

Connecting the Dots: Reconstructing an Organized Crime Network from Historical Records

David Byrne

byrned@st.johns.edu

St. John's University

Research Article

Keywords: Organized crime, criminal intelligence, link analysis, network analysis, historical intelligence records, intelligence analysis

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11261375/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Law enforcement records contain information about individuals, associations, locations, businesses, and activities that may be fragmented across documents. This study examines how network analysis and documentary coding can reconstruct these dispersed relationships using the historical intelligence files of a prominent U.S. Mafia crime figure from the 1950s. A standardized coding instrument identified 200 relational observations, of which 123 were designated for network inclusion and subjected to entity reconciliation and link analysis. The resulting network contained 103 distinct entities and 111 unique relational pairs, with 87 entities forming the principal connected component. Investigative and narrative records contributed disproportionately to the resulting network while addresses, locations, businesses, telephone information, and other non-person entities provided important connections alongside individual associates. Structural analysis further identified intermediary entities whose significance was not evident from frequency alone. The findings demonstrate how integrating heterogeneous historical records can reveal relationships and network structures that are less apparent when documents are examined independently. The study provides a replicable, evidence-bounded approach to reconstructing organized-crime networks from intelligence records and establishes a foundation for comparative and technology-assisted research.

1. Introduction

Organized crime presents a persistent intelligence challenge as criminal relationships are rarely documented in a single source. Investigations instead generate information across surveillance reports, criminal-history files, administrative documents, correctional records, and other sources to develop connections between entities. Gill (2000) emphasizes the organization and analysis of such information within the intelligence process, while Ratcliffe (2008) positions analytical integration as central to intelligence-led decision-making. Information sharing and integration is critical when intelligence is dispersed throughout organizational or jurisdictional boundaries Cotter (2017).

This problem is particularly significant for criminal networks. Individual records may identify people, addresses, businesses, locations, communications, and activities without revealing how these entities collectively relate. Incomplete information and uncertain boundaries represent persistent problems for criminal intelligence (Sparrow 1991), underscoring the value of relational approaches for understanding criminal networks (Morselli 2014). By integrating information from disparate records, social network analysis can reveal relationships and actors that might otherwise remain overlooked from varying disparate records (Burcher and Whelan 2018).

Link and network analysis provide systematic mechanisms for evaluating such relationships, drawing on the relational foundations of social network analysis to identify unforeseen associations among seemingly disparate entities (Wasserman and Faust 1994). The value is in revealing patterns of interaction and connectivity (Scott 2017), especially when network analysis can distinguish central

actors, intermediary positions, and relational structures that may be difficult to recognize through descriptive review alone (Borgatti et al. 2018).

This study examines the challenge of using historical intelligence records to ascertain networks and relationships concerning Vito Genovese, an American La Cosa Nostra boss who led the New York Mafia organization that became known as the Genovese crime family. Vito Genovese serves as the focal actor, and the records concerning him define the study corpus. Through documentary coding, entity reconciliation, link analysis, and network analysis, the study examines what becomes visible when dispersed relationships are analyzed collectively. In doing so, it evaluates both the contribution of different documentary sources and the network structure that emerges from their collective analysis.

2. Literature Review

2.1 Organized Crime as an Intelligence Problem

The structure of organized-crime networks is difficult to observe directly because criminal relationships are often concealed, fluid, and incompletely documented. Bouchard (2020) emphasizes the importance of collaboration and network boundaries in understanding organized crime, while Campana (2011) demonstrates how Mafia activities can become functionally dispersed across territories. Campana and Varese (2022) further indicate that representations of organized-crime networks are sensitive to the sources and boundaries from which they are constructed.

Criminal networks often balance operational effectiveness against detection. Morselli et al. (2007) describe an efficiency-security trade-off in which relationships supporting criminal activity can simultaneously create vulnerabilities. The most visible or highly connected actors therefore do not necessarily provide a complete representation of the network.

Empirical research demonstrates the capacity of network analysis to reveal structural patterns within fragmented criminal information. Natarajan (2006) reconstructed a heroin-distribution network from qualitative data, while Bright et al. (2012) applied network analysis to an Australian drug-trafficking syndicate. Geographic relationships also represent an important aspect of criminal networks. For instance, Calderoni et al. (2016) examined the international mobility of Italian Mafia groups, while Campana and Meneghini (2024) applied network analysis to examine organized-crime movement across local communities.

Network reconstruction nevertheless remains inhibited by incomplete evidence. Van der Hulst (2009) notes that criminal and intelligence networks are typically reconstructed from partial observations, while Sparrow (1991) identifies incomplete information, uncertain boundaries, and dynamic relationships as persistent analytical challenges. Burcher and Whelan (2018) further demonstrate that analysts can use network analysis to identify overlooked actors and investigative opportunities while remaining constrained by data limitations and investigator-generated bias. Organized crime can therefore be understood as both a criminal-network and a challenge to intelligence collection and analysis.

2.2 Criminal Intelligence and Analytical Synthesis

Criminal intelligence is distinguished from information by the analytical processes applied to collected material. Harfield and Harfield (2008) conceptualize intelligence as involving the acquisition, organization, retrieval, analysis, and interpretation of information, while Cope (2004) describes intelligence as information enhanced through analytical processing. Therefore, interpretation is necessary for information to guide law-enforcement action (Innes et al. 2005).

This analytical function gained increasing importance with the emergence of intelligence-led policing. Manning (2000) characterizes modern policing as increasingly dependent upon gathering, organizing, retrieving, and analyzing information, and Maguire and John (2006) associate intelligence-led policing with the strategic management of criminal problems and risks. Finally, Sheptycki (2002) emphasizes the identification and analysis of criminal threats. Collectively, these approaches position analysis between information collection and informed action.

Analytical synthesis requires disparate observations to be evaluated collectively. This process depends on evidence (McCue and Parker 2003) and adaptive reasoning under ambiguity and incomplete information is necessary (Puvathingal and Hantula 2012). Gentry (2016) further situates analytical judgment within the professionalization of intelligence analysis. The cognitive difficulties associated with incomplete and ambiguous evidence is often complicate the analytical process (Heuer 1999). Therefore, there is a need for structured techniques to aid in the evaluation of information to reveal alternative interpretations (Heuer and Pherson 2015). Marrin (2011) similarly frames analysis as a process for reducing uncertainty and supporting informed judgements and decision making.

