

The Historical Land Records of Basel - Ground Truth

Documentation of the Annotation

Ismail Prada Ziegler

August 18, 2025 (v0.1)

Contents

| | |
|---|-----------|
| Contents | I |
| 1 Introduction | 1 |
| 1.1 Background | 1 |
| 1.2 Annotation Principles | 1 |
| 1.3 Related Annotation Guidelines | 1 |
| 2 Entity Annotation | 2 |
| 2.1 Basic Principles | 2 |
| 2.2 Elements | 2 |
| 2.3 Entity Classes | 4 |
| 2.4 Description Classes | 4 |
| 3 Value Annotation | 8 |
| 3.1 Basic Principles | 8 |
| 3.2 Classes | 8 |
| 4 Relation and Event Annotation | 9 |
| 4.1 Basic Principles | 9 |
| 4.2 Elements | 9 |
| 4.3 Relationship Classes | 11 |
| 4.4 Event Classes | 15 |
| 5 Data Formats | 25 |
| 5.1 Introduction | 25 |
| 5.2 Simple TEI | 25 |
| 5.3 Nested TEI | 26 |
| 5.4 BeNASch Standard XML | 28 |
| List of Tables | 29 |
| Acronyms | 30 |

1 Introduction

1.1 Background

This documentation is meant to explain the annotations found within the Historical Land Records dataset. The Historical Land Records Ground Truth (HLR-GT) consists of 829 documents which were enriched with semantic annotation. The data set can be used as a base for machine learning and evaluation of automated annotation systems. It was created as part of the project *Economies of Space. Practices, Discourses and Actors on the Basel Real Estate Market (1400-1700)*. Please, make sure to cite our project if you use the data set.

1.2 Annotation Principles

The annotation of the Historical Land Records aims to be close to the text with little interpretation. As the goal of the framing project was to extract events from the documents, the annotation focuses to annotate entities which are actors and objects in these events as well as information such as dates, times and monetary values. To grasp the full content of the documents, the annotation could not be restricted to a flat layer, but instead features nested entities.

1.3 Related Annotation Guidelines

The Historical Land Record Ground Truth was the prototypical corpus that the *Berner (früh-)neuhochdeutsches Annotationsschema (BeNASch)* was based on. Some changes later added to BeNASch could not be applied to the HLR-GT and so it differs in some aspects. This, as well as the fact that the BeNASch Guidelines are currently not translated into english, requires this project-specific documentation. BeNASch itself was inspired by the ACE2005 Guidelines. Due to the nested nature of BeNASch, the annotations are easily reducable to use the HLR-GT data together with data with simpler annotation, such as the more commonly seen MUC guidelines.

2 Entity Annotation

2.1 Basic Principles

In the HLR-GT, all entity mentions of the specific entity classes (Person, Location, Organization, Facility, Geopolitical Entity) are annotated, even if they are only mentioned by a pronoun. The whole noun phrase that refers to an entity is what we call a **mention** or **reference**. Every reference features one (or very rarely multiple) **head** elements. The head is one of the *describing* elements inside a reference. We feature two more such elements which are **appositions** and **attributes**. Other entities might also be mentioned inside the original reference. This will usually point to some kind of relationship, which is further specified by a describing element. See chapter 4 for more information on how we handled relationships and events.

2.2 Elements

2.2.1 Reference

A reference is a noun phrase that refers to one or more entities. They must contain at least one head and can contain appositions and attributes. The most important classification is their entity class. A full list of entity classes can be found in section 2.3. Besides the entity class, annotations were also categorized by their numerus (singular entity or a group) and their specificity (referring to specific individuals or not).

| reference | entity class |
|-------------------------------|--------------|
| John the taylor | person |
| the house close to the church | facility |
| St. Augustin monastery | organization |

Table 2.1: Example References.

2.2.2 Head

The head is the central element of a reference and is usually the same as the grammatical head. A head may contain no other elements within. Each head features a class which is part of the list of *description classes* in section 2.4.

For organizations, determining which parts of a name are part of the head and which are part of an attribute or apposition was often difficult. We aimed for consistency inside the organizations themselves. When in doubt if a part of the reference was part of the head, we would decide to include it in the head.

| reference | head class |
|--------------------------------------|------------|
| John the taylor | name |
| the house close to the church | type |
| St. Augustin monastery | name |

Table 2.2: Example References with Head elements in bold.

