Ringtail
E-DISCOVERY SOFTWARE

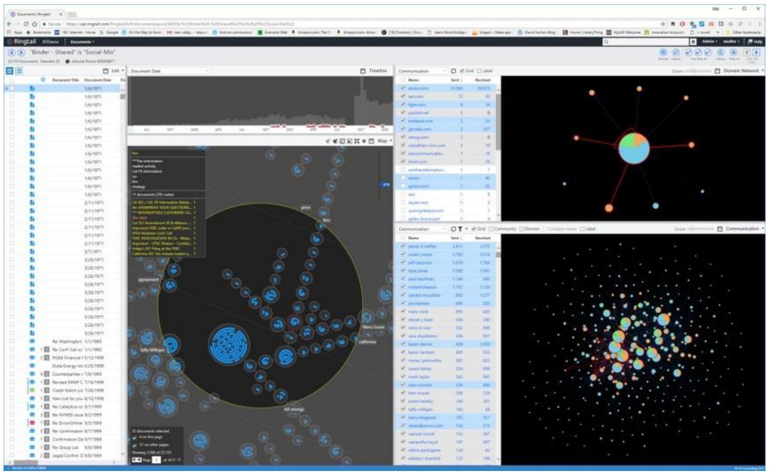
August 14, 2017

Ringtail Analytics – Investigation Use Case

Analytics is the systematic computational analysis of data. Data is comprised of two principle types: structured and unstructured. Unstructured data are typically documents while structured data can include items like phone records, wire transfers and other financial information. Some data types are a combination of the two such as emails and chat communications where the message is unstructured, but the address and time information is structured. For unstructured data, Ringtail uses semantic analysis to provide structure so that visualization of obscured relationships is possible. Ringtail uses network analytics to visualize relationships among people (email, phone logs, and SMS texts), organizations and financial transactions.

Importantly, Ringtail Analytics is not a separate module or feature set. Performing analytics in Ringtail does not require data migration or special handling. Ringtail Analytics isn't just integrated into the platform, it is the platform.

In addition to organizing and visualizing unstructured information (text, video, audio, images) at the meta (cubes, mines), macro (document map, people map, financial transaction map) and micro (document cluster, people communities) level, Ringtail Analytics includes predictive coding and continuous active learning technologies to surface relevant information. While Ringtail provides extensive search, document viewing, and robust coding for linear review for traditional litigation, Ringtail analytics and predictive coding excel at early case assessments, investigations and enforcement actions where you are not sure what you are looking for at the outset.



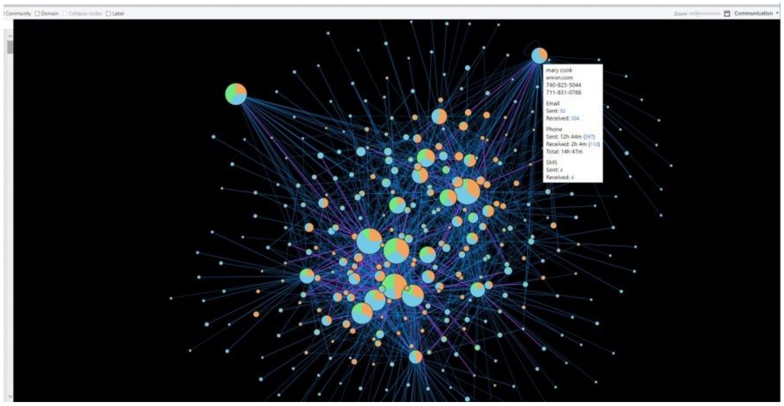
To illustrate the use of analytics for an investigation, this screen shot illustrates an example in which Ringtail configures nine panes within a workspace (five visible in this screen shot): Document List view, Timeline view, Map view, Domain Network view, and People Communication Network view. Another browser window would have Browse view, Document view, People view, and Coding view. Suspecting that the investigation is about inflating revenues, the investigator starts with the Domain Network view (upper right) and looks for external organizations with which the target organization is communicating. Noticing an abnormal volume of emails with a foreign distributor, she clicks on the connection between the two organizations to see who is communicating. Going to the People Communication view (lower right), she sees that the target's controller has many emails, phone calls and SMS texts with the CEO of the distributor. Wondering what concepts they were communicating about, she clicks on the link between the two custodians, and goes to the Map view to see how the emails cluster conceptually. She sees a promising cluster that includes the concepts revenue and recognition. She selects that document cluster and notices that the highlights on the Timeline view are within a two-week period right before the close of the quarter. She then investigates a few emails (in the Browse, View, Code panes) to confirm that the two individuals were talking about manipulating the revenue numbers.

Suspecting that she might also see the potential fraud reflected in the wire transfers that were ingested with the email and document information, she goes to the Financial Transactions view and sees that there are many transactions between dummy corporations.