The metaphor of *connecting the dots* describes this process: analysis draws conclusions by evaluating evidence, hypotheses, and arguments (Tecuci et al. 2016). Connecting disparate information is a complex analytical task (Bradel et al. 2013). In documentary intelligence, relevant observations may already exist, but analytical value emerges when meaningful relationships among them are identified.

2.3 Link Analysis and Criminal Networks

Social network analysis (SNA) provides a formal means of examining relationships rather than treating actors as isolated units. Wasserman and Faust (1994) established the relational foundations of SNA, while Scott (2017) emphasizes how interaction and connectivity contribute to broader social structure. Building on these foundations, Borgatti et al. (2018) further illustrate how network analysis can identify influential actors, clusters, and structural patterns that descriptive review alone may obscure.

Criminological applications demonstrate how network methods can identify structure within investigative and qualitative material. For example, Natarajan (2006) reconstructed a heroin-distribution network from qualitative data; Bright, Hughes, and Chalmers (2012) examined an Australian drug-trafficking syndicate; Heber (2009) analyzed drug-offender networks; and Morselli and Petit (2007) examined disruption within a drug-importation network. Collectively, these studies express the analytical

shift from identifying individual actors to examining the structure and significance of the relationships connecting them.

Network positions can also reveal significance not captured by frequency alone. Borgatti et al. (2018) distinguish highly connected actors from those occupying intermediary or brokerage positions, while Cavallaro et al. (2020) and da Cunha and Gonçalves (2018) demonstrate how structural analysis can identify dependencies and vulnerabilities within criminal networks. Consequently, an entity with relatively few direct ties may nonetheless occupy a structurally significant position within the network.

Such measures require cautious interpretation as imperfect data can affect centrality measures (Borgatti et al. 2006). Campana and Varese (2022) show that reconstructed organized-crime networks remain dependent on their evidentiary sources and boundaries. Drawing on mid-twentieth-century Federal Bureau of Narcotics records, DellaPosta (2017) similarly demonstrates how historical Mafia intelligence can support formal relational analysis while remaining bounded by the archival corpus from which the network was constructed.

2.4 Relationship-Centered Intelligence and Analytical Integration

Criminal intelligence frequently requires records created for different purposes to be examined collectively. Ericson and Haggerty (1997) describe policing as increasingly dependent on information generated across institutional systems, while Sheptycki (2002) emphasizes difficulties that arise when intelligence crosses organizational and jurisdictional boundaries. Consequently, individual documentary sources may therefore provide only partial representations of the same criminal environment.

Examining these partial representations collectively allows information to be compared and contextualized. Because evaluating evidence under conditions of uncertainty remains inherently challenging (Heuer 1999), structured approaches are necessary for systematically comparing information from multiple sources (Heuer and Pherson 2015). Entities derive meaning partly from their connections (Wasserman and Faust 1994), while relational structures can reveal patterns not apparent from individual attributes alone (Borgatti et al. 2018).

Working with multiple records requires more than simple access to additional information. Analysts must exercise judgement to assess the relevance and significance of available evidence (Heuer and Pherson 2015), while Lowenthal (2022) distinguishes collected information from intelligence produced through interpretation. Historical records provide a valuable context because they allow relationships to be reconstructed within a defined corpus. The central challenge lies in ascertaining relationships and connections from different types of records not designed for intelligence or investigative purposes to guide decision-making.

2.5 Research Gap and Study Contribution

Network approaches are well established in organized-crime research. Burcher and Whelan (2018) demonstrate the operational value of social network analysis for criminal intelligence, while Campana and Varese (2022) emphasize how data sources and network boundaries shape reconstructed organized-crime networks. Campana and Meneghini (2024) further demonstrate the value of relational approaches for examining organized-crime activity across geographic settings.

Documentary evidence is likewise widely used, but its methodological treatment requires transparency. Windle and Silke (2019) identify methodological concerns surrounding documentary sources in organized-crime research, while Chainey and Alonso Berbotto (2022) demonstrate how structured coding and document analysis can systematically extract and validate organized-crime information.

A gap remains between structured documentary extraction and network reconstruction. Existing research demonstrates how information can be coded and how network methods can reveal criminal relationships. However, less attention has been directed toward how heterogeneous records contribute to a unified network.

This study addresses this intersection by treating a bounded historical intelligence corpus as evidence for network reconstruction. Through standardized documentary coding, network-inclusion screening, entity reconciliation, link analysis, and structural network analysis, it examines both the contribution of individual source records and what becomes visible when their dispersed relationships are integrated.

Building upon this literature, the study addresses two research questions:

RQ1. What types of relational information are contained within the historical intelligence records concerning Vito Genovese, and how do different source documents and relationship types contribute to network reconstruction?

RQ2. What network structure emerges through entity reconciliation and integration of the documented relationships, and what does the reconstructed network reveal that is not readily apparent from individual source records?

3. Methodology

3.1 Research Design and Evidentiary Boundary

This study employed qualitative documentary analysis, structured relational coding, and network analysis to examine whether relationships dispersed across historical intelligence records could be reconstructed as an organized-crime network. Analysis proceeded from document review and coding through network screening, entity reconciliation, and network construction.

Vito Genovese, the La Cosa Nostra boss (Maas 1968) whose surname became the conventional law-enforcement designation for the Genovese crime family (US Senate Committee on the Judiciary 1983), served as the focal actor, while the documentary corpus concerning him established the evidentiary

boundary. Only information appearing within these records was eligible for coding; external sources were not used to supplement or expand relationships. All documented persons, addresses, businesses or organizations, locations, telephone numbers, aliases or identifiers, events or activities, vehicles or assets, and other entities were potentially eligible network nodes when involved in a coded relationship. This boundary ensured that the resulting network represented what could be reconstructed from the historical intelligence record rather than from retrospective knowledge.

3.2 Historical Source Records

The historical records analyzed in this study were accessed through a published compilation reproducing original law-enforcement and government intelligence records concerning organized crime (Carr 2012). The compilation served as the documentary repository for the study; the units of analysis were the individual historical records reproduced within it rather than the secondary narrative of the volume. The corpus was comprised of ten historical documents representing investigative, administrative, identification, and institutional records. These included investigative memoranda; FBI and NYPD identification, arrest, intelligence, and criminal-history records; an immigration and naturalization memorandum; and Bureau of Prisons identification, classification, criminal-history, personal, and business-background materials. The records documented varying combinations of identity, criminal history, associations, locations, family, employment, immigration, telephone, business, and investigative information.

