

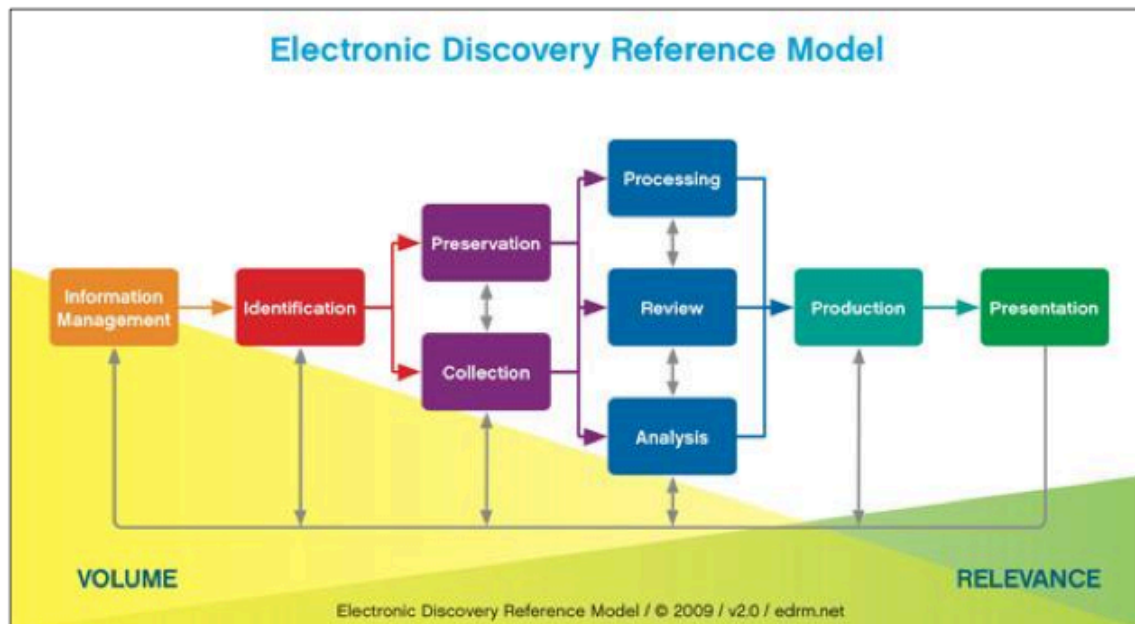
## Deprendo Technologies Product Scope Meeting

10/10/14

### Agenda:

- Introduce X-Lab Team
  - Define what segments of the EDRM model do we want to encompass in our first product
  - Begin preliminary ideal maximum functionality list per each process of the EDRM model within defined scope
  - Identify preliminary areas that are going to require further research and exploration
  - Begin preliminary list of ideal maximum types of ESI and locations to process
  - Identify Deprendo needs outside the actual eDiscovery software. (website, brand, etc.)
  - Gather initial thoughts on how we would want to integrate TeraCrunch from a high level system perspective
- 

### EDRM Model – standard eDiscovery process



**Information Management:** Getting your electronic house in order to mitigate risk & expenses should e-discovery become an issue, from initial creation of electronically stored information through its final disposition.

**Identification:** Locating potential sources of ESI & determining its scope, breadth & depth.

**Preservation:** Ensuring that ESI is protected against inappropriate alteration or destruction.

**Collection:** Gathering ESI for further use in the e-discovery process (processing, review, etc.).

**Processing:** Reducing the volume of ESI and converting it, if necessary, to forms more suitable for review & analysis.

**Review:** Evaluating ESI for relevance & privilege.

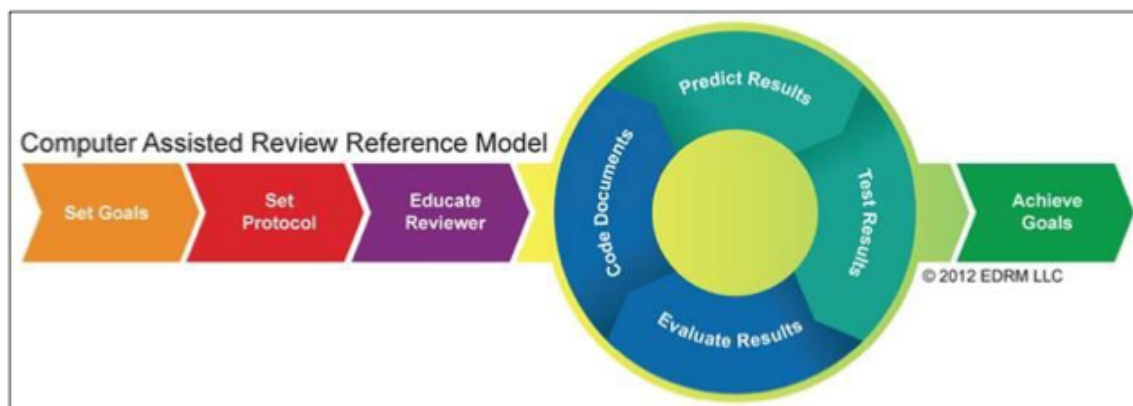
**Analysis:** Evaluating ESI for content & context, including key patterns, topics, people & discussion.