In a matter of minutes, the investigator has found a high likelihood of fraud. To confirm the fraud, she marks the highly suspect emails as relevant. Then she uses Continuous Active Learning ("CAL") to find the documents in the Ringtail repository that are most like those marked documents. As she identifies the emails and attachments that are relevant, the predictive model gets better at providing the next most relevant documents. Within an hour she has identified the type of fraud and the relevant documents and is prepared to go to the next phase of the investigation.

1.1.1. Social Network

While documents are an important source of information for litigation, investigations start by understanding who is talking with whom about particular topics. The structured metadata of emails (From, To, CC, BCC, Date, Time) provide the organizing information to show three types of communication relationships – people relationships, organization (domain) relationships, and people to organization relationships.



mary cook
enron.com
740-825-5044
711-831-0788

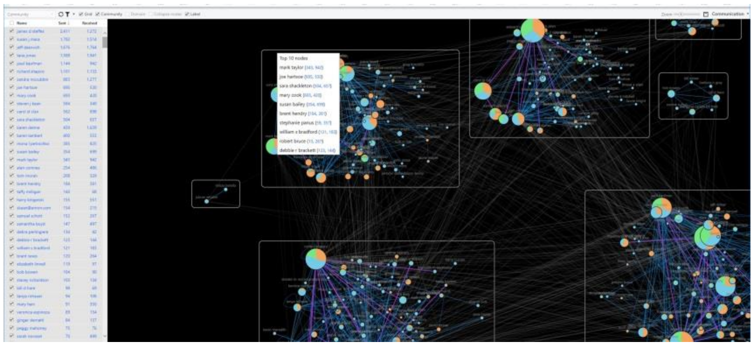
Email
Sent: 92
Received: 304

Phone
Sent: 12h 44m (597)
Received: 2h 4m (112)
Total: 14h 47m

SMS
Sent: 4
Received: 4

In the People Communication Network visualization, the size of a circular node represents how many emails, phone calls, and SMS texts are sent and received by an individual while the connecting lines (known as arcs) represent the flow of communication between two people. The nodes are color coded by type of communication (orange is email, blue is phone, and green is SMS text). Arcs are color coded by how many forms of communication are used between two people (nodes).

By exploring these nodes and arcs, an investigator can see who is communicating with whom. Clicking on a node or an arc will select the represented documents, calls, or texts and cause them to be highlighted in other visible panes (such as the document view, the Map view, and the Timeline view). Ringtail's referential integrity updates all panes based on a selection in any one pane that allows the user to see when communication between two people were exchanged (highlights in Timeline view) and what concepts were discussed (highlights in Map view). Specific emails or attachments are highlighted in the document view.



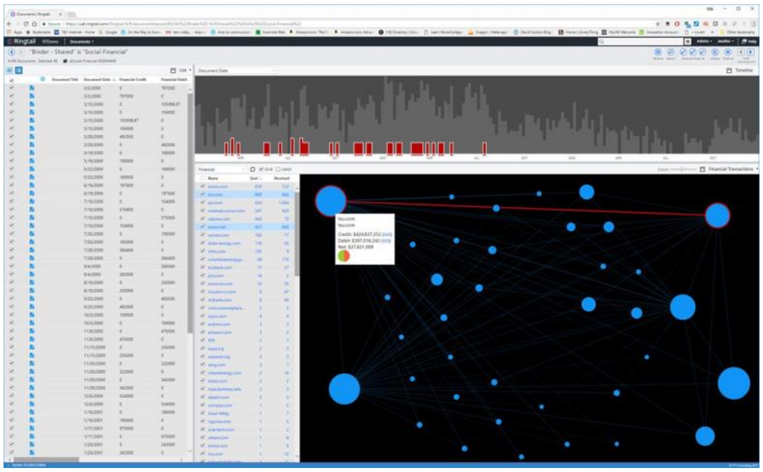
The social network community option pictured here provides an automated way to identify the informal networks of people that communicate often. Each group shows up in a box with the top ten participants in a community group identified.

In the Domain Network visualization, all the organizational entities or domains (the text that is to the right of the "@" in an email address) in the selected emails are represented. An organizational boundary is portrayed as a rectangle. The arcs (lines) between organizations represent the flow of emails between any two domains. By clicking on an organization or an arc, the items in the other panes are highlighted (people in the people relationships view, when in time on the Timeline view, and the documents and concepts in the Map view). By double clicking on an organization, the rectangle expands to show the people that are inside that organization (people to organization relationships).

In a future version of Ringtail, identity normalization (gathering together all the email addresses, phone numbers, SMS texts, and chat aliases associated with a single individual) will expand the visualization of the multiple media that individuals communicate through. The identity normalization will be both automated and allow for manual connections.

1.1.2. Financial Transaction Networks

In addition to document requests, investigators have access to financial information like wire transfers, accounts payable, accounts receivable, and stock transactions. The general form of a financial transaction is type (wire transfer, stock transaction, general ledger entry…), from (debit), to (credit), amount, and then additional fields. The financial transaction resembles the structure of email header information. Using the general form of a financial transaction combined with indicators of types of fraud, Ringtail uses the underlying network analytics to visualize the flows of money between organizational entities.



