THE XPRIORI REPORT

Forensically Sound Collections for Investigation, Information Governance and eDiscovery

In any investigation, information governance or eDiscovery activity, you need a "forensically sound" collection of electronic data (email, documents, drawings, etc.) and any other hard copy paper items – a collection that ensures their authenticity and so that they can be used in an evidentiary fashion if ultimately required. Here is the Xpriori approach:

What is "forensically sound" collection?

It is the extraction of electronic data by processes that ensure that the resulting information collected has not been altered in form or substance and that the surrounding information as to ownership, authorship, source location, times and dates are captured and stored in auditable and traceable fashion. In most cases this means that there are **two objectives that must be met**:

1. The acquisition and subsequent analysis has been undertaken with due regard to the preservation of the data in the state in which it was first discovered. And, this must be established with a rock solid audit trail;
2. Forensic data collection requires a reliable and demonstrably accurate process that is capable of being tested and verified by independent third parties and allows for admissibility in court or other proceedings should the need arise.

What are the requirements for collection in this case?

In most circumstances, you do not require exact copies of entire hard drives – i.e. old notion that only copies of each bit was required. You simply need to collect certain user created file folders and email files in order to demonstrate that the collection of data is as complete as possible. You need to assure any risk of "cherry picking" of information is obviated. By "cherry picking", we mean the collection of certain date to the exclusion of other data. The collection also needs to be of the documents in their original format.

In the case of government contracts, there can be regulatory inquiry that could lead to criminal as well as civil investigation, so in those instances, we propose a heightened standard of care in collection. Regulatory investigators have sweeping powers to request information of all kinds even remotely related to the matter at hand. Therefore, a broad and comprehensive scope of collection -- leaning to a collect all approach – is warranted**.** This also means that deletions of emails and other information can give rise to suspicion and, therefore, everything should be kept and accounted for.

What are the recommended procedures?

1. To assure a complete collection, all owners/custodians of data and potential places of storage of said data must be identified by interviewing IT/email administrators as well as the owners/custodians of data themselves.
2. The owners/custodians should be instructed not to delete or destroy any information.
3. All normal data purging or automated deletion functionality should be suspended.
4. A third party collection expert should do the collection to ensure consistency and efficiency of collection while minimizing operational impact on day to day activities – the man hours associated with this is greatly reduced and constrained using this approach. The tools the collection expert uses to collect the data will ensure that there is no alteration of any time or date stamps or ownership information thereby maintaining and establishing defensibility/evidentiary value should the need arise.
5. The data must be maintained in storage that will not allow for spoliation; preserved in pristine state, and be accompanied by detailed audit logs of all collection actions.
    a. To avoid spoliation, it is useful to store as read only, not allowing for any changes;
    b. Pristine condition requires:
        i. Data not be changed in the collection process;
        ii. File condition and metadata must be maintained as it was pre-collection;
        iii. File and folder structure must be maintained as it was on the original source
    c. Detailed audit logs must disclose:
        i. What was collected and not collected and why;
        ii. Who collected the data and when;
        iii. Details with respect to the source from which the information was collected; and
        iv. Hash values for all files collected.
6. Data should also be stored to an organizational taxonomy that can be configured to receive newly created or received information.

Who should do the collection?

It should be done by a team trained in the process and given access to the participants in the project by the highest authority in the customer. We have such a team. Once we have reviewed the legacy or existing information, to assure continuing preservation, we can help set up procedures that largely automate collection of the information from ongoing process. These procedures can help assure its classification to appropriate places under the storage taxonomy. This can be done with minimal human guidance.