The unit of coding was the documented relational observation within each reproduced historical record, not the published compilation itself. A source could therefore generate multiple observations, each retaining its identifier so that coded relationships remained traceable to the originating record.

3.3 Relational and Evidentiary Coding

A standardized coding instrument converted documentary information into Entity A, Relationship, and finally Entity B, Structure. For each observation, coders recorded the source document, entities and entity types, relationship and relationship category, evidentiary basis, whether the relationship was explicit or inferred, confidence, network-inclusion status, and explanatory notes. Relationship categories included criminal/investigative, residential, geographic/co-location, identity/alias, familial, legal/judicial, communication, business, employment, and other relationships.

Coding was initially inclusive so that source information was preserved before network eligibility was determined. Explicitly stated relationships were distinguished from analytical inferences. Observations were assigned high, moderate, or low confidence according to evidentiary clarity. All 200 coded observations were explicitly supported by the source records; 191 were rated high confidence and nine moderate confidences.

3.4 Network Screening and Entity Reconciliation

Coded observations were separately screened as *Yes*, *No*, or *Review* for network inclusion. This prevented contextual or administrative information from automatically becoming network edges. Of the

200 observations, 123 were designated for network inclusion, 71 were excluded from the network but retained as contextual information and six required additional review. The 123 network-ready observations formed the evidentiary base for reconciliation and network construction.

Network-ready entities were then reconciled to consolidate documented aliases, name variants, and formatting differences referring to the same underlying entity. A conservative rule was applied: references were merged only when supported by the source evidence; uncertain matches remained separate. Canonical identities were retained alongside an audit record of reconciliation decisions.

3.5 Network Construction and Structural Analysis

Each reconciled entity was represented as a node and each documented relationship as an edge. Repeated observations involving the same reconciled entities were retained as supporting evidence but collapsed into a unique relational pair for structural calculations. This prevented repeated documentation from artificially increasing connectivity. Persons and non-person entities, including addresses, businesses, locations, telephone numbers, and events, were retained as nodes when supported by the records.

The cleaned network contained 103 reconciled entities and 111 unique relational pairs across seven connected components. The largest component contained 87 entities and served as the primary network visualization. Structural analysis examined direct connections, degree centrality, betweenness centrality, articulation points, bridge edges, and relational clusters. These measures assessed connectivity, intermediary positions, structural dependency, and groupings. They were not interpreted as indicators of criminal rank or formal organizational authority. Community detection was applied to the largest component, with resulting communities interpreted as relational clusters rather than formal organized-crime crews.

The final link chart was constructed in i2 Analyst's Notebook using the cleaned, entity-reconciled network data; the software was used for visualization rather than for determining entity inclusion or relationship coding.

3.6 Analytical Strategy

Analysis proceeded in three stages. First, descriptive statistics characterized the complete 200-observation intelligence record. These included relationship categories and source-document contributions. Second, the 123 network-ready observations were examined to determine which relationship and document types contributed to network reconstruction. Third, the entity-reconciled network was analyzed for connectivity, centrality, structural connectors, bridges, and relational clusters.

This staged approach maintained a distinction between source information, network-eligible evidence, and reconstructed structure. The resulting network was interpreted as the intelligence picture supported by the bounded Genovese corpus rather than a comprehensive retrospective representation of the criminal organization.

Figure 1 summarizes the methodological progression from source records to the final network structure. The figure distinguishes the successive stages of documentary coding, network screening, entity reconciliation, and structural analysis while preserving the evidentiary boundary established by the source corpus.

Figure 1. Methodological Progression from Historical Intelligence Records to Network Reconstruction

As reflected in Fig. 1, the progressive refinement of the evidence is illustrated rather than a simple reduction of the dataset. The 200 coded observations preserved the broader intelligence record, while screening identified 123 observations suitable for network analysis. Entity reconciliation resolved aliases and name variants before repeated relationships were consolidated into 111 unique relational pairs. The resulting 87-entity principal component represented the largest structure supported by the historical records

4. Results

The ten historical source documents produced 200 coded observations, of which 123 were designated for network inclusion. Results proceed from the composition and screening of the coded evidence to source contribution and the structure of the entity-reconciled network.

4.1 Composition and Network Screening

The documentary analysis produced 200 coded observations across ten relationship categories. *Criminal/Investigative* information constituted the largest category (n = 68; 34.0%), followed by *Residential* (n = 27; 13.5%) and *Geographic/Co-location* relationships (n = 23; 11.5%). The remaining observations represented *Identity/Alias*, *Familial*, *Legal/Judicial*, *Communication*, *Business*, *Employment*, and *Other Relationships*.

The coded record extended beyond person-to-person associations. Persons were the originating entity in 186 observations (93.0%), while connected entities included events or activities (n = 57), persons (n = 45), addresses (n = 37), aliases or identifiers (n = 15), businesses or organizations (n = 14), locations (n = 13), telephone numbers (n = 10), other entities (n = 6), and vehicles or assets (n = 3). All 200 observations were explicitly documented; 191 (95.5%) received high-confidence ratings and nine (4.5%) moderate-confidence ratings.

Of the 200 observations, 123 (61.5%) were designated for network inclusion, 71 (35.5%) were retained as contextual information, and six (3.0%) remained pending review. Retention varied by relationship category. *Residential* and *Communication* observations were fully retained, while *Geographic/Co-location* (91.3%) and *Familial* (93.3%) observations were largely retained. By comparison, 26.5% of *Criminal/Investigative* and 30.8% of *Legal/Judicial* observations were retained as network-ready relationships.

Note

Light bars represent all coded observations (N = 200), and dark bars represent observations designated for network inclusion (N = 123). Percentages within each series describe the composition of the respective dataset; retention percentages indicate the proportion of coded observations within each category retained as network-ready relationships.

Figure 1 shows that network screening changed the composition of the relational evidence rather than uniformly reducing it. *Criminal/Investigative* information declined from 68 coded observations to 18 network-ready relationships, whereas *Residential* and *Communication* relationships were fully retained and *Geographic/Co-location* and *Familial* relationships remained largely represented.