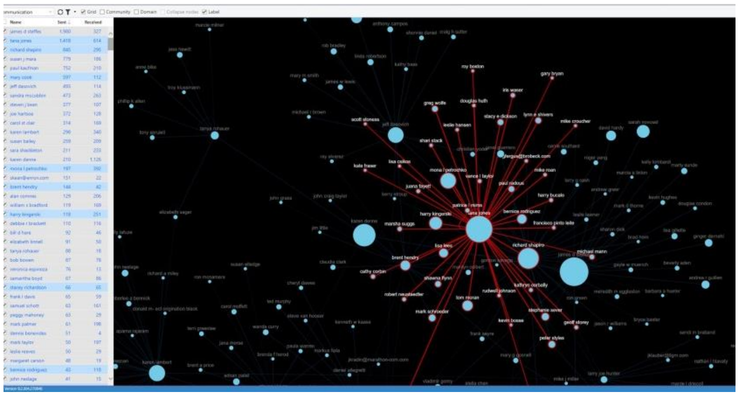
The initial Ringtail financial transaction tool looks for a common form of revenue enhancement fraud (also known as round tripping fraud). An investigator suspects that a company is trying to inflate revenues to increase stock prices by running sham transactions through dummy corporate entities. The indicators of this type of fraud are symmetric wire transfers, and two-way transactions through a same day return of a wire transfer. Ringtail shows the dollar flows and their directions. The larger nodes represent the entities with a high number of transactions and are the suspect dummy corporations.

The investigator can see the size and frequency of financial transactions by rolling the cursor over nodes or arcs (connecting lines). The pie chart on the rollover information give you a quick indicator of whether the transaction flow is symmetric or asymmetric (equal flows or unequal flows). By clicking on a node or arc, the detailed transactions are displayed in the document view pane and highlighted in the timeline pane. In the left pane, you can see that most of the transactions are symmetric (go to and from a company) and occur on the same day. These two transaction relationships are indicators of revenue enhancement fraud.

In future releases, additional financial transaction network visualizations will be added based on demand from the regulatory agencies, financial auditors and enterprise customers.

1.1.3. Phone Communication Network Analytics

Phone logs are another form of tracking social networks. Using either Call Data Recorder (CDR) phone logs or mobile phone data collections, phone calls and SMS text chats are analyzed like email information to represent the relationships between people and organizations. Using the length of calls and the directionality of calls, the investigator can identify suspect pairs who are in close relationships. The same communication visualization is used for emails, phone and chat. In this screen shot, email and SMS texts are turned off and only the phone information is shown. Where length of call information is available those call lengths are summed when rolling over nodes and arcs.



Name	Sent	Received
james d steffes	1,980	327
tana jones	1,418	614
richard shapiro	845	296
susan j mara	779	186
paul kaufman	752	210
mary cook	597	112
jeff dasovich	493	114
sandra mccubbin	473	263
steven j kean	377	107
joe hartsoe	372	128
carol st clair	314	169
karen lambert	296	340
susan bailey	259	209
sara shackleton	211	223
karen denne	210	1,126
mona l petrochko	197	392
skean@enron.com	151	22
brent hendry	144	42
alan comnes	129	206
william s bradford	119	169
harry kingerski	118	251
debbie r brackett	110	116
bill d hare	92	46
elizabeth linnell	91	58
tanya rohauer	88	18
bob bowen	87	78
veronica espinoza	76	13
samantha boyd	67	86
stacey richardson	66	65
frank l davis	65	59
samuel schott	63	161
peggy mahoney	63	29
mark palmer	61	198
dennis benevides	51	4
mark taylor	50	197
leslie reeves	50	29
margaret carson	48	19
bernice rodriguez	43	118
john neslage	41	15

In a future version of Ringtail, the different identities of individuals (nicknames, email addresses, phone numbers, SMS texts, chat aliases) will be found automatically where possible. With mobile phone information that has geographic location, a map of where calls are placed and received will be linked to the social network of callers.

1.1.4. Future Direction of Ringtail and Impact on SEC Requirements

Today Ringtail is primarily focused on the capture, analysis, and production of unstructured documents and financial transaction data. In the near term (2017/2018), Ringtail will add additional unstructured media types like audio, video and image analytics. Our analytics will process each of these unstructured data types to support traditional eDiscovery workflows and to support "Event 360" views for investigations (regulatory agencies, financial services, FOIA requests). Event 360 recognizes the importance of pulling together the following types of networked information – **semantic networks** (how documents and other media types are related in terms of their semantic content), **social networks** (how people and organizations are related through their exchange of messages in email, phone logs, Bloomberg chat, audio and video recordings), **event networks** (how events are related in the two dimensions of time – succession and intention), an extensive set of **financial transaction networks** (financial transactions have the same format as emails – date, time, from, to, and content – and can be related to and analyzed in a similar way to Ringtail's networked data analytics), and **geographic location networks**.