2.2.3 Apposition

Appositions share most characteristics with references but can never stand independently. They are also noun phrases and require an annotated head. They can also be further nested and contain other appositions and attributes. In the HLR-GT, appositions appear most frequently when describing persons more closely, such as their occupation or family relationships. Each apposition features a class which is part of the list of *description classes* in section 2.4.

| reference | apposition class |
|--|-----------------------------|
| John the taylor | occupation |
| his wife Anna | familial relationship |
| Hieronymus Bauer citizen of Basel | affiliation to organization |

Table 2.3: Example References with Appositions in bold, head of the apposition underlined.

There is one use of apposition in the HLR-GT which does not adhere to these guidelines, that is the mention of house names. Often we find this structure "the house, called Golden Star, located at ...". Due to some early decision on how to annotate the name of the house, the house names are annotated as appositions like this: "the house, called **Golden Star**, located at ...". An advantage of this inconsistency is that house names can be found by simply searching for all heads of the "name" class.

2.2.4 Attribute

Attributes annotate any spans which further describe the entity that the mention refers to which are not appositions. Attributes can come in many forms, but most commonly appear as relative clauses, genitive attributes and possessive articles. Attributes can contain references, but no describing elements. Each attribute features a class which is part of the list of *description classes* in section 2.4.

| reference | attribute class |
|--|-----------------|
| John the taylor who lives in minor Basel | location |
| the house, pays dues to the church St. Lienhard 5 lb. | dues |
| his¹ house | ownership |

Table 2.4: Example References with Attributes in bold.

2.2.5 Lists

A special type of references are lists. Lists mark references that are mentioned together and are joined by conjunctions or commata. Lists don't contain heads, but can contain appositions or attributes, which describe all references that are part of the list. Lists feature one or multiple entity classes which are simply an aggregate of the list members.

| list | entity class |
|---|----------------------|
| John Smith and his wife Anna <u>who live in minor Basel</u> | person |
| <u>the creditors</u> the church St. Lienhard and Sir Ulrich Roten | person, organization |

Table 2.5: Example Lists with references in bold and describing elements underlined.

As lists were introduced to the annotation system later on in the process to deal with attributes referring to multiple reference, they were not annotated in a very strict fashion. While we applied additional effort to rework older files, they might still be missing in some cases where they would normally be applicable.

2.3 Entity Classes

All references feature an entity class. These refer to the real-world entity that the reference mentions. In some cases, we also added subclasses, but please note that these were mostly not annotated in a strict fashion. We keep them included in the HLR-GT, but they should be used with care. An overview of the all entity classes with short descriptions is given in table 2.6.

2.4 Description Classes

All describing elements feature a description class. These classes describe *how* an entity is mentioned in the text. A few of these also feature subclasses, which were not strictly annotated. Note that while some classes correlate more strongly with certain elements or entity classes, we use a shared pool of classes for all describing elements. In one case "ownership" the meaning of the class changes depending on the entity that is being described. Due to this, any research interested in relations should use the relations and events described in chapter 4, where the directionality of these relations is explicit. See table 2.7 for a full overview of the description classes.

¹Please note that there is also a reference to the owner nested within the attribute.

| class | description | subclasses |
|-----------------------------|---|---|
| Person | Any reference to a person. | - |
| Facility | Any reference to a human built structure such as houses, streets, fountains. | Street, Religious, District, Public Square, Gatehouse, Defense, Occupational, Fountain, Partial, Bridge, Governmental |
| Organization | Any reference to an organization or institution that is not at the same time a geopolitical entity. Distinguished from person groups by some form of constant or formal membership. | Religious, Governmental, Occupational, (Noble) Family, Academic |
| Geo-Political Entity | Any reference to an entity which can be referenced as both a political actor as well as a location. Is split into three classes in the corpus to distinguish mention as a political entity, geographic entity or general mention. | - |
| Location | Any reference to locations which are neither geo-political nor human made entities (facilities). | River, Lake, Tree |
| Unknown | Any reference where the entity type can not be determined for some reason (e.g. HTR errors). | - |

Table 2.6: Overview of all entity classes and descriptions in order of frequency in the HLR-GT.

Table 2.7: Overview of entity description classes.

| class | description |
|------------------------------------|--|
| Name of the Entity (<i>name</i>) | A personal name or named reference. |
| Type | A closer description of the entity class. E.g. "house" in a facility reference, "monastery" in a organization reference or "city" in a geo-political reference. |
| Location | A spatial description indicating where the entity is. |
| Pronoun | A pronoun used to refer to the entity (e.g., he, she, they). |
| Self-Mention | A pronoun mention where the author referred to themselves (e.g. "we sold the house"). Subclass of <i>Pronoun</i> . |
| Occupation | Description of the entity by its profession or trade. |
| Work for an Organization | As <i>Occupation</i> , but implying work for a specific organization. Subclass of <i>Occupation</i> . |
| Representation of an Organization | Describes the entity working in a function where it has some kind of representative status in an organization, such as the secular administrator of a monastery after the reformation ("Schaffner"). Subclass of <i>Occupation</i> . |
| Ownership | The entity is either referenced owning something (common for persons or organizations) or being owned by someone (facilities). |
| Familial Relationship | The entity is referenced by a familial relationship such as parent-child, husband-wife or being siblings. |
| Is Dead | The descriptor indicates that the entity is dead in the time-context of the document. |
| Dues | Describes the financial obligations any owner of the referenced facility must fulfill, usually to a landlord or a creditor. |
| Title | A title held by the entity (e.g., Lord, Master, Doctor). |
| Academic Title | Subclass of <i>Title</i> to mark academic titles specifically. |
| Includes | Describes parts that the entity is made of. In case of a group-reference describes the members of the group. |
| Is Part Of | Describes the entity as a component of another entity. |
| Organizational Affiliation | Describes the entity as member or in association with an organization. |
| Has Member | Describes the members of an organizational entity. |

Continued on next page

| class | description |
|----------------------|--|
| Heir | Describes the entity as the heir to something or of someone. |
| Role in Event | Describes the entity by their function in a mentioned event, such as being a seller in a purchase event. |
| Mention by Numbering | Describes a group of entities by their count, e.g. "the two". |
| Representative | Describes the entity as a representative of another entity. |
| Generic Relationship | Vague or general relationships between entities, e.g. neighbors. |
| Is Owed Money | Describes the entity as a creditor or to be paid by others. |
| Other | Miscellaneous descriptors which don't fit any other class. |
| Is Sick | Describes the entity as ill or suffering from a condition. |
| Was Sold | Describes a sale of the entity. |
| Inherited From | Describes a transfer of the entity by inheritance. |
| Unknown | The nature of the description is unclear (e.g. due to HTR errors). |

3 Value Annotation

3.1 Basic Principles

Values mark quantifiable information such as dates, time and monetary values. Values never contain other entities or values. Values can be annotated in references, appositions or attributes, e.g. a monetary value when the dues owed to someone are described. While BeNASch features more complex values, the HLR-GT uses only a simpler subset of them.

3.2 Classes

3.2.1 Money

While this value class is called money, it encompasses other bartering goods as well. For example, payment of the dues "zur Weisung" were often required in the form of bread or chicken.

3.2.2 Date

The date class marks specific points in time, such as a specific day or year.

3.2.3 Time

While the time class can be used for any finite or quantifiable timespans, e.g. "five days". The only relevant use of it is in the subclass *time_recurring* (e.g. "yearly", "monthly") to denote the yearly dues.

4 Relation and Event Annotation

4.1 Basic Principles

Relationships and events work in the same fashion in BeNASch, both feature one or multiple actors or objects. While relationships describe a state of being between two or more entities or values, events describe an interaction (or simply an action) in a specific (even if unknown) moment in time. Relationships and events are connected through the previously mentioned annotations by their anchor, trigger and roles.

4.2 Elements

4.2.1 Classification and Anchor

Every event and relationship features a main class to describe the type of action or interaction. All relationship and event classes used in the HLR-GT are listed in sections 4.3 and 4.4. Besides the main class, each event and relationship features classification by polarity, tense and modality. In some cases, we classified specific changes in modality as a new event class, such being the case for "testament" and "inheritance", where "testament" could have been annotated simply as "inheritance" with an "intent" modality class. Furthermore every event and relationship is anchored in an annotation layer. For example, a dues attribute may be the anchor to a dues relationship. All relevant elements to the event usually share the same annotation layer (e.g. all inside the same attribute), with the exception of the entity being described by the layer, which usually also takes part in the event or relationship.¹ The anchor may also be the document level. For each event and relationship, a start and end position is also given. Note that these were only in rare cases annotated explicitly and instead inferred from the relevant elements partaking in the event.

4.2.2 Roles

The minimum requirement for any relationship or event is the annotation of at least one role. A role is usually linked to an entity reference or a value, but sometimes can mark a span of text independent of entities and values. Entities in roles can be seen as participants in the event, while the role designates *how* the entities are taking part in it.

¹Two times "usually" because there were some rare cases where this restriction could not be adhered to without loss of information.

4.2.3 Trigger

Some events and relationships feature a trigger (no more than one). If the event or relationship is described inside a reference or apposition, the head element is automatically the trigger.² In other cases, the triggers were only annotated for designated events. A trigger is the central signifier to an event and usually the head of the syntactic structure of the linguistic element that the event is mentioned in. In case of the HLR-GT, the trigger annotation was not always done with complete consistency, which should be considered when working with the triggers. If multiple events share the same trigger, they are marked as an event group. Event groups often also share other features such as roles or modifiers.

4.2.4 Example

"Ulrich Steiner the taylor purchases from Anna, Hans Schaads rip. widow, the house located at the Eisengasse, because she inherited that house recently, for 35 fl."

In this example, we'll focus on some specific events in the example document text above to further the readers understanding of events.

Simple Relationship The relation between "Anna" and "Hans Schaad" is a perfect example of a simple relation with two actors and one easily recognizable trigger.

- class: Familial Relationship
- anchor: "Hans Schaads rip. widow"
- trigger: "widow"
- roles:
 - Family Member A: "Anna, Hans Schaads rip. widow"
 - Family Member B: "Hans Schaads rip."

Complex Event The main event of the document, the purchase of the property is a bit longer:

- class: Property Purchase
- anchor: document level
- trigger: "purchases"
- roles:
 - Buyer: "Ulrich Steiner the taylor"
 - Seller: "Anna, Hans Schaads rip. widow"

²For example in the reference "his wife Anna" a familial relation between "his" and "Anna" is present, anchored to the apposition "his wife". "wife" is the trigger to this relation.

- Property: "the house located at the Eisengasse"
- Cause: "she inherited that house recently" (this role would refer to the inheritance event!)
- Price: "35 fl."

4.3 Relationship Classes

Disclaimer: The descriptions of states and roles in this section were written by an AI system, based on previous German-language documentation, and were proof-read and corrected by a human.

civic-affiliation

The *civic-affiliation* relationship denotes a person's formal connection to a town, city, or other geopolitical entity, e.g. citizenship ("burger") or residency ("hindersass").

- **geopolitical-entity**: The town or administrative unit to which the person is affiliated.
- **resident**: The individual holding the civic status or affiliation.

debt

The *debt* relationship reflects an outstanding financial obligation between two parties.

- **capital**: The principal amount owed.
- **creditor**: The party to whom the debt is owed.
- **debtor**: The party responsible for repaying the debt.

Other roles: related-property, accumulated-interest, capital-detail

due-obligation

The *due-obligation* relationship captures a standing financial obligation, usually tied primarily to a property.

- **interest**: The financial value that is due periodically.

- **property**: The property which owes the due.
- **payer**: The individual or party responsible for making the payment.
- **beneficiary**: The recipient entitled to receive the payment.
- **interval**: The time interval in which the interest is due (almost always yearly).
- **due-date**: The specific date on which the payment is expected each interval.
- **previous-due**: A reference to an earlier payment obligation, sometimes occurring when the paying party changes.
- **previous-payer**: Added information that someone else paid the interest before the current payer.
- **capital**: The principal amount associated with the obligation, usually given when the due is repayable.
- **interest-redeemable**: Information that or how the interest can be repaid.
- **consenting**: An agent providing approval related to the obligation, usually when a modification occurs.

Other roles: interest-conv, interest-detail, interval-conv, other, payment-mode, distribution, effective-payer, alternative-payer, proxy, confirmation, prev-interest, quant, hereditary-owner, part-redeemable, property-detail, previous-property, date-due-detail, alternative-interest

employment

The *employment* relationship marks the relationship between a person or group offering labor and another offering compensation in exchange.

- **employee**: The agent performing work or service under the employment agreement.
- **employer**: The agent providing compensation and overseeing the work.

family

The *family* relationship denotes a familial relationship between two or more individuals.

- **family-a**: One person involved in the family relationship.
- **family-b**: Another person involved in the family relationship.

heirship

The *heirship* relationship describes one entity as being the heir of another.

- **heir**: The individual or party inheriting from someone.
- **decedent**: The deceased individual.
- **property**: A property being transferred through inheritance.

Other roles: property-detail

holds-landed-title

The *holds-landed-title* relationship describes a situation where an individual or group holds an official or territorial title associated with land or governance.

- **title**: The landed or organizational title being held.
- **holder**: The individual or group in possession of the title.

membership

The *membership* relationship indicates that an agent is a member of an organization.

- **organization**: The group, institution, or body to which the member belongs.
- **member**: The individual or agent who holds membership in the organization.

other-interpersonal-relation

The *other-interpersonal-relation* relationship describes a connection between two persons that does not fall into more specific categories such as family or employment.

- **person-a**: One of the individuals involved in the relation.
- **person-b**: The other individual involved in the relation.

ownership

The *ownership* relationship describes the legal or factual possession of a property by an individual or entity, optionally distinguishing an owner from a tenant.

- **owner:** The individual or party who holds ownership of the property. Unless a tenant role is present this does not distinguish between an entity and simply annotates any entity which is named in a possessive sense in regards to the property. Usually this will be a tenant, not the landowner.
- **property:** The item or real estate held under ownership.
- **tenant:** The individual or party using or leasing the property.

part-whole

The *part-whole* relationship most often designates one structure being part of a larger structure (e.g. a garden as part of a property), but can also refer to compositional relationships, e.g. an agency being part of a government.

- **part:** The constituent element or component.
- **whole:** The encompassing or composite entity.

representation

The *representation* relationship indicates that one individual or party acts on behalf of another in legal, economic, or administrative matters.

- **represented:** The individual or party being represented.
- **representative:** The individual or party acting on behalf of the other.

topological

The *topological* relationship represents a spatial relationship between two locations or places. Due to our projects research goals, topological relationships only serve to mark text which holds text with that kind of semantic value, but does not further classify what kind of topological relationship exists.

- **location-a**: The first location in the spatial relation.
- **location-b**: The second location in the spatial relation.

workplace

The *workplace* relationship represents a connection between a worker and their place of work.

- **worker**: The person engaged in work.
- **location**: The place where the work is carried out.

4.4 Event Classes

Disclaimer: The descriptions of events and roles in this section were written by an AI system, based on previous German-language documentation, and were proof-read and corrected by a human.

Generic roles for all events

- **date**: The date(s) when the event occurred.
- **place**: The location where the event occurred.
- **cause**: A cause that lead to the event.
- **consequence**: A consequence of the event.
- **other**: Any additional contextual detail.

bequest

The *bequest* event describes individual bequeathing property to one another.

- **bequeather**: The agent who gives the property.
- **beneficiary**: The agent who receives the property.
- **property**: The property being transferred as a bequest.

Other roles: property-detail.

bid

The *bid* event refers to an offer made by a party to acquire a property, usually in the context of auction-like scenarios which typically occurs after a property seizure, but documented only in few cases in our corpus.

- **bid**: The monetary offer made for the property.
- **bidder**: The individual or party making the offer.
- **property**: The asset that is the object of the bidding.

consent

The *consent* event marks an individual's formal approval of an action, transaction, or agreement, sometimes addressed to another party and referencing a specific subject matter.

- **consenting**: The person or party granting approval.
- **subject**: The action or issue to which consent is given.
- **addressee**: The recipient of the consent.

construction

The *construction* event records the creation or modification of a structure.

- **builder**: The person or party responsible for the construction activity.
- **property**: The location or parcel associated with the construction.
- **structure**: The actual building or structure being constructed.

death

The *death* event marks the passing of an individual, usually included to explain changes in ownership, obligations, or rights.

- **deceased:** The individual who has died.

declaration

The *declaration* event represents the formal act of acknowledging rights, claims, or intentions relevant to legal or financial matters.

- **proclaimer:** The person or entity making the declaration.
- **content:** The statement or subject matter being declared.
- **addressee:** The person or group to whom the declaration is directed.

due-payment

The *due-payment* event captures payments that are given in context of due obligations. Usually the information about the due obligation itself is included in this event. This event is often found as as a description of the obligation with multiple dates added when the payments were made. We followed this pattern in our annotation and representation of the information, using a single event group for all recurring payments in context of the same obligation by the same paying individual.

- **interest:** The financial value that is due periodically.
- **property:** The property which owes the due.
- **payer:** The individual or party responsible for making the payment.
- **beneficiary:** The recipient entitled to receive the payment.
- **interval:** The time interval in which the interest is due (almost always yearly).
- **date-missing:** An date for which no payment is recorded.
- **due-date:** The specific date on which the payment is expected each interval.

- **previous-due:** A reference to an earlier payment obligation, sometimes occurring when the paying party changes.
- **previous-payer:** Added information that someone else paid the interest before the current payer.
- **capital:** The principal amount associated with the obligation, usually given when the due is repayable.
- **interest-redeemable:** Information that or how the interest can be repaid.

Other roles: interest-conv, interval-conv, other, payment-mode, distribution, effective-payer, alternative-payer, proxy, confirmation, prev-interest, quant, property-detail, previous-property, date-due-detail, alternative-interest

endowment

The *endowment* event captures the act of bestowing property, often as a gift or donation, potentially with conditions or reciprocal expectations.

- **property:** The property being endowed.
- **benefactor:** The agent granting or donating the endowment.
- **beneficiary:** The agent receiving the endowment.
- **consideration:** Any form of reciprocal action, obligation, or condition associated with the endowment.
- **consideration-object:** A specific physical entity associated with the consideration.

heritable-lease-grant

The *heritable-lease-grant* event describes the granting of a lease by an entity who holds hereditary rights to the land, typically involving the transfer of property use or ownership rights in exchange for non-redeemable due obligations.

- **lessee:** The individual or party receiving the heritable lease, often the future tenant.
- **property:** The property being granted under the lease.
- **grantor:** The individual or party granting the lease.

- **due:** The due obligation that is tied to the property in return.
- **price:** In rare cases a fixed price to be payed for the lease.

Other roles: grantor-repr, context

inheritance

The *inheritance* event describes the transfer of property from a deceased individual to an heir.

- **heir:** The individual or party receiving the inheritance.
- **decedent:** The deceased individual from whom the property is inherited.
- **property:** The property being transferred.

Other roles: property-detail

litigation

The *litigation* event captures a legal conflict between two parties.

- **party1:** One party involved in the legal dispute, usually the plaintiff.
- **party2:** The opposing party in the dispute, usually the defendant.
- **subject:** The subject matter of the litigation.
- **court:** The institution where the dispute is adjudicated.
- **decision:** The resolution or consequence of the case.

Other roles: delay

offer

The *offer* event describes a situation in which one party proposes to acquire a property from another party for a stated price.

- **offeror:** The individual or party making the offer.
- **offeree:** The individual or party receiving the offer.

- **property:** The property or object that is the subject of the offer.
- **price:** The proposed payment for the property.

Other roles: event-detail

payment

The *payment* event represents the transfer of a monetary sum between parties. This event specifically only addresses cases in which the payment was made in a context which is not already captured through some other event (e.g. *property-purchase* or *due-payment*).

- **sum:** The amount of money involved in the transaction.
- **payer:** The person or party responsible for making the payment.
- **beneficiary:** The recipient of the payment.
- **due-date:** The date on which the payment is expected to be made.
- **cause:** The underlying reason for the payment.

Other roles: alternative-payment

pledge

The *pledge* event marks the act of offering property as collateral in exchange for a loan in contexts that are not covered by *rent-purchase*, for example because the financial obligation is not named or a fixed sum is named instead of an obligation to pay interest.

- **property:** The asset being pledged as security for a debt or obligation.
- **pledger:** The individual or party who offers the property as collateral.
- **pledgee:** The individual or institution receiving the pledge, typically the lender or creditor.
- **capital:** The principal monetary value associated with the pledge.
- **property:** A second mention of pledged asset, possibly used to represent cases of multiple or complex collateral arrangements.

receipt

The *receipt* event marks the acknowledgement of one party receiving a payment from another.

- **payer:** The individual or party making the payment.
- **receiver:** The individual or party receiving the payment.

redemption

The *redemption* event describes the repayment of a loan. As redemption events often describe the due obligation which is absolved by paying it back, all roles from *due-obligation* can also occur in this event.

- **beneficiary:** The individual or party receiving the payment.
- **redeemer:** The individual or party returning the loan.
- **capital:** The sum of money repaid.
- **collateral:** The property pledged as security.
- **due:** The obligation that is absolved by paying the loan.

Other roles: interest-conv, capital-original, interest-remaining

rent-purchase

The *rent-purchase* event describes the purchase of a rent (effectively a *due-obligation* by one party by another party. Unlike the due-obligation resulting from a *heritable-grant-lease*, the seller can pay off the loan to end the obligation. This event may feature roles from the *due-obligation* class whenever the annotation of such an event was omitted.

- **seller:** The individual or party selling a rent.
- **buyer:** The individual or party buying a rent.
- **collateral:** The property pledged as collateral.
- **price:** The amount of money paid by the buyer.
- **due:** The obligation to pay recurring interest.
- **consenting:** Any involved party consenting to the transaction.

Other roles: prev-creditor

revocation

The *revocation* event represents the formal annulment of some contract.

- **proclaimer:** The party who originally proclaimed the content being revoked.
- **revoker:** The party formally revoking the proclamation.
- **content:** The subject matter or act being revoked.

property-purchase

The *property-purchase* event refers to the acquisition of property in exchange for money.

- **buyer:** The acquiring party.
- **seller:** The party relinquishing ownership of the property.
- **witness:** An party attesting to the transaction.
- **consenting:** A party whose agreement is necessary.
- **property:** The object of the purchase.
- **price:** The value exchanged in the transaction.

Other roles: payment-mode, property-detail, offer, payment-detail, price-conv, condition, context, event, rights, seizure, support, temp, time-past.

sanction

The *sanction* event describes a legal or administrative judgment or consequence imposed by an authority, usually following some *litigation* which may or may not be mentioned.

- **court:** The institution issuing the sanction.
- **sanctioned:** The individual or party subject to the sanction.
- **party1:** Involved parties in the process, usually the plaintiff of the original litigation.
- **party2:** Involved parties in the process, usually the defendant of the original litigation.
- **procedure:** The proceedings preceding the decision, usually litigation event.
- **decision:** The outcome or imposed measure.

Other roles: decision-alternative.

seizure

The *seizure* event indicates the legal appropriation of property, usually for debt recovery.

- **property**: The object being seized.
- **claimant**: The party initiating the seizure.
- **debtor**: The individual or entity whose property is seized.
- **executor**: The party carrying out the seizure.
- **capital**: The financial basis for the seizure.
- **date-affirmed**: The date when the seizure is confirmed, given only in rare cases.

Other roles may include: offer.

settlement

The *settlement* event refers to the resolution of a dispute between parties.

- **subject**: The dispute or matter being resolved.
- **party1**: Involved parties in the process, usually the plaintiff of the original litigation.
- **party2**: Involved parties in the process, usually the defendant of the original litigation.

testament

The *testament* event indicates someone declaring his or her heir.

- **bequeather**: The person making the testament.
- **beneficiary**: The intended recipient of the property.
- **property**: The assets to be inherited.

Other roles may include: considered, property-detail.

transfer

The *transfer* event represents the transfer of ownership to a property from one party to another.

- **property**: The property being transferred.
- **owner-prev**: The previous owner of the property.
- **owner-new**: The new owner of the property.
- **value**: The monetary value associated with the transfer.

Other roles: owner-new-detail.

5 Data Formats

5.1 Introduction

In this section of the documentation, we explain what parts of the annotation system are represented in what way in each of the offered data formats. We published the data set in three different formats to adjust to different user types. Two formats are encoded according to TEI guidelines. The first of those is a **simple TEI** format which replicates typical flat named entity annotations as seen in many edition projects in the digital space.¹ A bit more complicated is the **nested TEI** format, which extends the simple TEI format by representing nested objects as well as including all entities, not only named ones. Finally the **BeNASch** format includes the full annotation in a stand-off system. This is the only format which also includes event and relation information.

5.2 Simple TEI

The simple TEI format only represents a small subset of the full annotation. This format is similar to most historical digital editions and should thus be readable by people who know TEI even without studying this document. Table 5.1 lists which tags are in use and which BeNASch annotation they refer to.

Example

```
<persName>Richart Keller</persName> priester caplan  
der <orgName>Pfankilchen s . Martii</orgName>  
verkaufen an <persName>Wolfen Koch</persName> den Moler  
das Hus u . Hofstatt am <placeName>Vischmarkt</placeName>  
glegen , u . ouch <placeName>zum Resslin</placeName> gnant ist ,  
sum <measure type="currency">258 fl .</measure>
```

¹See for example the Collection of Swiss Law Sources Online.

| Tag | Description | Corresponding elements in BeNASch |
|---|-------------------------|---|
| <code><persName/></code> | Name of a Person | head elements of the name class that are nested inside references of the person entity class. |
| <code><orgName/></code> | Name of an Organization | head elements of the name class that are nested inside references of the organization or GPE-ORG entity class. |
| <code><placeName/></code> | Name of a Place | head elements of the name class that are nested inside references of the facility , location or GPE (except GPE-ORG) entity classes. |
| <code><date/></code> | Date | value elements of the date class. |
| <code><measure type="currency"/></code> | Mediums of Exchange | value elements of the money class. |

Table 5.1: List of tags in the simple TEI format with their corresponding concepts in BeNASch.

5.3 Nested TEI

The nested TEI tries to represent all entity annotation elements that are anchored in the text through TEI. This requires nested annotation and annotation of certain syntactic features such as appositions in TEI. As this is usually not done in TEI, no properly fitting tags could be found for some elements, but we tried to substitute them as well as possible while still adhering to the TEI guidelines. Table 5.2 lists which tags are in use and which BeNASch annotation they refer to. Each reference also features a **ref** attribute which groups references to the same entity (on a document level).

| Tag | Description | Corresponding elements in BeNASch |
|---|------------------------------|--|
| <code><rs type="person"/></code> | Reference to a Person | person class references . |
| <code><rs type="organization"/></code> | Reference to an Organization | organization and GPE-ORG class references . |
| <code><rs type="place"/></code> | Reference to a Place | location and GPE (except GPE-ORG) class references . |
| <code><seg type="apposition"/></code> | Apposition | All apposition elements. |
| <code><seg type="?"/></code> | Attribute | attribute elements with the type referring to the class of the attribute. |
| <code><name/></code> | Proper Noun | head elements of the name class. |
| <code><roleName type="?"/></code> | Other Head Elements | head elements not of the name class with the type referring to the class of the head. |
| <code><date/></code> | Date | value elements of the date class. |
| <code><measure type="currency"/></code> | Mediums of Exchange | value elements of the money class. |

Table 5.2: List of tags in the nested TEI format with their corresponding concepts in BeNASch.

Example

```

<rs ref="01" type="person"><name>Richart Keller</name>
<seg type="apposition"><roleName type="occ">priester</roleName></seg>
<seg type="apposition"><roleName type="occ">caplan</roleName>
<rs ref="02" type="organization">der <name>Pfankilchen s . Martii</
  name></rs></seg></rs>
verkaufen an <rs ref="08" type="person"><name>Wolfen Koch</name>
<seg type="apposition">den <roleName type="occ">Moler</roleName></seg>
  </rs>
<rs ref="11" type="place">das <roleName type="type">Hus u . Hofstatt</
  roleName>

```

```
<seg type="loc"><rs ref="12" type="place">am <name>Vischmarkt</name></rs> glegen</seg>
, u . ouch <seg type="apposition"><name>zum Resslerin</name></seg>
gnant ist</rs> , sum <measure type="currency">258 fl .</measure>
```

5.4 BeNASch Standard XML

The final file format is encoded in the standard format decided on in the BeNASch working groups. This format is made to work with a number of scripts, so it can be easily converted to other file formats such as the TEIs described above or the column format for input in machine learning frameworks. This format features three sections beside the Heading which contains the same metadata as the TEI files: Text, Spans, and Events.

Text The text is represented as a sequence of token elements. Each token element features a `token_id` which the spans in the next section refer to.

Spans The spans section represents all elements anchored directly to the text. This includes all entity annotations, but also additional annotations used by events and relations, such as triggers and additional information. Each `span` element includes references to the tokens from the `text` section, a unique id (on document level), information about the classification, and the included text in a short string to ease readability.

Events The events section lists all extracted relations and events. Refer to chapter 2 for more information about events, relations, trigger and roles. Each event features a unique `event_id` (on document level). `eventGroup` elements point to their anchor elements using a `ref` attribute. `trigger` and `role` elements use the `ref` attribute to point to their respective `span` elements.

List of Tables

| | | |
|-----|--|----|
| 2.1 | Example References. | 2 |
| 2.2 | Example References with Head elements in bold. | 3 |
| 2.3 | Example References with Appositions in bold, head of the apposition underlined. | 3 |
| 2.4 | Example References with Attributes in bold. | 3 |
| 2.5 | Example Lists with references in bold and describing elements underlined. | 4 |
| 2.6 | Overview of all entity classes and descriptions in order of frequency in the HLR-GT. | 5 |
| 2.7 | Overview of entity description classes. | 6 |
| 5.1 | List of tags in the simple TEI format with their corresponding concepts in Be- NASch. | 26 |
| 5.2 | List of tags in the nested TEI format with their corresponding concepts in Be- NASch. | 27 |

Acronyms

HTR Handwritten Text Recognition