4.2 Contribution of Source Records

Individual records varied substantially in both information volume and network yield. The largest source produced 49 coded observations, of which 44 were retained for network analysis (89.8%); another investigative memorandum contributed 21 of 23 (91.3%). All 17 observations contained in the NYPD arrest/intelligence bulletin were network ready. In contrast, the criminal-history compilation produced 27 observations, but only six network-ready relationships (22.2%), while the NYPD criminal-identification record contributed one of 13 (7.7%). Two institutional records contributed to no network-ready relationships.

Grouping records by documentary function produced a similar pattern. Investigative memoranda generated 65 network relationships from 72 coded observations (90.3%), while the police intelligence/arrest bulletin yielded 17 of 17. The Personal/Business background record yielded 15 of 20 relationships (75.0%), and the Administrative/Naturalization record yielded 12 of 20 (60.0%). Identification/Criminal-history records generated 56 observations but only 14 network relationships (25.0%).

Collectively, the principal relational sources produced 109 of the 123 network-ready relationships (88.6%). Thus, the volume of coded information within a source did not necessarily correspond with its contribution to network reconstruction.

4.3 Reconstructed Network Structure

Following entity reconciliation, the 123 network-ready observations produced a cleaned network of 103 distinct entities and 111 unique relational pairs distributed across seven connected components. The principal component contained 87 entities (84.5% of the complete network), and overall network density was .021.

The network-ready evidence remained heterogeneous. *Residential* relationships were the largest category (n = 27; 22.0%), followed by *Geographic/Co-location* (n = 21; 17.1%) and *Criminal/Investigative* relationships (n = 18; 14.6%). *Residential* and *Geographic/Co-location* relationships together accounted

for 48 of the 123 network-ready observations (39.0%), while *Familial, Communication, Business, and Employment relationships* provided additional relational structure.

Vito Genovese occupied the dominant structural position, with 43 direct connections, degree centrality of .422, and betweenness centrality of .645. His wife, Anna Genovese, had ten direct connections, followed by Frasca with seven. Several secondary persons and non-person entities also occupied notable structural positions.

Table 1

Selected Principal Entities and Their Structural Roles in the Reconstructed Intelligence Network

Entity	Entity Type	Direct Connections (n)	Degree Centrality	Betweenness Centrality	Structural Connector
Vito Genovese	Person	43	.422	.645	Yes
Anna Genovese	Person	10	.098	.169	Yes
Frasca	Person	7	.069	.080	Yes
Ocean Blvd.	Address	5	.049	.082	Yes
Miranda	Person	5	.049	.020	Yes
Smurra	Person	5	.049	.020	Yes
Farm	Location	4	.039	.052	Yes
Profaci	Person	4	.039	.035	Yes
DeFeo	Person	4	.039	.012	No
Callendriello	Person	3	.029	.074	Yes

Note

The table presents selected entities illustrating prominent connectivity, intermediary position, and structural-connector roles within the reconstructed network; it is not an exhaustive listing of network entities. Degree centrality represents normalized direct connectivity, while betweenness centrality reflects intermediary position along network paths.

Structural significance was not determined solely by direct connectivity. Callendriello had three direct connections but betweenness centrality of .074, while Ocean Blvd (.082) and Farm (.052) also occupied intermediary positions. The network contained 28 articulation points and 78 bridge edges, indicating dependence on particular persons, places, and relationships connecting otherwise separated portions of the sparse network.

4.4 Relational Clusters and Integrated Network Structure

Community analysis of the 87-entity principal component identified 14 relational clusters. These communities are distinct from the seven connected components of the complete network: the latter represent structurally disconnected portions of the network, whereas the collection represents relational groupings within the principal component.

The principal entities summarized in Table 1 represent only a subset of the broader network; examination of relational communities identified additional persons, businesses, places, and other entities that contributed to individual cluster structure.

Several larger communities were apparent. The largest was centered on Vito Genovese and integrated business, familial, residential, investigative, and activity-related information. Other prominent communities included an Anna Genovese grouping characterized by familial and residential ties; an investigative grouping connecting Frasca, Miranda, Smurra, De Feo, and the Broome Street–Mulberry St. vicinity; and a Profaci-related grouping involving the Farm, Callendriello, Familietti, and Sea Coast Liquor Distributors. These communities emerged through integration of relationships documented across multiple records rather than from information contained in a single source.

Figure 2 extends beyond the selected entities reported in Table 1 by visualizing the complete 87-entity principal component, illustrating the broader pattern of documented relationships, structural connectivity, and relational communities revealed through network reconstruction. Callout labels are descriptive aids used to identify prominent relational communities and should not be interpreted as formal organizational units or established crews.

Note

Node shape represents entity type; lines represent documented relationships between entities and greater line width reflects direct connectivity. *Person, business/organization, and place labels are abbreviated for visual clarity; Identifier/Communication and Other Entity nodes are represented by symbol only. Cluster designations identify relational communities within the 87-entity principal component. Visualization created using i2 Analyst's Notebook.*

Figure 2 visualizes the complete 87-entity principal component, preserving the documented relational paths connecting the focal actor to surrounding entities and relational communities.

Taken together, the results demonstrate that network reconstruction depended not simply on the volume of historical information available, but on the relational content retained through screening and integration. The Discussion considers these findings in relation to organized-crime intelligence, analytical synthesis, and criminal-network research.

4.5 Investigative Intelligence Revealed Through Network Reconstruction

The reconstructed network demonstrates how relational analysis could have provided investigative value beyond reviewing the historical records individually. Cluster analysis separated the principal component into distinct relational communities, allowing groups of people, places, businesses, communications, and activities appearing across different records to be examined collectively. Rather than treating each documented association as an isolated lead, the resulting structure identifies where relationships converge and which entities connect otherwise differentiated portions of the network.

Several findings illustrate this value. The Anna Genovese grouping concentrates on familial and residential relationships around an actor who also occupied the second-highest structural position in the network. The Frasca-related grouping brings together people, places, and investigative information that becomes more readily interpretable when viewed as a relational community. Similarly, the Profaci-related grouping connects individuals with a business and place-based entities that were dispersed across the underlying records. These patterns could have directed investigators toward relationships and entities warranting comparison across case files rather than evaluation within a single document.

The analysis also identifies potential investigative priorities that are not apparent from connection counts alone. Callendriello, for example, had only three direct connections but comparatively high betweenness centrality (.074), while 130 Ocean Blvd (.082) and the Farm (.052) also occupied intermediary positions. Such entities could warrant attention because they connect portions of the documented network despite having relatively few direct relationships. The identification of 28 articulation points and 78 bridge edges similarly highlight persons, places, and relationships upon which portions of the reconstructed network depended.

