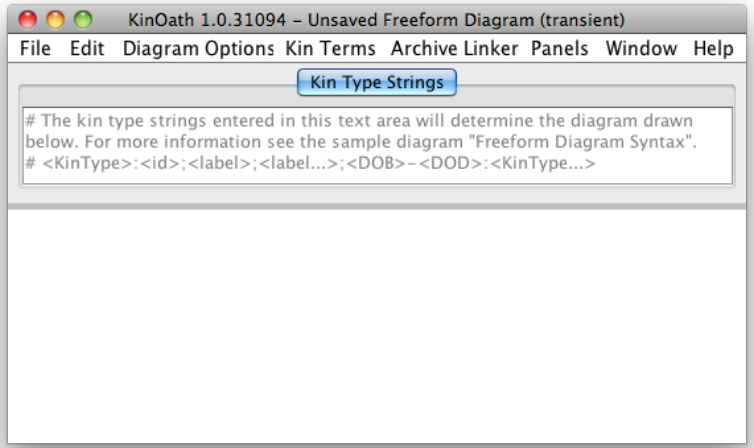


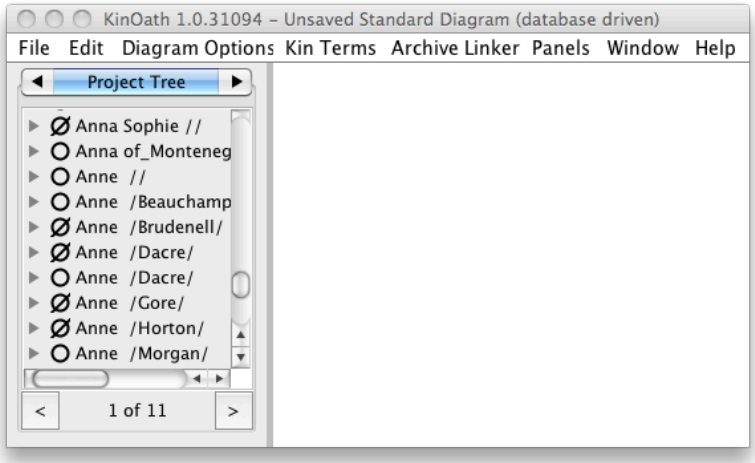
Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Freeform vs Database Diagrams

- Freeform diagrams have no kinship data records.
- The data only exists on the diagram.
- The kin type strings or kin terms definitions are the only data source for the diagram.

- Database diagrams display data from the local project.
- There are kin data files for each entity.
- Changes to kin data is reflected on all diagrams using that project data.
- Data can be imported from GEDCOM, TIP or CSV.



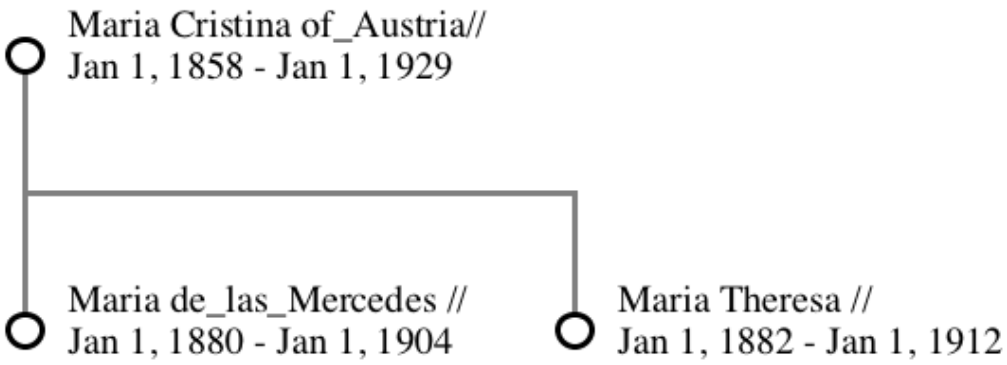


Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Kin Type Query Syntax

- Kin type queries are used to retrieve individuals and their relations from the database.
- The kin types following a query will add any matching kin.
- For example to following query returns Maria and her daughters.

x[Maria Cristina of_Austria]D



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Multiple Query Parameters

- Multiple queries can be used per kin type.
- Each condition can use:

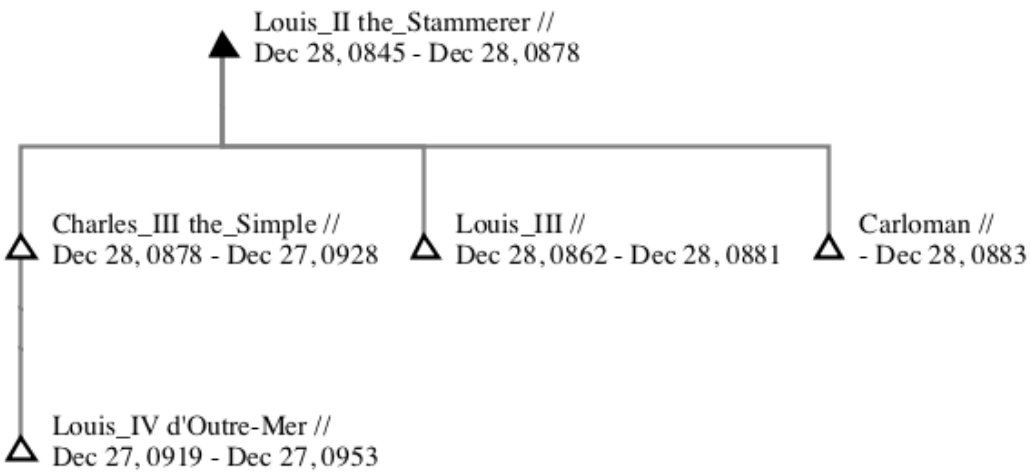
  = contains

  == exact match

  > greater than

  < less than

E[DateOfBirth<0850][INDI.TITL=King of France]CC



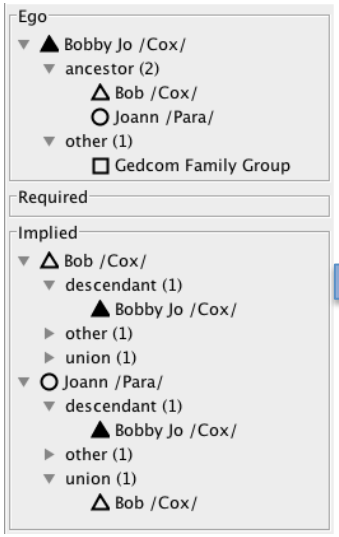
Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Dragging to the Query Text

- The table is visible when an entity is selected on the tree or diagram.
- Fields from this table can be dragged to the query area to construct queries.
- Entities in the tree can also be dragged into the query, in this case the entity ID is added to the query.

E[Entity.Identifier=70596686bd36929ace97868ec7f412a8]

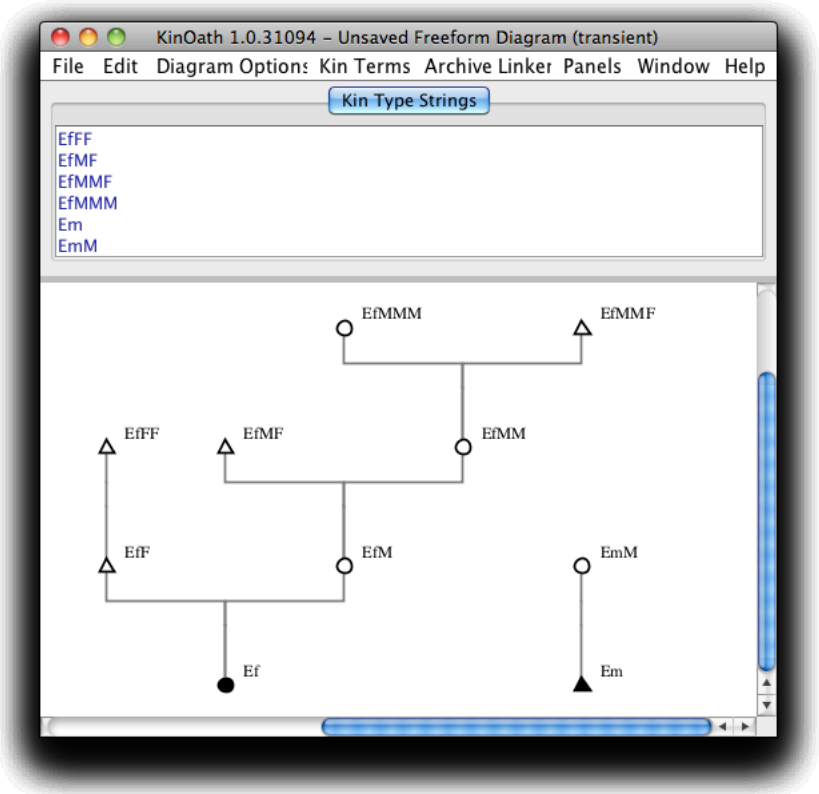
E[INDI.NAME=Louis_II the_Stammerer //]