**Production:** Delivering ESI to others in appropriate forms & using appropriate delivery mechanisms.

**Presentation:** Displaying ESI before audiences (at depositions, hearings, trials, etc.), especially in native & near-native forms, to elicit further information, validate existing facts or positions, or persuade an audience.

---

The CARRM Model is the current process of *Predictive Coding* implementation



**Set Goals:** The process of deciding the outcome of the Computer Assisted Review process for a specific case. Some of the outcomes may be:

- Reduction and culling of not-relevant documents;
- Prioritization of the most substantive documents; and
- Quality control of the human reviewers.

**Set Protocol:** The process of building the human coding rules that take into account the use of CAR technology. CAR technology must be taught about the document collection by having the human reviewers submit documents to be used as examples of a particular category, e.g. Relevant documents. Creating a coding protocol that can properly incorporate the fact pattern of the case and the training requirements of the CAR system takes place at this stage. An example of a protocol determination is to decide how to treat the coding of family documents during the CAR training process.

**Educate Reviewer:** The process of transferring the review protocol information to the human reviewers prior to the start of the CAR Review.

**Code Documents:** The process of human reviewers applying subjective coding decisions to documents in an effort to adequately train the CAR system to “understand” the boundaries of a category, e.g. Relevancy.

**Predict Results:** The process of the CAR system applying the information “learned” from the human reviewers and classifying a selected document corpus with pre-determined labels.

**Test Results:** The process of human reviewers using a validation process, typically statistical sampling, in an effort to create a meaningful metric of CAR performance. The metrics can take many forms, they may include estimates in defect counts in the classified population, or use information retrieval metrics like Precision, Recall and F1.

**Evaluate Results:** The process of the review team deciding if the CAR system has achieved the goals of anticipated by the review team.

**Achieve Goals:** The process of ending the CAR workflow and moving to the next phase in the review lifecycle, e.g. Privilege Review.

---

## High Level Notes from “eDisclosure Systems Buyers Guide 2013 Edition”:

### **Whats Hot:**

- Ability to process and analysis audio files with the same sophistication as any other ESI
- Ability to collect and process data from social networks
- Ability to “just read” small amounts of ESI
- Ability to redact documents in their native form

### **Problems currently in the industry:**

- Treating things like email families as a single entity which prevents the unitization of things like attachments for relevancy processing
- Issues working in native formats. See 4.5.5 of eDisclosureSystemsBuyersGuideV1\_2.pdf pg 20

## High Level Notes from ACEDS webinar “Lawyer Competency in the age of E-Discovery”:

- Where is the industry going? – Predictive coding, managed and taught by “Key” lawyer in case
  - Need “information” based model in modern eDiscovery as opposed to the prior art based on a “document” model
  - The protection of privileged information and client communication. Particularly with 3<sup>rd</sup> party hosting of data removing privileged status
  - 1/3 of what lawyers manually characterize as relevant is irrelevant and vise versa
  - Judges are really hungry for the implementation of Predictive Coding but lawyers are afraid of it. Our training resources and user experience are going to be extremely important
  - DO NOT keyword cull documents prior to predictive coding as the manual assumptions in the keyword cull can cripple the data set prior to predictive coding getting a chance
  - High quality keyword search is currently easier to defend because of the very known and familiar process to replicate results and define search parameters as opposed to the fuzz realm of predictive coding. How can we combat this?
-

**Suggested Follow Up:**

READ(AC, HB): "Ten things that trouble judges about e-discovery" by Craig Ball 2010

READ(AC, HB): Brown v. Tellermate Holdings LTD. Case No. 2:11-cv-1122 2014

READ(HB): CTRL Guidelines

OUTLINE(X-Lab): Project processes and practices (PM tool, security, source control, documentation, etc.)