These findings do not establish that historical investigators failed to recognize these relationships. Rather, they demonstrate the additional intelligence picture that becomes available when fragmented documentary evidence is reconciled and analyzed as an integrated network. Had such relational analysis been available contemporaneously, it could have supported lead prioritization, cross-record comparison, identification of intermediary entities, and assessment of relationships extending beyond the focal actor.

5. Discussion

The intelligence value of network reconstruction lies not merely in visualizing known associations, but in changing the unit of analytical attention from individual records and actors to relational communities, intermediary entities, and cross-community connections. In this study, relationships that appeared fragmented across individual historical records became structurally meaningful when reconciled and examined collectively. The resulting network identified not only highly connected actors but also intermediary persons and places whose significance was less apparent from direct connection counts alone.

5.1 Research Questions and Principal Findings

The findings provide direct answers to both research questions. **RQ1** examined the relationships contained within the historical records and the contribution of different sources to the resulting network. The corpus produced 200 coded observations, of which 123 were designated for network inclusion. *Residential* and *Geographic/Co-location* relationships accounted for 39.0% of the network-ready evidence, while documents yield varied substantially. *Investigative Memoranda* produced 65 network relationships from 72 observations (90.3%), compared with 14 from 56 *Identification/Criminal-History* observations (25.0%). Information volume therefore did not necessarily correspond with network-reconstruction value.

RQ2 examined the structure produced through entity reconciliation and integration. The resulting network contained 103 distinct entities and 111 unique relational pairs across seven components, with 87 entities forming the principal component. Although Vito Genovese occupied the dominant structural position, integration revealed intermediary persons and non-person entities, relational clusters, and connecting paths distributed across multiple records. The reconstructed network therefore exposed structure that was less apparent when the documents were examined independently.

5.2 From Documentary Information to Relational Intelligence

The findings reinforce the distinction between collecting information and producing intelligence. Cope (2004) and Innes et al. (2005) locate analytical value in the interpretation and integration of information rather than accumulation alone. Here, only 123 of 200 coded observations were designated network-ready, preserving contextual information without automatically converting every documented fact into a relationship.

This screening also reflects the structured approach advocated by Chainey and Alonso Berbotto (2022), who emphasize the importance of explicit coding and quality assurance when extracting organized-crime information from source documentary. Retaining source identifiers and evidentiary status preserved traceability, thus producing evidence-bounded analytical reconstruction rather than a visualization of accumulated historical information.

5.3 The Collective Contribution of Heterogeneous Records

The findings demonstrate the complementary value of heterogeneous records. Investigative, identification, administrative, correctional, and criminal-history documents contributed different forms of evidence, connecting individuals with residences, businesses, legal proceedings, communications, and geographic locations.

This supports Campana and Varese's (2022) argument that observed network structure is partly shaped by its underlying sources. Records created for different purposes contributed different pieces of the reconstructed network. Consistent with Campana and Antonopoulos (2024), entities that appeared routine in isolation, such as address or business, could acquire greater analytical significance through their relationships with entities documented elsewhere.

5.4 Connecting the Dots Across Records

The integrated network gives the study's *connecting the dots* concept empirical form. Bradel et al. (2013) describe this analytical task as establishing connections across dispersed information to produce a more coherent understanding. The Genovese corpus demonstrated a comparable process: separately documented relationships became structurally meaningful when assembled within a common network.

Integration also exposed paths among secondary entities rather than merely aggregating references to Genovese. Profaci, for example, connected with a telephone number, address, the Farm, and Callendriello; the Farm connected separately with a location, Familietti, and Frasca. Other evidence connected Frasca with Miranda and Smurra. Such paths illustrate what integration adds beyond sequential document review.

5.5 Structural Position and Analytical Significance

The cleaned network demonstrates why analytical significance cannot be reduced to frequency. Degree centrality captures direct connectivity, while betweenness identifies entities positioned along paths connecting other portions of the network. Borgatti et al. (2018) emphasize this distinction between connectivity and intermediary position, while Morselli et al.'s (2007) efficiency-security argument suggests why less-connected actors may nevertheless occupy important positions within covert networks.

The identified bridges and articulation points reinforce this distinction. Structural connectors included not only individuals but businesses, addresses, locations, and other entities. These measures describe relational position, and should not be interpreted as evidence of criminal rank or organizational authority.

5.6 Network Structure Beyond the Focal Actor

Although Vito Genovese was intentionally established as the focal actor, the network extended beyond direct Genovese-to-associate relationships. Seven connected components emerged, with 87 of 103 entities forming the principal component.

This accords with the relational perspective described by Campana and Antonopoulos (2024) and the geographic network approach demonstrated by Campana and Meneghini (2024). Secondary connections among associates, addresses, businesses, family members, and locations expanded the analytical picture beyond a conventional list of known associates. Nevertheless, resulting clusters should be interpreted conservatively. Consistent with Fielding's (2017) caution regarding secrecy and observed network organization, they represent concentrations within the available evidence rather than established Mafia crews or formal organizational units.

5.7 What the Integrated Network Adds

The principal contribution is not the discovery that Vito Genovese had documented associates; the historical records already establish those associations. Rather, the analysis demonstrates how dispersed

observations combine into a broader intelligence picture.

This distinction is consistent with Campana and Varese's (2022) warning that reconstructed networks remain products of their sources and evidentiary boundaries. The present network is therefore not a complete representation of the Genovese criminal organization. Within that boundary, however, integration reveals how independently documented people, places, businesses, communications, and events intersect. In this sense, the study gives *connecting the dots* a concrete analytical form: documentary records provide the observations, coding preserves their evidentiary origins, reconciliation resolves recurring entities, and network analysis reveals their collective structure.

5.8 Limitations

The network represents relationships documented within the bounded Genovese corpus rather than a complete reconstruction of the criminal organization. Campana and Varese (2022) emphasize that network structure is shaped by source selection and evidentiary boundaries. Accordingly, an entity or relationship's absence from the present network indicates only that it was not documented within the defined corpus, rather than that it did not exist.

The records were also created for investigative, administrative, correctional, and identification purposes rather than network analysis, producing substantial variation in relational yield. This source dependence reflects concerns regarding incomplete criminal-network data identified by Borgatti et al. (2006) and methodological challenges associated with documentary evidence discussed by Windle and Silke (2019).