Field Name	Value
NAME	Louis_II the_Stammerer //
TITL	King of France
SEX	M
BIRT.DATE	846
DEAT.DATE	879

Freeform Diagrams

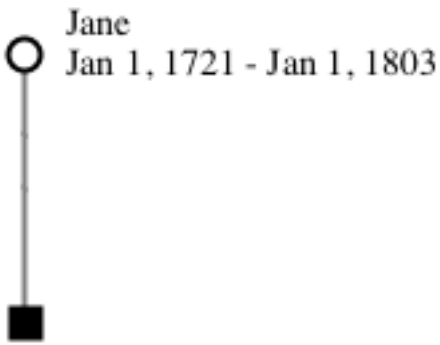
- Freeform diagrams are constructed simply by entering kin type strings.
- They have no database.
- The data only exists on the diagram.
- Kin Terms can also be used to generate these diagrams.

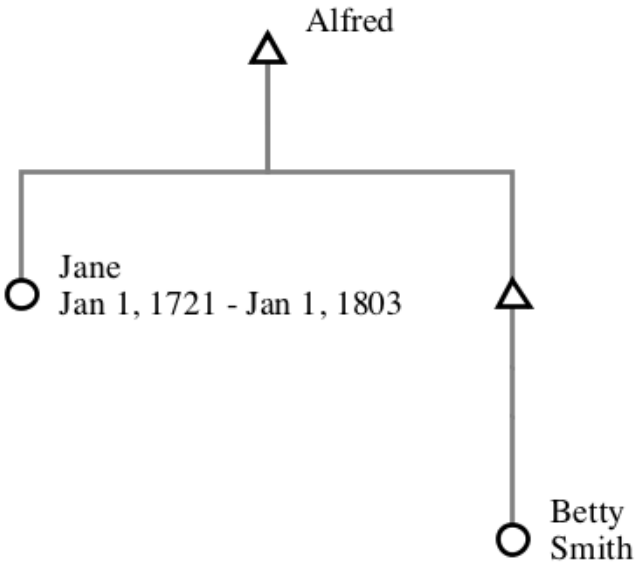


Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Freeform Diagram Syntax

- The syntax allows labels, dates and identifiers to be specified.
- <KinType>:<id>;<label>;<label...>;<DOB>-<DOD>:<KinType...>
- EM:Jane;1721-1803:
- f:#3;Jane;1721-1803:F:Alfred:SD:Betty;Smith:

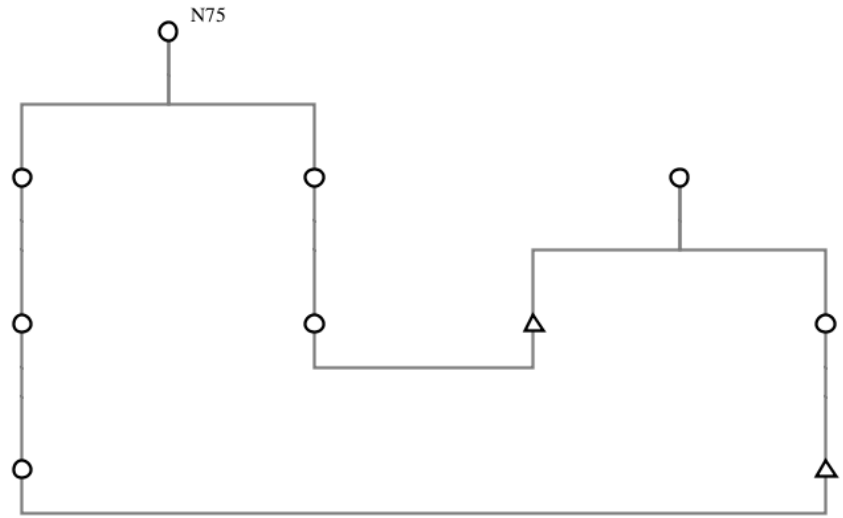




Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

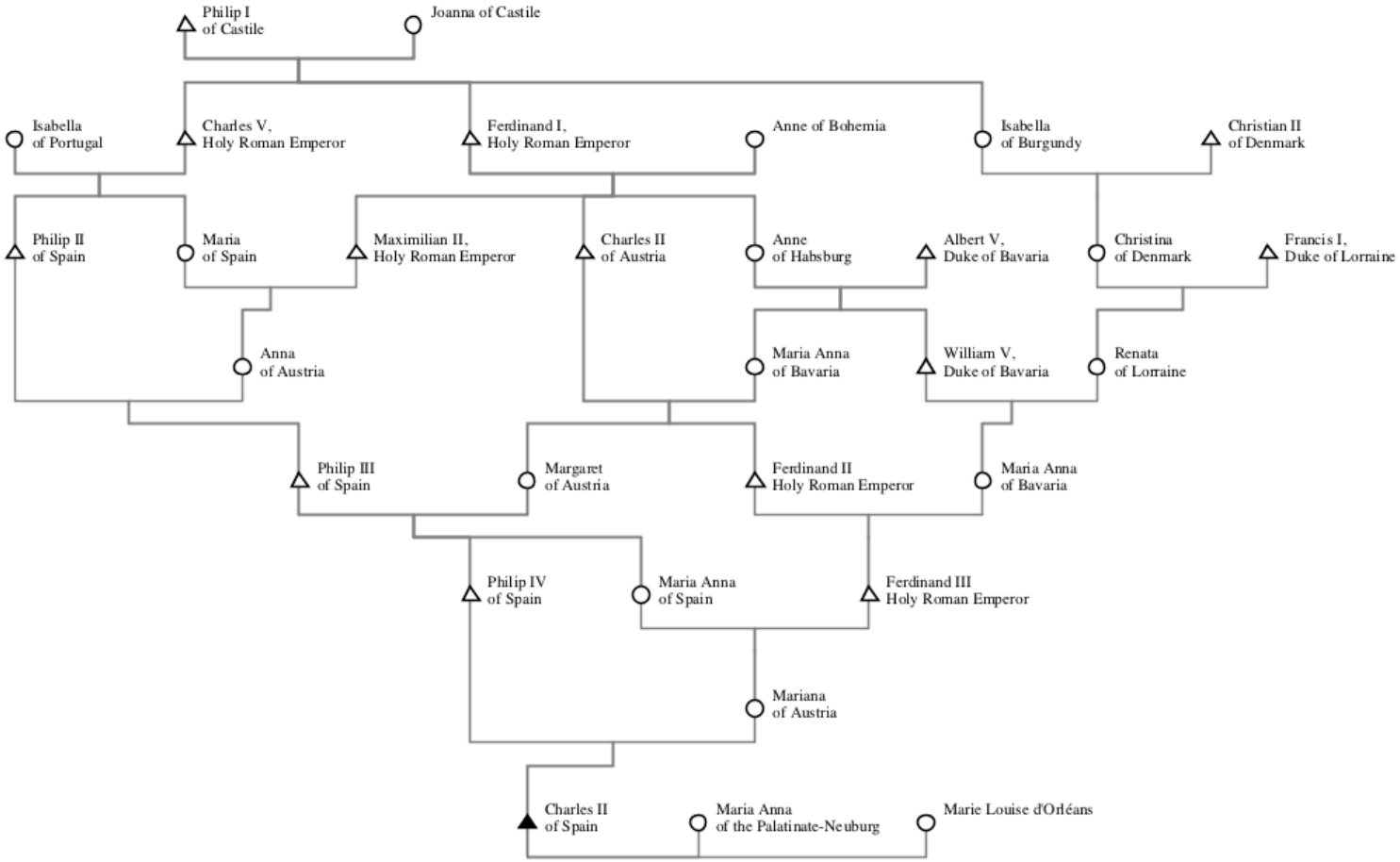
Freeform Diagram Syntax

- By using the <id> parameter, preceding individuals can be referred to later in the kin type string.
- Here id #75 is used as a back reference to the first individual, creating a matrimonial ring.
- f:#75;N75:DDHMDSWMMM:#75:



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

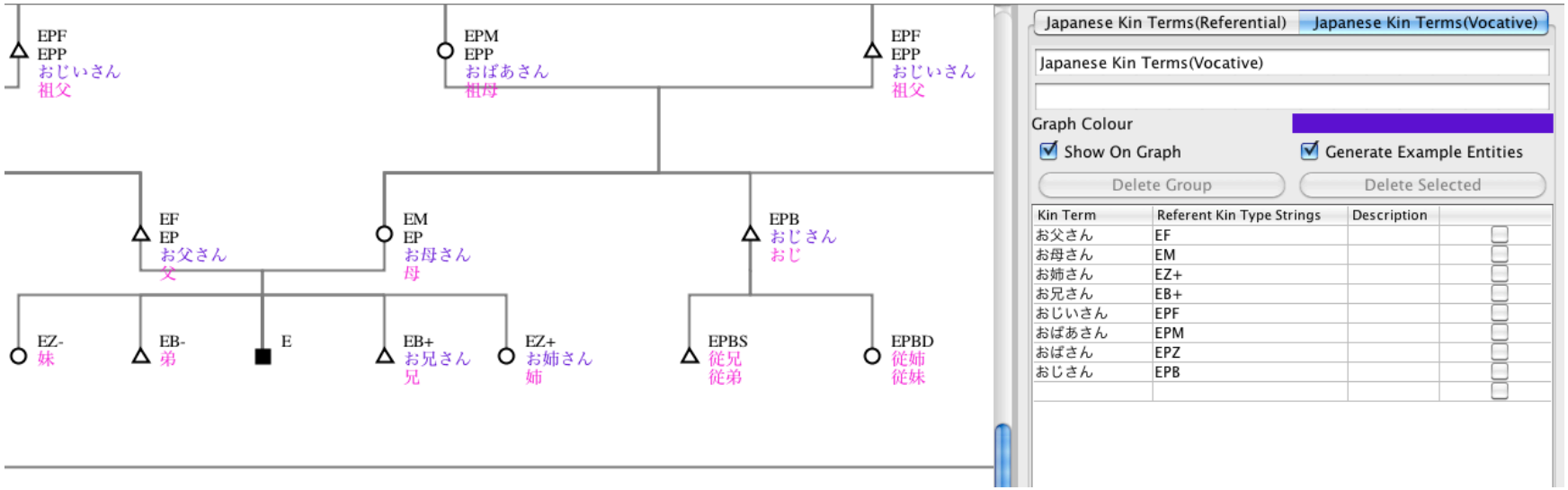
Freeform Example: Charles II of Spain



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Kin Terms

- Kin terms can be entered into a diagram.
- They are defined with kin type strings
- A freeform diagram can be generated from these kin terms.
- They can be imported and exported.
- In the future it will be possible to overlay these kin terms onto a database driven diagram.
- There are a number of example diagrams in the application.



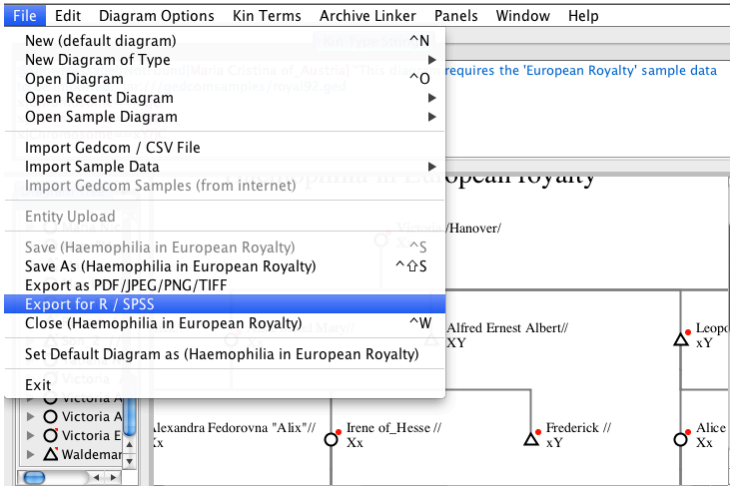
Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

R and SPSS

- The kinship data displayed on a standard diagram can be exported to a text file for use in R or SPSS etc.

