# Network Construction Methodology

There are several technical considerations when projecting an affiliation network from historical data. Some of these major issues are discussed here.

## Unknown creators

Because this study is focused on the interactions between individuals, it is not possible to include artworks that have no identified creators. For example, some artworks in the RKM dataset are listed with as created by "anoniem" ("anonymous") or "onbekend" ("unknown"). These links are ignored. However, if the same object includes links to any other identified creators, those creators are duly connected in the network model. The BM dataset does not attempt to mark anonymous or unknown artists.

Similarly, artworks with no date whatsoever cannot incorporated into this analysis.

## Production roles

While there are other avenues for participation in the production of a print, such as a sovereign issuing a privilege to a publisher, or an intermediate draughtsman, or a calligrapher providing an inscription. However, it would be difficult to usefully incorporate such orthogonal roles into the macro-scale analysis that is the subject of this work. For one, the amount of richly described links between, for example, prints and the poets whose inscriptions have been used for them is small indeed. Moreover, the outputs of network analysis become less intelligible the more heterogeneous their nodes are.

The RKM uses a more granular set of terms to designate that an artwork was done after the design of an artist, by a publisher, or by an engraver. For this project, I categorize these into three major classes:

1. after: "naar ontwerp van", "naar tekening van", "naar schilderij van", "naar schildering van", "naar eigen ontwerp van", "naar prent van", "naar beeld van", "ontwerper", "cartoraaf", "tekenaar", "inspirator"
2. printmaker: "prentmaker", "vervaardiger", "graveur", "bloksnijder"
3. publisher: "uitvoerder", "supervisie", "verkoper", "boekverkoper", "handelaar", "uitvoerder", "prentverkoper", "boekhandelaar"

The RKM also indicates when associations between objects and artists have been rejected --- these relationships are not incorporated into the network data examined here. Many of their artist-object links have qualifiers, e.g. "documented on the object", or "forgery after". Of these qualifications, this study has accepted the following: "vermeld op object", "eigenhandig gesigneerd", "mogelijk", "toegeschreven aan", "kopie naar", "naar", "mogelijk kopie naar".

## Link direction

Networks may be constructed with either directed or undirected edges between nodes. The decision to construct a network as either directed or undirected will affect many different network metrics, and has a direct bearing on the mapping between network concept and real-world meaning. A directed edge is appropriate when the relationship being represented is one-way, or asymmetrical; for example, a letter sent from individual A to individual B is a directed connection. In an affiliation network such as the one that we have constructed based on artists' involvement with objects, we are dealing with symmetric relationships ("printmaker A participated with publisher B" "publisher B participated with printmaker A"). For this reason, the print production networks in this study will be undirected.

## Link weight

What of the cases where two individuals are tied together by multiple objects? This is where edge weight must be considered. As with directionality, the weight, or strength of connections between any two nodes, can affect many different measures of network properties. The printmaking network presents a challenging weighting problem. On the one hand, it would be easy to weight each connection between two individuals based on the number of prints they were both involved in during a certain span of time; 1 for 1 print, 10 for 10 prints, 100 for 100 prints. However, this raises two issues:

1. This analysis is based on two collection databases, therefore the basic object is not a print *per se*, but an impression *of* that print. Because of this foundation, edge weights based on the number of indexed impressions would be inordinately sensitive to the vagaries of collecting histories in the BM and the RKM. To claim an edge weight based solely on the number of surviving impressions would be to base that weight on a generally arbitrary number subject to a great number of historical contingencies; [@parshall1998, 20--21.]
2. Even if we were able to count how many distinct prints (as opposed to surviving impressions) connected two individuals, to weight edges based on this number would be equivalent to arguing that two individuals who happened to be involved in a book illustration project with fifty different prints are, in a directly meaningful way, *fifty times* as connected to each other as two individuals who are linked by one print alone. This is a dubious claim to make, especially in light of the previous issue.

Therefore, in this study I will be constructing *unweighted* networks that describe only the existence or non-existence of a relationship, and make no claims about its relative strength. This is, undoubtedly, a simplification of reality. However, I argue that, within the bounds of the data now available, it is sufficient to use unweighted networks, as it will still capture the most important aspect of the network (when and between whom a connection existed) while avoiding introducing premature complexity based on problematic assumptions. This simplified network preserves the basic network topology (the existence or non-existence of a relationship between two parties at a certain point in time) while removing the biased weighting of a particular collection, which may have many impressions of the same print.

## Alive vs. dead artists

Constructing each time slice of the network in this way does simplify one important aspect of print production: that printmakers could reproduce designs of artists who were no longer alive, and, furthermore, that publishers could purchase and republish plates by printmakers who were no longer alive. In anticipation of this problem, I ran all analyses in this chapter and the following ones twice: once using this "alive only" method of time slice construction, and once constructing each slice by including all artists who had been born up to that date, including ones who had already died. the choice to exclude dead artists from the network analysis had only a marginal effect on the observed statistics such as centralization. For the sake of visual legibility, all the figures shown here, unless otherwise noted, are based on networks constructed using the "alive only" method.