Finally, entity reconciliation and network-inclusion decisions required analytical judgment. Conservative rules reduced unsupported connections but may have left genuine relationships unresolved. The network should therefore be interpreted as evidence-bounded analytical reconstruction, not a definitive representation of the historical criminal network.

5.9 Implications for Contemporary Intelligence Practice and Future Research

The findings have implications beyond the historical Genovese case. Burcher and Whelan (2018) demonstrate the value of network analysis for criminal intelligence, while Chainey and Alonso Berbotto (2022) emphasize reproducible procedures for extracting organized-crime information from documentary sources. This study combines these principles through structured coding, source traceability, entity reconciliation, and network reconstruction.

This approach may be useful when analysts confront records containing recurring people, addresses, businesses, communications, and activities. Rather than treating link analysis as a final visualization, it can form part of a documented analytical process in which relationships remain connected to their evidentiary sources.

Future research should apply the method to additional organized-crime corpora. Applying the same procedures to records concerning leaders of other Mafia families would permit comparison of source yield, network composition, centrality, bridging entities, and relational clusters. Such comparisons could help distinguish Genovese-specific findings from broader patterns in historical organized-crime intelligence.

Subsequent research could also test whether AI-assisted extraction and entity reconciliation reproduce manually identified relationships while maintaining evidentiary traceability. Such applications should be built upon, rather than replace the manually coded baseline established here.

6. Conclusion

Organized-crime intelligence is often distributed across records that provide only partial views of people, places, activities, and associations. This study demonstrated how a bounded collection of historical intelligence records concerning Vito Genovese could be coded, reconciled, and reconstructed as a network.

The findings show that the analytical value of these records extends beyond the information contained in any single document. Integrating relationships across investigative, administrative, identification, and institutional sources revealed connections among people, addresses, businesses, locations, communications, and activities that became more apparent when examined collectively. The resulting network should not be viewed as a complete representation of the Genovese criminal organization, but as an evidence-bounded reconstruction of what the corpus supports.

Ultimately, *connecting the dots* requires more than collecting the *dots*. It requires a transparent process for determining which information represents meaningful relationships, reconciling recurring entities, preserving evidentiary provenance, and examining the resulting structure. By demonstrating that process with historical organized-crime records, this study provides a foundation for comparative network research and future applications of emerging analytical technologies.

Declarations

Competing Interests:

The authors have no relevant financial or non-financial interests to disclose.

Funding:

No funding was received for conducting this study.

Author Contribution

The authors were responsible for study conception and design, methodology, data analysis and interpretation, and preparation and revision of the manuscript. D.B.: Conceptualization, methodology, formal analysis, investigation, data curation, visualization, writing, original draft, writing, review and editing, supervision, and project administration; S.M.: Investigation, formal analysis, and writing, review, and editing; G.T.: Investigation, data curation, formal analysis, and writing, review and editing. All authors reviewed the manuscript.

Acknowledgement

The author acknowledges Peter Paul Grech and Hailey Oswick, members of the Investigative Research & Innovative Scholars (IRIS) program at St. John's University, for their contributions to the development of the network visualizations using i2 Analyst's Notebook.

Data Availability

The historical intelligence records analyzed in this study are available through the published documentary sources cited in the manuscript. The coded and entity-reconciled dataset generated from these records and used for network reconstruction and analysis is available from the corresponding author upon reasonable request.

References

1. Borgatti SP, Carley KM, Krackhardt D (2006) On the robustness of centrality measures under conditions of imperfect data. *Soc Netw* 28(2):124–136. <https://doi.org/10.1016/j.socnet.2005.05.001>
2. Borgatti SP, Everett MG, Johnson JC (2018) *Analyzing social networks*, 2nd edn. SAGE, London
3. Bouchard M (2020) Collaboration and boundaries in organized crime: a network perspective. *Crime Justice* 49:425–469. <https://doi.org/10.1086/708435>
4. Bradel L, Self JZ, Endert A, Hossain MS, North C, Ramakrishnan N (2013) How analysts cognitively “connect the dots.” In: 2013 IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE, Piscataway, pp 24–26. <https://doi.org/10.1109/ISI.2013.6578780>
5. Bright DA, Hughes CE, Chalmers J (2012) Illuminating dark networks: a social network analysis of an Australian drug trafficking syndicate. *Crime Law Soc Change* 57(2):151–176. <https://doi.org/10.1007/s10611-011-9336-z>
6. Burcher M, Whelan C (2018) Social network analysis as a tool for criminal intelligence: understanding its potential from the perspectives of intelligence analysts. *Trends Organ Crime* 21(3):278–294. <https://doi.org/10.1007/s12117-017-9313-8>

7. Calderoni F, Berlusconi G, Garofalo L, Giommoni L, Sarno F (2016) The Italian mafias in the world: a systematic assessment of the mobility of criminal groups. *Eur J Criminol* 13(4):413–433. <https://doi.org/10.1177/1477370815623570>
8. Campana P (2011) Eavesdropping on the Mob: the functional diversification of Mafia activities across territories. *Eur J Criminol* 8(3):213–228. <https://doi.org/10.1177/1477370811403442>
9. Campana P, Antonopoulos GA (2024) A relational approach to organised crime: introduction to the special issue on crime and networks. *Trends Organ Crime* 27(3):229–234. <https://doi.org/10.1007/s12117-024-09534-4>
10. Campana P, Meneghini C (2024) Organised crime movement across local communities: a network approach. *Trends Organ Crime* 27(3):286–313. <https://doi.org/10.1007/s12117-024-09531-7>
11. Campana P, Varese F (2022) Studying organized crime networks: data sources, boundaries and the limits of structural measures. *Soc Netw* 69:149–159. <https://doi.org/10.1016/j.socnet.2020.03.002>
12. Carr C (2012) New York police files on the Mafia: a collection of New York’s five families’ files, arrest records and mug shots from the 1940s and 1950s. Hosehead Productions, USA
13. Cavallaro L, Ficara A, De Meo P, Fiumara G, Catanese S, Bagdasar O, Song W, Liotta A (2020) Disrupting resilient criminal networks through data analysis: the case of Sicilian Mafia. *PLoS One* 15(8):e0236476. <https://doi.org/10.1371/journal.pone.0236476>
14. Chainey SP, Alonso Berbotto A (2022) A structured methodical process for populating a crime script of organized crime activity using OSINT. *Trends Organ Crime* 25(3):272–300. <https://doi.org/10.1007/s12117-021-09428-9>
15. Cope N (2004) Intelligence led policing or policing led intelligence?: integrating volume crime analysis into policing. *Br J Criminol* 44(2):188–203. <https://doi.org/10.1093/bjc/44.2.188>
16. Cotter RS (2017) Police intelligence: connecting-the-dots in a network society. *Policing and Society* 27(2):173–187. <https://doi.org/10.1080/10439463.2015.1040794>
17. da Cunha BR, Gonçalves S (2018) Topology, robustness, and structural controllability of the Brazilian Federal Police criminal intelligence network. *Appl Netw Sci* 3:36. <https://doi.org/10.1007/s41109-018-0092-1>
18. DellaPosta D (2017) Network closure and integration in the mid-20th century American mafia. *Soc Netw* 51:148–157. <https://doi.org/10.1016/j.socnet.2016.11.005>
19. Ericson RV, Haggerty KD (1997) *Policing the risk society*. University of Toronto Press, Toronto. <https://doi.org/10.3138/9781442678590>
20. Fielding NG (2017) The shaping of covert social networks: isolating the effects of secrecy. *Trends Organ Crime* 20(1–2):16–30. <https://doi.org/10.1007/s12117-016-9277-0>
21. Gentry JA (2016) The “professionalization” of intelligence analysis: A skeptical perspective. *International Journal of Intelligence and Counterintelligence* 29(4):643–676. <https://doi.org/10.1080/08850607.2016.1177393>

22. Gill P (2000) Rounding up the usual suspects? Developments in contemporary law enforcement intelligence. Routledge, Abingdon. <https://doi.org/10.4324/9781315186016>
23. Harfield C, Harfield K (2008) Intelligence: investigation, community and partnership. Oxford University Press, Oxford
24. Heber A (2009) The networks of drug offenders. *Trends Organ Crime* 12(1):1–20. <https://doi.org/10.1007/s12117-008-9055-8>
25. Heuer RJ Jr (1999) Psychology of intelligence analysis. Center for the Study of Intelligence, Central Intelligence Agency, Washington, DC
26. Heuer RJ Jr, Pherson RH (2015) Structured analytic techniques for intelligence analysis, 2nd edn. CQ Press, Washington, DC
27. Innes M, Fielding N, Cope N (2005) The appliance of science?: The theory and practice of crime intelligence analysis. *British Journal of Criminology* 45(1):39–57. <https://doi.org/10.1093/bjc/azh053>
28. Lowenthal MM (2022) Intelligence: from secrets to policy, 9th edn. CQ Press, Thousand Oaks, CA
29. Maas P (1968) The Valachi papers. GP Putnam's Sons, New York
30. Maguire M, John T (2006) Intelligence led policing, managerialism and community engagement: competing priorities and the role of the National Intelligence Model in the UK. *Policing and Society* 16(1):67–85. <https://doi.org/10.1080/10439460500399791>
31. Manning PK (2000) Policing new social spaces. In: Sheptycki JWE (ed) *Issues in transnational policing*. Routledge, London, pp 177–200. <https://doi.org/10.4324/9780203186091>
32. Marrin S (2011) Improving intelligence analysis: bridging the gap between scholarship and practice. Routledge, London. <https://doi.org/10.4324/9780203810200>
33. McCue C, Parker A (2003) Connecting the dots: data mining and predictive analytics in law enforcement and intelligence analysis. *Police Chief* 70(10):115–118,120,122
34. Morselli C (ed) (2014) *Crime and networks*. Routledge, New York. <https://doi.org/10.4324/9781315885018>
35. Morselli C, Giguère C, Petit K (2007) The efficiency/security trade-off in criminal networks. *Soc Netw* 29(1):143–153. <https://doi.org/10.1016/j.socnet.2006.05.001>
36. Morselli C, Petit K (2007) Law-enforcement disruption of a drug importation network. *Glob Crime* 8(2):109–130. <https://doi.org/10.1080/17440570701362208>
37. Natarajan M (2006) Understanding the structure of a large heroin distribution network: a quantitative analysis of qualitative data. *J Quant Criminol* 22(2):171–192. <https://doi.org/10.1007/s10940-006-9007-x>
38. Puvathingal BJ, Hantula DA (2012) Revisiting the psychology of intelligence analysis: from rational actors to adaptive thinkers. *Am Psychol* 67(3):199–210. <https://doi.org/10.1037/a0026640>
39. Ratcliffe JH (2008) *Intelligence-led policing*. Willan Publishing, Cullompton

40. Scott J (2017) Social network analysis, 4th edn. SAGE, London.
<https://doi.org/10.4135/9781529716597>

41. Sheptycki JWE (2002) In search of transnational policing: towards a sociology of global policing. Ashgate, Aldershot.

42. Sparrow MK (1991) The application of network analysis to criminal intelligence: an assessment of the prospects. Soc Netw 13(3):251–274. [https://doi.org/10.1016/0378-8733\(91\)90008-H](https://doi.org/10.1016/0378-8733(91)90008-H)

43. Tecuci G, Schum DA, Marcu D, Boicu M (2016) Intelligence analysis as discovery of evidence, hypotheses, and arguments: connecting the dots. Cambridge University Press, Cambridge.
<https://doi.org/10.1017/CBO9781316388488>

44. US Senate Committee on the Judiciary (1983) Organized crime in America: hearings before the Committee on the Judiciary, United States Senate, Ninety-eighth Congress, first session, on organized crime in America, May 20 and July 11, 1983. US Government Printing Office, Washington, DC

45. van der Hulst RC (2009) Introduction to social network analysis (SNA) as an investigative tool. Trends Organ Crime 12(2):101–121. <https://doi.org/10.1007/s12117-008-9057-6>

46. Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511815478>

47. Windle J, Silke A (2019) Is drawing from the state ‘state of the art’? A review of organised crime research data collection and analysis, 2004–2018. Trends Organ Crime 22(4):394–413.
<https://doi.org/10.1007/s12117-018-9356-5>

Figures

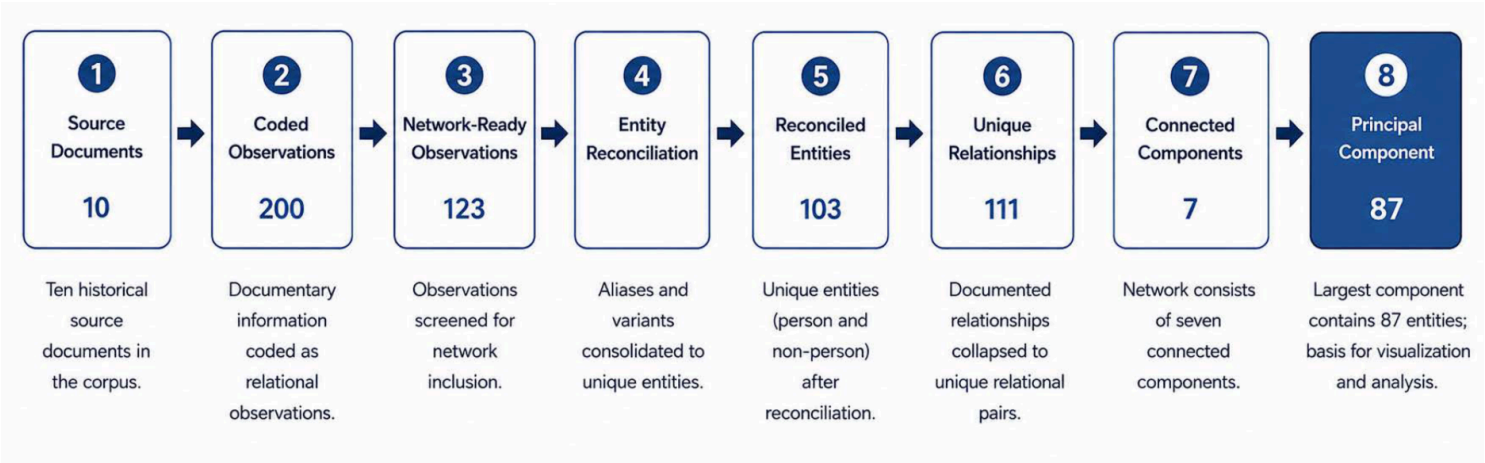


Figure 1

Methodological Progression from Historical Intelligence Records to Network Reconstruction

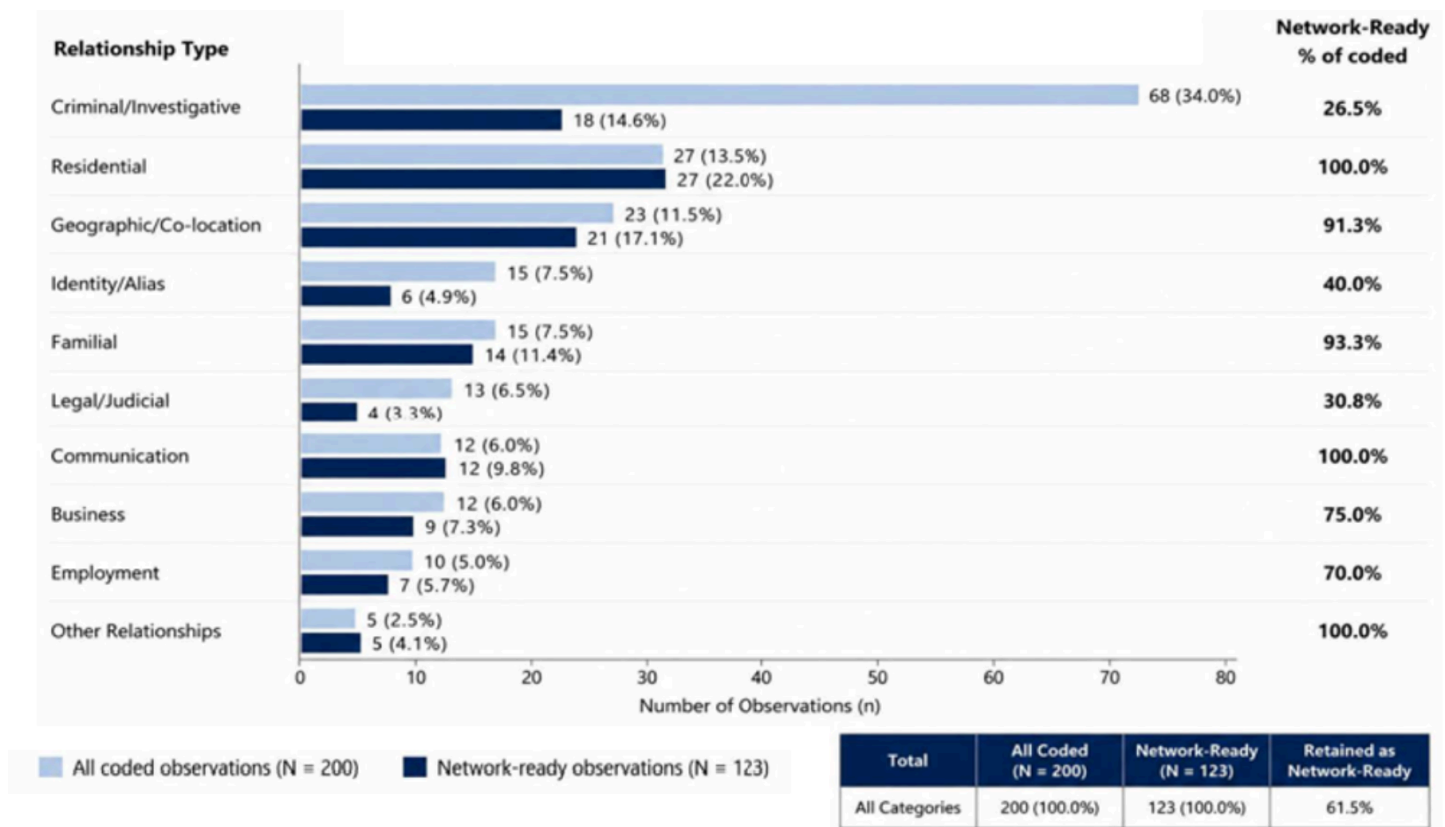


Figure 2

Figure 1. Comparison of Coded and Network-Ready Relationships by Category

Note: Light bars represent all coded observations (N = 200), and dark bars represent observations designated for network inclusion (N = 123). Percentages within each series describe the composition of the respective dataset; retention percentages indicate the proportion of coded observations within each category retained as network-ready relationships