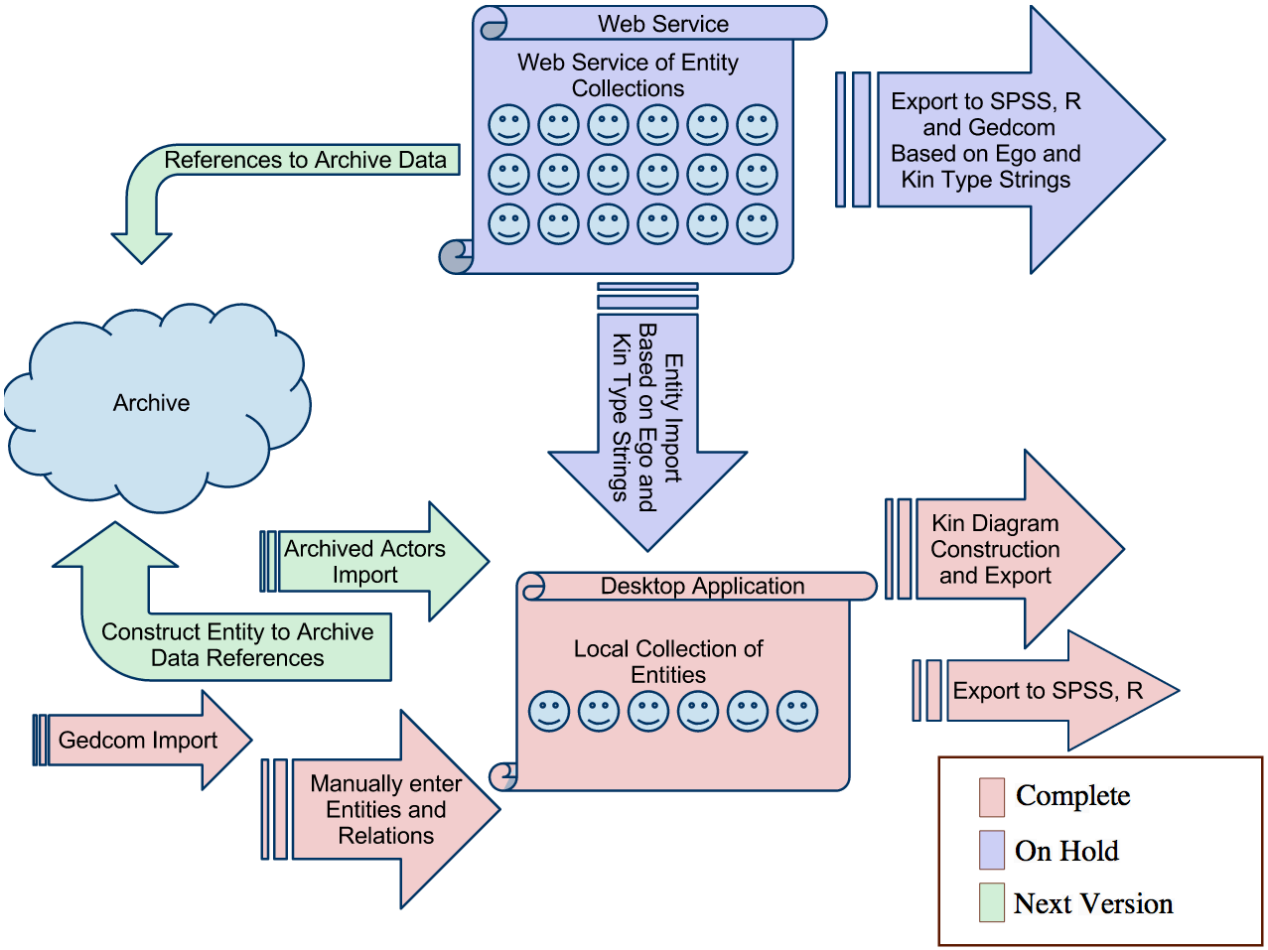
```r
dataFrame <- read.table("/home/peter/Desktop/KinOathDemoR/exportedData.tab",header=T)

# get all affected parent ids
affectedIds = dataFrame$id[dataFrame$symbol_redmarker == 1]
# get all children of affected
childrenOfAffected = dataFrame[dataFrame$momid %in% affectedIds || dataFrame$dadid %in% affectedIds]
# print names of affecteds' children
childrenOfAffected$label_0
```



