

REED London Person Entity “Crosswalk” between CWRC and EATS (Developing Proposal)

[REED London challenge:](#)

[Approach:](#)

[“Data slice”:](#)

[Establishing entities in CWRC:](#)

[Workflow:](#)

[REED Entity Template:](#)

[Tagging Entities in CWRC-Writer:](#)

[Building Crosswalk Between CWRC and EATS:](#)

[Developing EATS Connection to REED London Entities:](#)

[Expansion to Other Entities, Review of Entity Types:](#)

REED London challenge:

how do we create “REED authority” in CWRC for over 2000+ people referred to in Inns of Court ~ with the goal that we are making those people discoverable on the semantic web and not duplicating entities that already exist (with EATS, as well as established authorities including OCLC/VIAF, Wikidata, SNAC, ODNB, etc.)? Further, how do we ensure that REED London and REED Online entities remain in sync so that we do not duplicate efforts?¹

Approach:

REED London team creates an authority list within CWRC/CWRC-Writer that prioritizes REED entities while avoiding duplication of entities beyond the REED-iverse. That list is used when tagging entities in CWRC-Writer and tracks with EATS IDs via shared spreadsheet (see ‘Building Crosswalk..’ below).

¹ Note that this proposal concentrates on person entities. People and organizations might be handled in a similar fashion (see ‘Expansion ..’ below).

“Data slice”:

In 2018 REED London is focusing on people referenced in *Inns of Court* records as follows:

- Gray’s Inn records
- Middle Temple records
- 1593-1612 (Shakespeare’s floruit)
- 1612-14 (masque seasons)

This data slice includes ~800 disambiguated people (disambiguated in the sense that they are identified with at least a distinctive last name and floruit date).

Establishing entities in CWRC:

Adapting the TEI-XML template developed by Orlando (see screenshot below, with apologies for resolution) we will create a TEI-XML file for each person, with the following guidelines:

- 1) For people who who already have a OCLC/VIAF, Wikidata, ODNB DOI, etc. URI, we will supply ****only**** last name, first name, and idno with sameAs attribute, as demonstrated for Robert Ker:

```
<idno type="VIAF" cert="high" sameAs="https://viaf.org/viaf/22030889/" />
```

Note that Ker also has ODNB DOI and Wikidata URIs, and so we can also add sameAs refs for those if it is helpful (actually, Wikidata appears to aggregate the others, so including that would probably be in our best interest). We need no other information about this person - their birth/death dates, relatives, etc. - because those have already been established.

- 2) For people who already have an EATS URI as well as one of the above, we will supply ****only**** last name, first name, and sameAs attribute, as demonstrated for Elizabeth I:

```
<idno type="VIAF" cert="high" sameAs="http://viaf.org/viaf/97107753/" />
```

```
<idno type="REED" cert="high" sameAs="https://ereed.library.utoronto.ca/eats/entity/11772/" />
```

Again, no other information is needed, and now we’re creating a link between EATS, VIAF, and CWRC to ensure that we are not establishing multiple Elizabeths (CWRC already uses QE1’s VIAF ID, so we are not creating another one ... just adding her to the “list” of available REED London entities for tagging in CWRC-Writer. - see “Find a Person ..’ screenshot below)

- a) For the 17 other people who already have an EATS ID **and** appear in *Inns of Court*, all have an existing VIAF or Wikidata ID, so we will continue this process throughout.
- 3) For people who appear in *Inns of Court* and are also referenced in *REED: Patrons & Performances* but do not have an EATS ID, we will create an entity and use REED as the bibliographical source, as demonstrated for Edward Courtney:

```
<ref source="REED" target="https://reed.library.utoronto.ca/node/315995" />
```

Note that as far as we can tell people who appear in *Inns of Court* and have a P&P ref

also have a published URI with VIAF or Wikidata, so we would also include that as sameAs, as above.

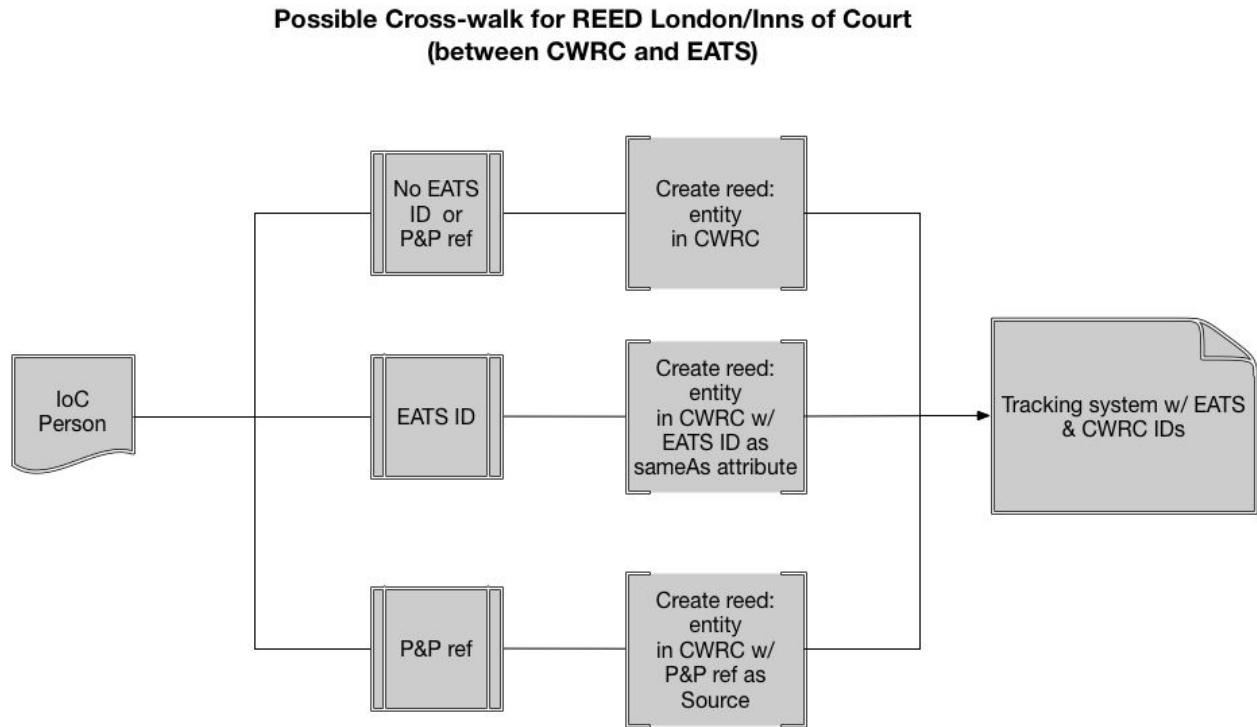
It might be helpful to establish a sameAs when P&P is ingested into EATS so as not to create duplicates of already established entities. If helpful this could be research that a REED London research assistant at Bucknell might undertake in summer/fall 2018. That might be a better use of Kathy's time - the same will be done with the Inns of Court people who are still being vetted - a better use of Kim's time, as well!

- 4) Where there is no discoverable authority name CWRC will establish the authority for REED London. The very minimum of information we need for this is a last name and floruit date. For example: *Francis East, Gray's Inn (1595)*; or even ... *Fitzherbert (1598)* Obviously we'll want to get more information than this, and in most cases we'll have an affiliation (an Inn) and/or an event and/or an occupation and/or a family relationship. If we discover after the fact that there is already an authority, we can edit the entity from within CWRC-Writer "edit entity" function, adding "sameAs" attribute as above).

Because these CWRC-created entities will all have a published URI (HTTP protocol) used as a Public Subject Indicator, we can maintain them as active links in a Google spreadsheet shared with REED staff. CWRC will host the REED London entities with the same commitment they have made to Orlando - meaning that these can be considered "permanent" "stable" URIs - however REED London continues to be developed. That permanence is important because we are using these to make the entities discoverable to other projects, (e.g., Six Degrees of Francis Bacon, EMLO, etc.).

Workflow:

Here is a rudimentary workflow visualization showing how we might address person entity creation and aggregation (with apologies for my lack of diagram-making skills):



REED Entity Template:

Here is a (bad resolution) screenshot of the TEI-XML template for person entities.

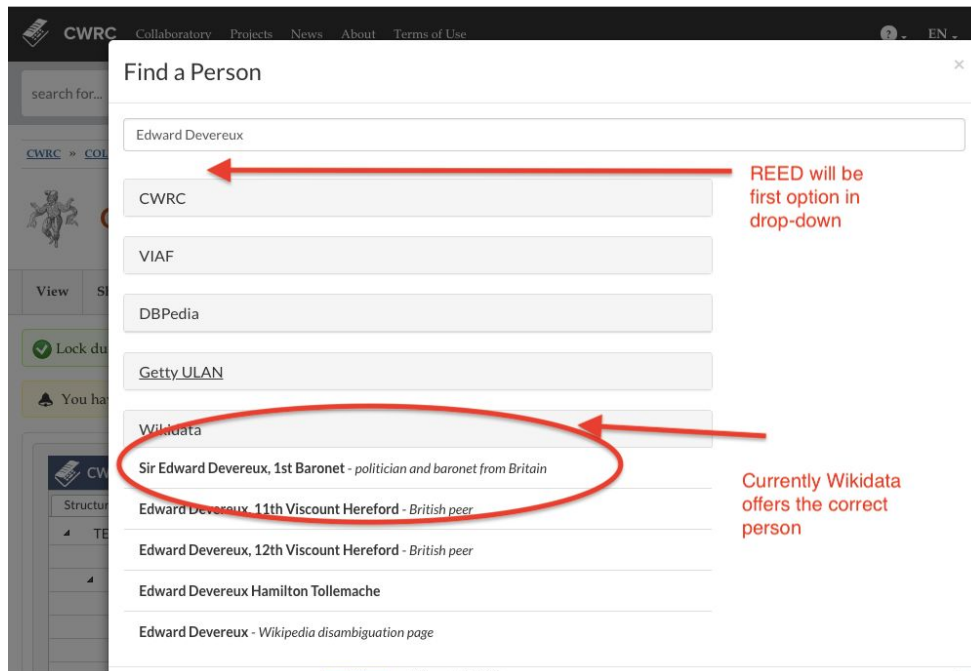
```
<listPerson>
  <person xsl:id="PID_entity_record" sex="f">
    <idno type="Authority_name" cert="high">VIAF URI</idno>
    <persName type="standard">
      <name>Last Name, First Name </name>
    </persName>
    <persName type="preferred" xsl:lang="ENG">
      <name type="forename">Forename</name>
      <name type="middlename">Middle Name</name>
      <name type="surname">Last Name</name>
      <genName>(e.g. the IIrd, senior, junior, etc.)</genName>
      <roleName>official title or rank</roleName>
    </persName>
    <persName type="variant" role="variant_name_type" xsl:lang="ENG">
      <name type="forename">Forename</name>
      <name type="middlename">Middle Name</name>
      <name type="surname">Last Name 2</name>
      <genName>(e.g. the IIrd, senior, junior, etc.)</genName>
      <roleName>official title or rank</roleName>
    </persName>
    <note>
      <label>Named preferred by: </label>
      <desc>
        <orgName ref="URI">Name of the project that uses the variant name</orgName>
      </desc>
    </note>
    <persName>
    <birth>
      <date cert="high" source="https://www.w3.org/TR/NOTE-datetime" when="2000-10-30">D.O.B </date>, <placeName>Place of birth</placeName>
      <note>
      </note>
    </birth>
    <death>
      <date cert="high" notBefore="2017-07-21">?
        <!--format for living persons; for dead persons, structure the same as for D.O.B.--></date>,
        <placeName>Place of death</placeName>
      <note>
      </note>
    </death>
    <floruit>
      <date>
      </date>
    </floruit>
    <sex source="http://sparql.ciirc.ca/ontology/ciirc.html#Gender" value="GenderQueer" cert="high"/>
    <trait type="factuality" cert="high">
      <ab>real/fictional
      <!--form should contain a dropdown with three options: real, fictional, both, unknown--></ab>
    </trait>
    <nationality cert="high" ref="http://sparql.ciirc.ca/ontology/ciirc.html">Nationality
      (zero or more)</nationality>
    <occupation cert="high">Occupation (zero or more)Use Orlando picklist
      values</occupation>
    <affiliation cert="high">User input</affiliation>
    <note type="general" xsl:lang="ENG"> This is a general note about the person. </note>
    <note type="project-specific" xsl:lang="ENG">
      <note>This is a project-specific note about the person.</note>
      <respons locus="value">
        <desc>
          <orgName ref="URI">Project (e.g.OWRC, REED, Orlando)</orgName>
        </desc>
      </respons>
    </note>
  </person>

```

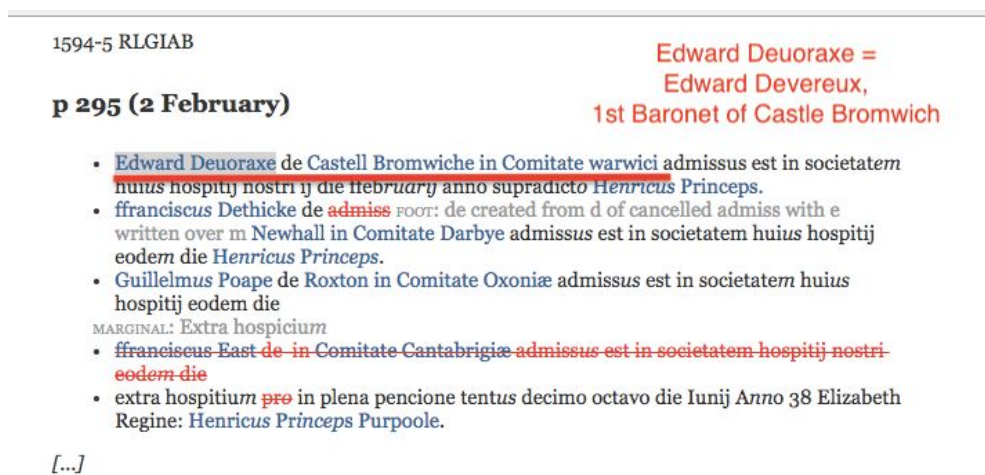
We can include as much or as little information as we have available, as long as we have a last name and floruit date. We can also include variant spellings as listed in the *Inns of Court* index.

Tagging Entities in CWRC-Writer:

Here are suggestions about how CWRC would allow an editor to associate a name in a text with the correct entity (REED would be first option, then VIAF, then DBPedia, etc.



Here is how the editor would see and tag the entity from within CWRC-Writer (alt spellings for names will be included wherever possible so that we can minimize confusion due to non-standard spelling.) If a person is listed as “REED” that would be the first option in the drop-down. If a person has not been added to REED, the choices for searching would continue (VIAF, Wikidata, etc.)



Here is how the editor would “see” the person in a record. The entity tag for variant “Edward Deuoraxe” would point at Edward Devereux. Additionally, if the editor wanted to tag “he” or “Sir Edward” or “the Baronet”, they could point that to the entity ~ it appears that this is how EATS manages tagging, too.

Building Crosswalk Between CWRC and EATS:

Using a shared Google spreadsheet, we can document and track entity creation and aggregation. For example:

	A	B	C	D	E	F	G	H
1	Standard Name	CWRC/REED London ID	sameAs	EXISTING AUTHORITY	sameAs	EATS ID	Source	Patrons & Performances
2	Elizabeth I, Queen of England	https://commons.cwrc.ca/cwrc:d713a-eeb-e817-40cc-9b3a-b4cd63228c31	Y	http://viaf.org/viaf/97107753	Y	https://ereed.library.utoronto.ca/eats/entity/11772		
3	Ker, Robert	https://commons.cwrc.ca/reed:##	Y	https://viaf.org/viaf/22030889	n/a			
4	Courtney, Edward	https://commons.cwrc.ca/reed:##	n/a		n/a		Y	https://reed.library.utoronto.ca/node/315995
5	East, Francis	https://commons.cwrc.ca/reed:##	n/a		n/a		n/a	
6	Fitzherbert, ...	https://commons.cwrc.ca/reed:##	n/a		n/a		n/a	
7								

*Elizabeth I is already identified in CWRC using a cwrc: ID with sameAs attribute with VIAF. She also has an EATS ID, so all three will now be “sameAs”

*Robert Ker has a VIAF ID, so we will create a reed: ID with sameAs attribute.

*Edward Courtney will have a reed: ID and Source reference to Patrons & Performances

*Francis East and ... Fitzherbert will both have reed: ID.


If, for example, EATS should discover Robert Ker or Francis East in Hampshire - or EATS-ify Edward Courtney, it would make sense for an EATS ID with sameAs connection with reed: to be created in EATS. It would then be a maintenance task to make sure that sameAs attributes are maintained going forward.

Developing EATS Connection to REED London Entities:

A review of the EATS interface suggests that a field could be added for a sameAs attribute - perhaps in a way similar to the current bibliographic citation field (usually used for the ODNB URL). It also appears that EATS fields can be 'added', (for additional names, citations, etc.) so if there were a way that the field could be used for VIAF IDs, then REED Online people would be discoverable to other existing projects. See this "stub" in CWRC for Elizabeth I:

Elizabeth I, Queen of England

View	Edit	Share	Lock	Manage
-------------	-------------	--------------	-------------	---------------

 You have 714 unstamped log entries. Please [review these here](#).

Information

ABOUT

ACCESS CONDITIONS
Use of this public-domain resource is governed by the [Creative Commons Attribution-NonCommercial 4.0 International License](#). (use and reproduction)

CONTRIBUTING PROJECT(S)

- VIAF

IDENTITY

PREFERRED NAME(S)

- Elizabeth I, Queen of England

SAME AS

- <http://viaf.org/viaf/214146998503118941898/>

Expansion to Other Entities, Review of Entity Types:

Currently there are no London places or organizations that have EATS IDs (e.g., the Banqueting House, Gray's Inn, etc.). These place and organization entities will be created in CWRC in a similar fashion as people, and the URIs shared with REED staff for tracking purposes.

The entity types and relationships used in EATS map well against TEI, so with minor translations from EATS to TEI those should not be an issue. In cases where new entity types (e.g., Thing_animal, or the like) are referred to in *Inns of Court* or other REED London collections/editions, then a proposal/review system would be established through REED.