id	momid	dadid	sex	ego	dob	dod	label_0	symbol_triangle	symbol_blackstrikethrough	symbol_redmarker	symbol_circle	UniqueIdentifier
1	21	42	1	0	1900	1904	Henry //	1	1	1	0	iid_1abe7dabd1c70628f07963e95f9eb0f3
2	37	43	2	0	1846/05/25	1923/06/09	Helena Augusta Victoria//	0	1	0	1	iid_ed611b28b31aaec4c49801aab67955c5
3	17	44	1	0			Son_2 //	1	0	1	0	iid_f165b998134ee9c87fcf7f77afcdcb30
4	21	42	1	0	1889	1945	Waldemar //	1	1	1	0	iid_01af3ba990cdf7c5fb99c99663310fb5
5	37	43	2	0	1857/04/14	1944/10/26	Beatrice Mary Victoria//	0	1	1	1	iid_d573d6fdbe6bc99b0dee7584a73f1c4c
6	29	45	2	0	1897/06	1918/07/18	Tatiana Nicholovna //	0	1	0	1	iid_61f27021af082a3ae385fa409ae578ad
7	14	46	1	0	1870	1873	Frederick //	1	1	1	0	iid_a8a378106e81dd1a9358709ed2140d9a
8	32	47	1	0	1907	1938	Alphonso of_Cavadonga //	1	1	1	0	iid_b79466e3fda68b430682cf359d3e989f
9	37	43	1	0	1841/11/09	1910/05/06	Edward_VII /Wettin/	1	1	0	0	iid_15cfe4ec699efd384eb3892d68164be4
10	37	43	1	0	1850/05/01	1942/01/16	Arthur William Patrick//	1	1	0	0	iid_55158ffef1ef49fb5e03dcc7c97e856b
11	37	43	2	0	1848/03/18	1939/12/03	Louise Caroline Alberta//	0	1	0	1	iid_04c5146f7a9f5602148fe498a2719272
12	32	47	2	0	1911		Maria Christina //	0	0	0	1	iid_38fe23e9e9198f80ec2bea370315527d
13	29	45	1	0	1904/08/12	1918/07/18	Alexis Nicolaievich /Romanov/	1	1	0	0	iid_980067ddf26ec47693a733a733c02694
14	37	43	2	0	1843/04/25	1878/12/14	Alice Maud Mary//	0	1	1	1	iid_67530193a6b8b2c0a1033446c23b55fe
15	14	46	2	0	1874	1878	Mary "May" //	0	1	0	1	iid_7b6fba6d2514f5ea6a2c0133f6481faf
16	5	48	1	0	1889	1922	Leopold //	1	1	1	0	iid_076ea3ded22a12f1dddf4bee9d9f5cd2
17	49	41	2	0	1883	1981/01	Alice of_Athlone //	0	1	1	1	iid_49c88dd77a2264b6aa613021e236eb84
18	5	48	1	0	1886	1960/02/23	Alexander of_Carisbrooke //	1	1	0	0	iid_7b59019be9dc376171a1ebaf6080225a
19	14	46	1	0	1868	1937	Ernest Louis of_Hesse//	1	1	0	0	iid_42b0e40a5d3ed02b29c62a3df8d29482
20	32	47	1	0	1914	1934	Gonzalo //	1	1	1	0	iid_d5699b66a887d30bb18939c592c88d62
21	14	46	2	0	1866	1953	Irene of_Hesse //	0	1	1	1	iid_c7be01fab14a0b39a944cfbc7b1af370
22	21	42	1	0			Child_#3 //	1	0	0	0	iid_2297e359f40a11761ca4a82b4d49f5a0
23	37	43	2	0	1840/11/21	1901/08/05	Victoria Adelaide Mary//	0	1	0	1	iid_cfeb646cc3dff46971e79f0c3b0a853f
24	32	47	1	0	1913/06		Don Juan of_Spain//	1	0	0	0	iid_d0e1f42101710cd32d6ae07940f3169e
25	29	45	2	0	1901/06	1918/07/18	Anastasia Nicholovna /Romanov/	0	1	0	1	iid_76a8b10d50c99d531a226f479ca8c3d0
26	37	43	1	0	1844/08/06	1900/07/30	Alfred Ernest Albert//	1	1	0	0	iid_cefe73aa656fef02c994ac13ec32e453
27	14	46	2	0	1864	1918/07/17	Elizabeth "Ella" //	0	1	0	1	iid_439671a86d8cf7ef7b64603d17c045c6
28	29	45	2	0	1899/05	1918/07/18	Maria Nicholovna /Romanov/	0	1	0	1	iid_ac0255ad17cb3b3f7ceef79a562c98fa
29	14	46	2	0	1872/06/06	1918/07/16	Alexandra Fedorovna "Alix"//	0	1	1	1	iid_2cf36745e277a718028035c9547243d8
30	5	48	1	0	1891	1914	Maurice //	1	0	0	0	iid_8719d4e8c7f5f343aa17007ee4ad4c59
31	14	46	2	0	1863	1950	Victoria Alberta of_Hesse//	0	1	0	1	iid_7b00f979e251465c17088bf0d19a06aa
32	5	48	2	0	1887	1969	Victoria Eugenie "Ena"//	0	1	1	1	iid_6ef08317358fc7605bba50ba947c5fa1
33	32	47	1	0	1908	1975	James //	1	1	0	0	iid_f270191b00089852dc552275213190fb
34	32	47	2	0	1909		Beatrice //	0	0	0	1	iid_92724995a7618cec9cff9c8d177b0b5a
35	29	45	2	0	1895/11	1918/07/18	Olga Nicholovna /Romanov/	0	1	0	1	iid_f236828d50909413df16564a9c44ba10
36	17	44	2	0	1906		May Cambridge //	0	0	0	1	iid_5e6b6b2dbaa03bf4544a255deecec69a
37	50	51	2	0	1819/05/24	1901/01/22	Victoria /Hanover/	0	1	1	1	iid_87b8650d6336994f9f24654fdb214c71
38	49	41	1	0	1884	1954	Charles Edward //	1	1	0	0	iid_7efe2ede2b0df488c46bc7d6f8b11779
39	32	47	1	0			Don Jamie //	1	0	0	0	iid_3dcede17c537c1aa8d74ba0a4badf972
40	17	44	1	0		1928	Rupert //	1	1	1	0	iid_8b561f64886573a2b5164f687dd02a24
41	37	43	1	0	1853/04/07	1884/03/28	Leopold George Duncan//	1	1	1	0	iid_72fd2eb4dded435312589416 1fcfc8dc

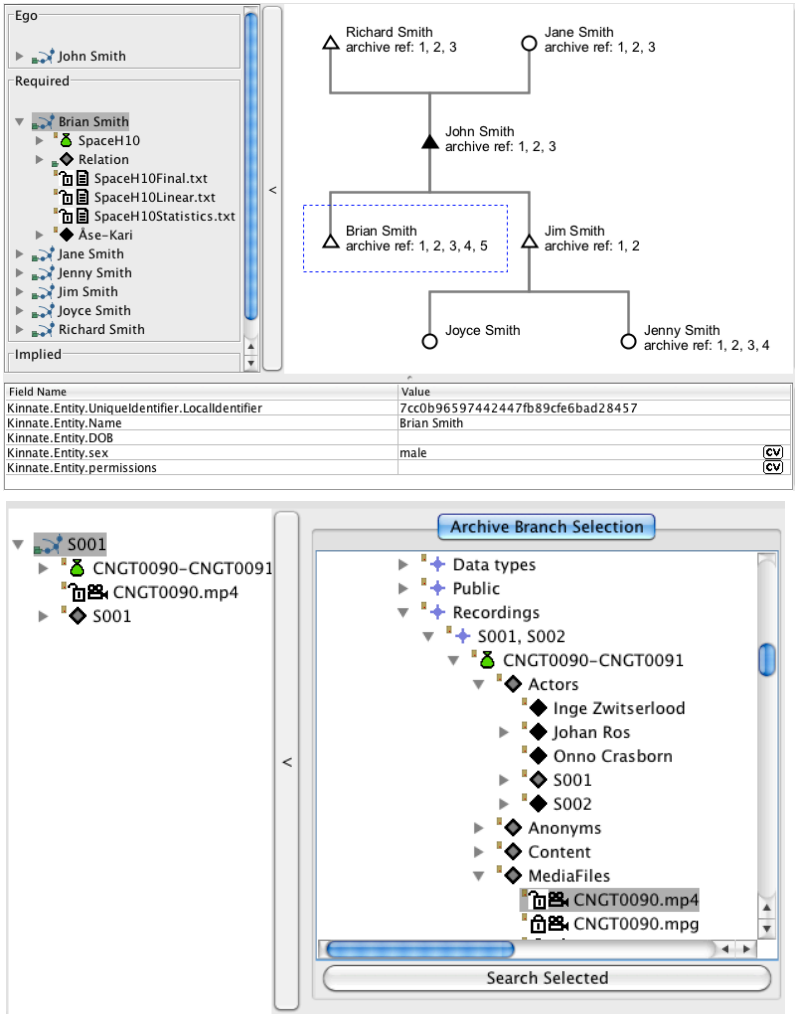
Current Status



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Archive Linking

- This is on hold until a technical agreement is reached on how to persistently link kinship individuals to the archive metadata and data.
- As a result these features will be available in initially as a plugin for the 1.1 version.
- This plugin approach will also allow custom linkers for additional data repositories.
- It is also hoped that we will be allocated time to produce a light weight kinship panel for use in other applications, such as Arbil and ELAN. Such that links can be made and utilised directly in these applications.



Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Plugin Structure in Version 1.1

- The core application is being treated as a thin application.
- On top of this a plugin structure has be added.
- This allows for example:
  - External projects to create modules
  - Custom data sources
  - Specialized importers / exporters
  - Data filters / processors / transformers.
  - A greater variety of graph sorting algorithms, such as sort by clan vs descendancy etc.
- Another benefit of this plugin structure is that new features do not need to have such an impact on the core applications and its stability.

Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen

Conclusion

- The stable version 1.0 of KinOath Kinship Archiver has already been field tested.
- Version 1.1 is currently being tested, it also responds to the initial field testing and will be the new stable version shortly.
- Archive linking will be available initially as a plugin, allowing more flexibility.
- In the current stable you can:
  - Create, view, import, export kinship data.
  - Create, view, import, export kinterms data
  - Query large datasets based on free text and kintypes.
  - Quickly create complex diagrams.
  - Customise the kin data, kinterms and kintypes to suit your needs.
- The plugin structure is now available in 1.1 to further extend the flexibility of the application.

Peter.Withers@mpi.nl The Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen