Seven Steps to Client Success
Understanding the Flow of eDiscovery

ORANGE LEGAL TECHNOLOGIES

Seven Steps to Client Success

An Overview of Electronic Discovery

- Understanding Electronic Discovery

  Definitions and Descriptions

- Understanding the Flow of Electronic Discovery

  Tasks and Techniques

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Electronically Stored Information (ESI)?

"…all information on computers."

While not specifically defined in the Federal Rules of Civil Procedure, electronically stored information, or ESI, is defined in the November 2006 issue of *The Third Branch* (Newsletter of the Federal Courts) simply as "all information on computers".

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Data?

"…info suitable for processing."

In the world of computer science, data is "numerical or other information represented in form suitable for processing by computer".

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Description

- How is Data Structured?

"…85% of enterprise data is unstructured."

Unstructured data refers to ESI that does not have an assigned format and significance. Examples of "unstructured data" may include audio, video and unstructured text such as the body of an email or word processor documents.

Structured data refers to ESI in which every bit of information has an assigned format and significance. Examples of "structured data" may include databases such as QuickBooks, SQL or Access.

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Description

- What are the typical Data Formats?

"…a format is a particular way to encode information for storage in a computer file."

Textual | Still Images | Moving Images | Sound | Web Archive | Generic

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Description

- How do these typical Data Formats present themselves?

Textual | .doc .pdf .txt .wpd .xls .ppt .html

Still Images | .bmp .gif .jpg .tiff

Moving Images | .avi .flv .mov .mpeg .swf .wmv

Sound | .au .mp3 .mp4 .ra .wav .wma

Web Archive | .arc .mhtml .warc

Generic | UTF-8 (Unicode)

ORANGE LEGAL TECHNOLOGIES

Definition

- What is Metadata?

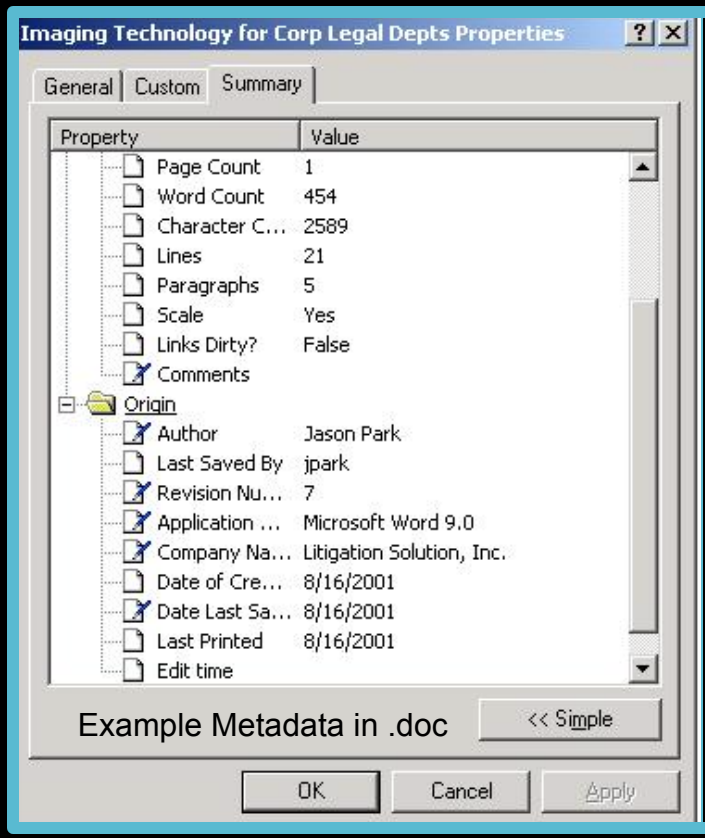
"…used to describe data or information."

Metadata are used to describe data or information. Metadata can describe just about anything you find on a computer, and the term is often used to refer to information that is not typically visible to document users.

Understanding Electronic Discovery

Description

- How does Metadata from these Data Formats present itself?



Example Metadata Fields in a Standard Project

BegDoc#	Comments
EndDoc#	Categories
BegAttach	Company
EndAttach	DocExt
PgCount	Email_Subject
Custodian	From
Folder	To
Filename	Cc
DateCreated	Bcc
DateLastMod	DateSent
Title	TimeSent
Author	DateRcvd
Subject	TimeRcvd
Keywords	Attach

Understanding Electronic Discovery

Description

- Where might Data be found?

"On any device that can store ESI."

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Electronic Discovery?

"…finding, preparing, and sharing ESI for use in the legal process."

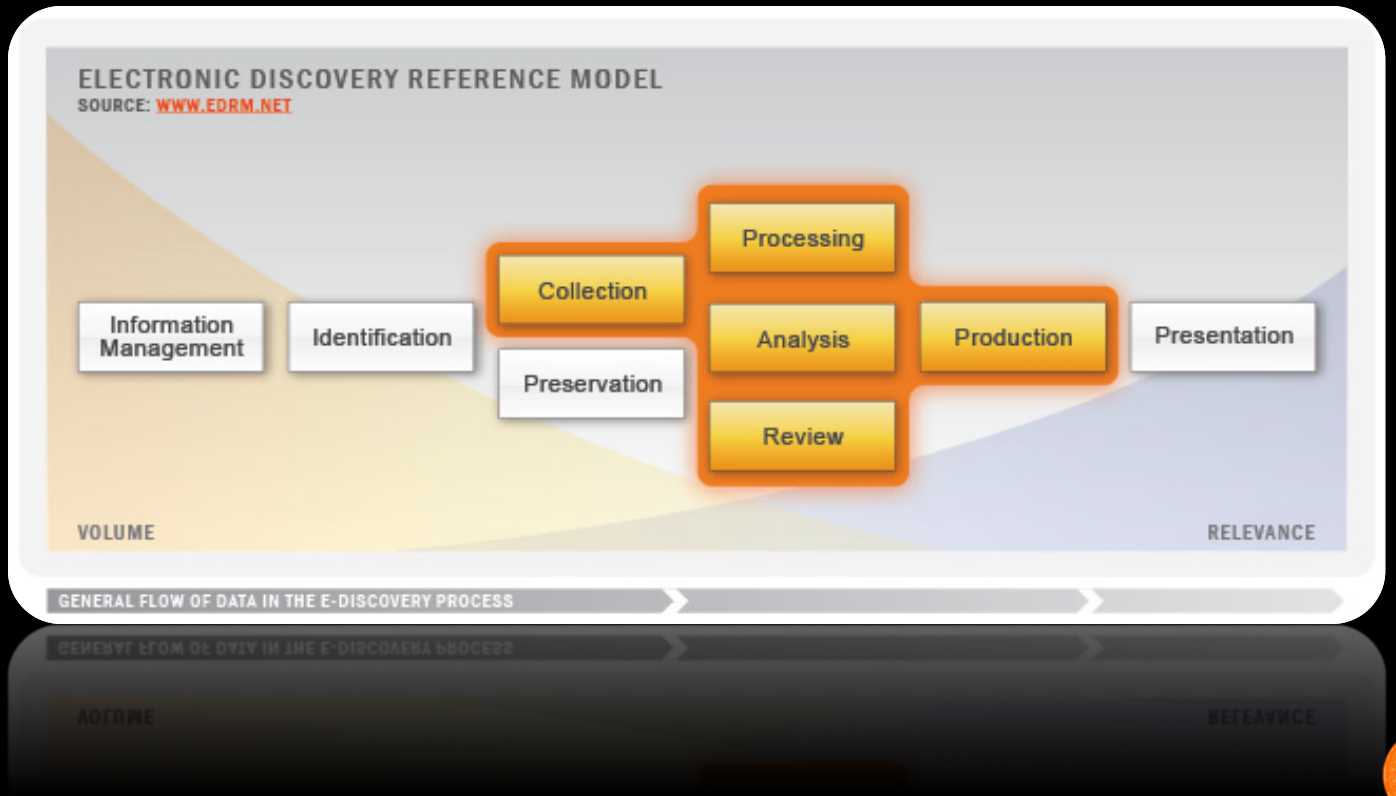
Electronic Discovery is the process of collecting (also called "harvesting"), preparing, reviewing, and producing electronic documents in the context of the legal process.

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is the Electronic Discovery Reference Model (EDRM)?



ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Collection?

"…the acquisition of ESI in eDiscovery."

The collection of data can be simply defined as the acquisition of potentially relevant electronically stored information (ESI) in the conduct of electronic discovery.

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is a "Forensically Sound" Data Collection?

"…the acquisition of ESI in a manner that does not materially alter the source evidence…"

The forensically sound collection of data does not materially alter the source evidence except to the minimum extent necessary to obtain the data. The manner used to obtain the evidence must be documented.

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Data Ingestion?

"…the process of introducing data into an eDiscovery platform."

The ingestion of data is simply the process of introducing data from proprietary industry formats and/or standard data transfer protocols into an eDiscovery platform to allow for the execution of eDiscovery tasks.

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Analytics?

"…early organization, understanding, and prioritization of information."

Analytics in the realm of electronic discovery is the leveraging of data through the use of particular functional processes to include locating, indexing, and searching data to enable early case assessment.

ORANGE LEGAL TECHNOLOGIES

Understanding Electronic Discovery

Definition

- What is Processing?

"…processing prepares files for review, production, and subsequent use."

"Processing" is any operation or set of operations which is performed upon data, whether or not by automatic means, such as Chain of Custody Security and Tracking, Data Staging, Data Filtering, De-duplication, Metadata Extracting, Full Text Extractions, Exception Handling, Data Conversion, and Load File Production.

ORANGE LEGAL TECHNOLOGIES

Definition

- What is Review?

"…the objective of review is to determine file relevancy and/or responsiveness."

"Review" can be defined as the culling process that produces a dataset of potentially responsive documents that are then examined and evaluated for a final selection of relevant and/or responsive documents and assertion of privilege, confidentiality, etc., as appropriate.

Understanding Electronic Discovery

Definition

- What is Production?

"…the preparation and delivery of data based on legal request."

As defined by The Sedona Conference, "Production" is "the process of delivering to another party, or making available for that party's review, documents and/or ESI deemed responsive to a discovery request". In even simpler terms, "Production" can be understood as the "delivery of data or information in response to an interrogatory, subpoena or discovery order or a similar legal process."

ORANGE LEGAL TECHNOLOGIES

Seven Steps to Client Success

An Overview of Electronic Discovery

- Understanding Electronic Discovery

  Definitions and Descriptions

- Understanding the Flow of Electronic Discovery

  Tasks and Techniques

ORANGE LEGAL TECHNOLOGIES

Understanding the Flow of Electronic Discovery

Seven Key Steps to eDiscovery

- Data Inventory
- Data Collection
- Data Ingestion
- Data Analytics
- Data Processing
- Data Review
- Data Production

Electronic Discovery is the process of collecting, preparing, reviewing, and producing electronic documents in the context of the legal process.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

Questions to Consider in Inventorying ESI

- What is the scope of applicable data?
- What is the structure of applicable data?
- What is the format of applicable data?
- What is the state of applicable data?
- How is client data interconnected?
- How does the client access active data?
- How does the client maintain static data?
- How much client data will be acted upon?
- Is client data encrypted?

As part of initiating a project, a full inventory of potential client ESI is undertaken using the aforementioned framework with the goal of identifying appropriate data from client data stores for the forensically sound conduct of electronic discovery.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

What is the scope of applicable data?

- Entity Scope - Entities that may have had individuals involved in the creation, review, and/or response of data that may contain relevant info for the matter at hand.

- Custodian Scope - Individuals who may have been involved in the creation, review, and/or response of data that may contain relevant information for the matter at hand.

- Data Steward Scope - Individuals who have Information Technology management responsibilities for the entities and individuals determined to be relevant to the matter at hand and/or individuals who maintain access rights to the applications and equipment used by these entities and organizations.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

What is the scope of applicable data? (Continued)

- Geographical Scope - The geographical locales of the entities and individuals that may have been involved in the creation, review, and/or response of communications and/or documents relevant to the matter at hand as well as the locales of the equipment used to support creation, transmission, review, and storage of these communications and/or documents.

- Time Frame Scope - The period of time in which relevant information may have been created, reviewed, and/or responded to for the matter at hand.

- Volume Scope - The estimated volume of data that may contain relevant information for the matter at hand.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

What is the structure of applicable data?

Unstructured data refers to ESI that does not have an assigned format and significance. Examples of "unstructured data" may include audio, video and unstructured text such as the body of an email or word processor documents.

Structured data refers to ESI in which every bit of information has an assigned format and significance. Examples of "structured data" may include databases such as QuickBooks, SQL or Access.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

What is the format of applicable data?

- Textual - Content works consisting primarily of text.

- Still Images - Images that convey their meaning in visual terms, e.g. pictorial images, photographs, posters, graphs, diagrams, documentary architectural drawings. Formats for such images may be bitmapped (sometimes called raster), vector, or some combination of the two. A bitmapped image is an array of dots (usually called pixels, from picture elements, when referring to screen display), the type of image produced by a digital camera or a scanner. Vector images are made up of scalable objects—lines, curves, and shapes—defined in mathematical terms, often with typographic insertions.

- Moving Images - A variety of media-independent digital moving image formats and their implementations.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

What is the format of applicable data? (Continued)

- Sound - Media-independent sound content that can be broken into two format sub-categories. The first sub-category consists of formats that represent recorded sound, often called waveform sound. The second sub-category consists of formats that provide data to support dynamic construction of sound through combinations of software and hardware.

- Web Archive - Content in formats that hold the results of a crawl of a Web site or set of Web sites and captures them in the form disseminated to users.

- Generic - Content in widely acceptable generic formats to include but not limited to specifications for wrappers (e.g., RIFF and ISO_BMFF), bundling formats (e.g., METS and AES-31), and encodings (e.g., UTF-8 and IEEE 754-1985).

ORANGE LEGAL TECHNOLOGIES

Data Inventory

What is the state of applicable data?

- Active Data resides on the hard drives or optical drives of computer systems, is readily visible to the operating system and/or application software with which it was created and is immediately accessible to users without deletion, modification or reconstruction.

- Static Data (or Archival Data) is info that is not directly accessible to the user of a computer system but that the organization maintains for long-term storage and record keeping purposes. Static data may be written to removable media such as a CD, magneto-optical media, tape or other electronic storage device, or may be maintained on system hard drives in compressed formats.

- Residual Data is not active on a computer system. Residual data includes (1) data found on media free space; (2) data found in file slack space; and (3) data within files that has functionally been deleted in that it is not visible using the application with which the file was created, without use of undelete or special data recovery techniques.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

How is client data interconnected?

- Non-Networked - Data is not interconnected to a group of computers.

- Personal Area Network (PAN) - A computer network used for communication among computer devices close to one person. Some examples of devices that may be used in a PAN are printers, fax machines, telephones, PDAs, or scanners.

- Local Area Network (LAN) - A network covering a small geographic area, like a home, office, or building. Current LANs are most likely to be based on Ethernet technology.

- Campus Area Network (CAN) - A network that connects two or more LANs but that is limited to a specific and contiguous geographical area such as a college campus, industrial complex, or a military base.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

How is client data interconnected? (Continued)

- Metro Area Network (MAN) – A network that connects two or more Local Area Networks or Campus Area Networks together but does not extend beyond the boundaries of the immediate town, city, or metropolitan area.

- Wide Area Network (WAN) - A WAN is a data communications network that covers a relatively broad geographic area (i.e. one city to another and one country to another country) and that often uses transmission facilities provided by common carriers, such as telephone companies.

- InterNetwork (Intranets, Extranets, Internet) - Two or more networks or network segments connected using devices that operate at layer 3 (the 'network' layer) of the OSI Basic Reference Model. Any interconnection among or between public, private, commercial, industrial, or governmental networks may also be defined as an internetwork.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

How does the client access active state data?

- Direct Attached Storage (DAS) - Direct-attached storage (DAS) refers to a digital storage system directly attached to a server or workstation, without a storage network in between. It is a retronym, mainly used to differentiate non-networked storage from SAN and NAS.

- Network Attached Storage (NAS) - Network Attached Storage (NAS) is a file-level computer data storage connected to a computer network providing data access to heterogeneous network clients.

- Storage Area Network (SAN) - A storage area network (SAN) is an architecture to attach remote computer storage devices (such as disk arrays, tape libraries and optical jukeboxes) to servers in such a way that, to the operating system, the devices appear as locally attached.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

How does the client maintain static state data?

- Semi Conductor Based Storage Media
  Memory Cards, USB Flash Drives, PDAs, Digital Audio Players, Digital Cameras, Mobile Phones, Copiers

- Magnetic Based Storage Media
  Floppy Disk, Hard Disk, Magnetic Tape

- Optical and Magneto Optical Storage Media
  CD, CD-ROM, DVD, BD-R, BL-RE, HD DVD, CD-R, DVD-R, DVD+R, CD-RW, DVD-RW, DVD+RW, DVD-RAM, UDO

ORANGE LEGAL TECHNOLOGIES

Data Inventory

How much client data will be acted upon?

- Compressed Data
  Data not having undergone a process of transformation from one representation to another, smaller representation from which the original, or a close approximation to it, can be recovered.

- Uncompressed Data
  Data having undergone a process of transformation from one representation to another, smaller representation from which the original, or a close approximation to it, can be recovered. Typically determined by Algorithm Complexity and Amount Of Compression.

ORANGE LEGAL TECHNOLOGIES

Data Inventory

Is the client data encrypted?

- Encrypted Data
  Data having undergone a procedure that renders the contents of a computer message or file unintelligible to anyone not authorized to read it. The data is encoded mathematically with a string of characters called a data encryption key.

- Unencrypted Data
  Data not having undergone a procedure that renders the contents of a computer message or file unintelligible to anyone not authorized to read it. The data is encoded mathematically with a string of characters called a data encryption key.

ORANGE LEGAL TECHNOLOGIES

Data Collection

Sources to Consider in ESI Collections

- Active, Online Data (Typically Accessible)
- Nearline Data (Typically Accessible)
- Offline Storage (Sometimes Accessible, Sometimes Unreasonably Accessible)
- Backup Tapes (Typically Unreasonably Accessible)
- Erased, Fragmented, or Damaged Data (Typically Unreasonably Accessible)

For the purposes of time and cost considerations, ESI is typically categorized as either reasonably accessible or not reasonably accessible. Generally speaking, a party need not provide discovery of ESI from sources that the party identifies as not reasonably accessible because of undue burden or cost (FRCP Rule 26 (b)(2)(B).

ORANGE LEGAL TECHNOLOGIES

Data Collection

Key Capabilities for Complete Collections

- Fixed Storage Collection - Manual+Active Data Copy/Forensic Imaging
- Portable Storage Collection - Manual Copy/Forensic Imaging
- Back Up Tape Restoration
- Automated Network Discovery of Devices/Repositories and Data

ORANGE LEGAL TECHNOLOGIES

Data Ingestion

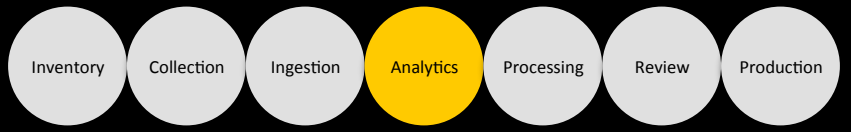
Stages to be Considered as part of Data Ingestion

- Copying and storage of original ESI files on closed and isolated network server.
- Storage of original media and ESI files in a forensically sound manner.
- Storage of copied ESI files for use in further eDiscovery processing.

Ingestion of client data into an eDiscovery platform typically begins with data staging. Data staging is the process by which original ESI files are copied, isolated, and stored in a forensically sound manner for future use.

ORANGE LEGAL TECHNOLOGIES

Data Analytics



Key Tasks Accomplished with Data Analytics

- Data Indexing provides a comprehensive index that includes full text and metadata attributes and can quickly be queried online to organize, understand, and assess available data.

- Data Reduction is accomplished through the combined used of culling and filtering technologies that provide system file, data range, extension, custodian, and key word filtering as well as the application of near duplicate identification.

- Data Understanding is facilitated with unique features to include interesting phrase finder and conversation thread linking technologies allowing for analysis of data within context of its use.

- Early Case/Data Assessment is the combined leveraging of the preparation, indexing, organization, and understanding of analytics to provide users with the ability to balance opportunities, risks, and costs in preparation for litigation, audits, and investigations.

ORANGE LEGAL TECHNOLOGIES

Data Processing

Key Tasks Accomplished with Data Processing

- Data Filtering provides the capability to filter data by date ranges, extensions, custodians, and key words as well as allows for system file filtering against the NIST database using the MD5 hashing algorithm.

- Data Deduplication is provided using the MD5 hashing standard and can be accomplished throughout processing at both the global and/or the document family group level.

- Metadata Extraction allows for the efficient capture of system, file, and field metadata for most unstructured data formats.

- Full Text Extraction is conducted and is augmented as required by streamlined exception handling procedures to support secondary extractions via OCR and print driver text recognition.

ORANGE LEGAL TECHNOLOGIES

Data Processing

Key Tasks Accomplished with Data Processing (Continued)

- Data Conversion allows for the full conversion of native file formats into high quality TIFF images and PDF documents while also supporting native file linking.

- Load File Preparation allows for the seamless production of standard output files based on XML, Pass Through, Image, Native, and Proprietary Database Load Formats to ensure ease of use with industry standard review tools.

- Custom Database Development allows for the proactive development of custom databases that enable the usage of non-standard review tools and technologies.

- Exception Reporting allows the client to generate reports that explain in detail all types of exceptions encountered, what causes them, why they occur, what can be done to remedy them, and what can be done to prevent them in the future. There are usually standard processes in place to re-process exceptions in place after they have been repaired, decrypted or otherwise remedied.

ORANGE LEGAL TECHNOLOGIES

Data Review

Primary Types of Data Review

- First Level - The primary purpose of first level document review typically to review documents and determine whether or not they're "responsive" or "non-responsive" as they pertain to a specific legal case or issue. In essence, first level document review forms part of the discovery phase of litigation.

- Second Level - Second Level Review, sometimes referred to as "Privilege Review", is one of the most critical and sensitive aspects of the document review process and usually involves the use of senior reviewers and/or review by the litigators actually involved in the matter under review.

ORANGE LEGAL TECHNOLOGIES

Data Review

Key Considerations for Data Review

- Develop a Center of Gravity for Creating Data Sets
  Data Grouped by Custodian, Chronology and/or Concept
- Develop a Methodology for Searching Data Sets
  Key Word Searching (Boolean Logic, Proximity Connectors, Stemming, Wildcards)
  Conceptual Searching (Linguistics, Latent Semantics, Bayesian Technology)
- Develop a Data Set Review Approach/Protocol
  Determine Review Approach (Native, Near Native, Near Paper (Image/TIFF))
  Identify, Segregate and Log Documents as Responsive, Privileged, Significant and/or Confidential
- Validate Data Set Creation, Search and Review Approach/Protocol
  Establish a Quality Control Process to Test for Completeness and Accuracy

ORANGE LEGAL TECHNOLOGIES

Data Review

Key Discovery Review Platform Requirements

- Centralization - Review tools should allow for time efficient for complex searches against large volumes of documents from a centralized review architecture.
- Scalability - Review tools should allow the client to take full advantage of all available processing power regardless of the size of the data set being reviewed or the complexity of the review queries. The investment protection provided by a scalable and centralized review architecture ensures that growing capacity requirements do not adversely affect review capability.
- Security - Review tools should be secured and supported with forensically sound processes and protocols for both physical and digital security
- Usability - Review tools should be able to be easily accessed and intuitively used by multiple reviewers, from multiple locations, potentially on different review teams.

ORANGE LEGAL TECHNOLOGIES

Data Production

Primary Formats for ESI Production

- **Native File Formats** Files produced in the format they were created and maintained are known as native production. In a native production, MS Word documents are produced as .doc files, MS Excel files are produced as .xls files, and Adobe files are produced as .pdf files, etc. Native format is often recommended for files that were not created for printing such as spreadsheets and small databases. For some file types the native format may be the only way to adequately produce the documents.

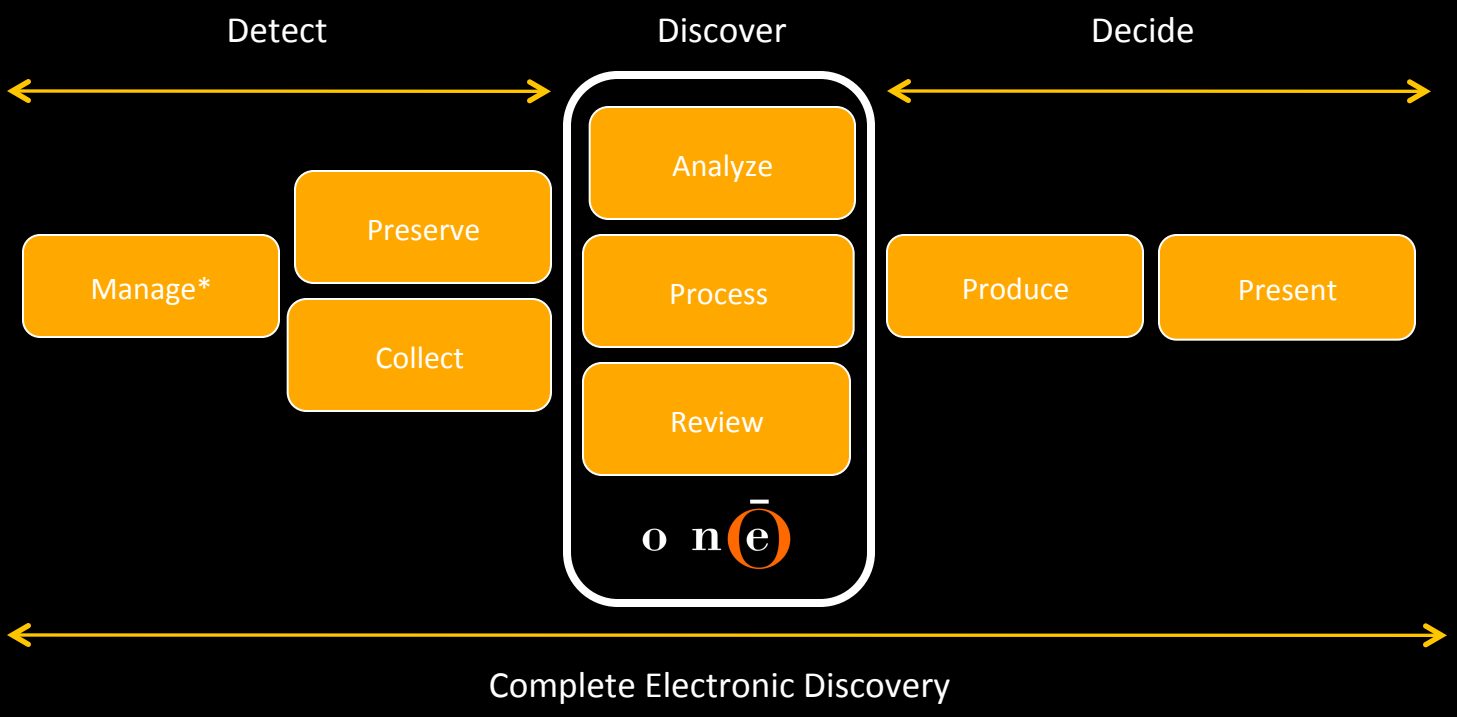
- **Near Native File Formats** Some files, including most e-mail, cannot be reviewed for production and/or produced without some form of conversion. Most e-mail files must be extracted and converted into individual files for document review and production. As a result, the original format is altered and they are no longer in native format. There is no standard format for near-native file productions. Files are typically converted to a structured text format such as .html or xml. Other common e-mail formats include .msg and .eml.

- **Near Paper Formats** ESI can also be produced in a near paper format. Rendering an image is the process of converting ESI or scanning paper into a non-editable digital file. During this process a "picture" is taken of the file as it exists or would exist in paper format. Expertise in the field of electronic discovery and image rendering tools are necessary to minimize this risk.
- **Paper Format** A paper production is just what it sounds like: paper is produced as paper or ESI is printed to paper and the paper is produced. When producing ESI in paper, it is recommended to utilize someone with expertise in the field of e-discovery and image rendering tools to minimize this risk during the printing or image rendering process.

ORANGE LEGAL TECHNOLOGIES

Orange Legal Technologies and eDiscovery

Orange Legal Technologies delivers eDiscovery services that support the litigation, audit, and investigation requirements of legal professionals.



* Information Management and Identification in the Electronic Discovery Reference Model

Orange Legal Technologies and eDiscovery

Orange Legal Technologies

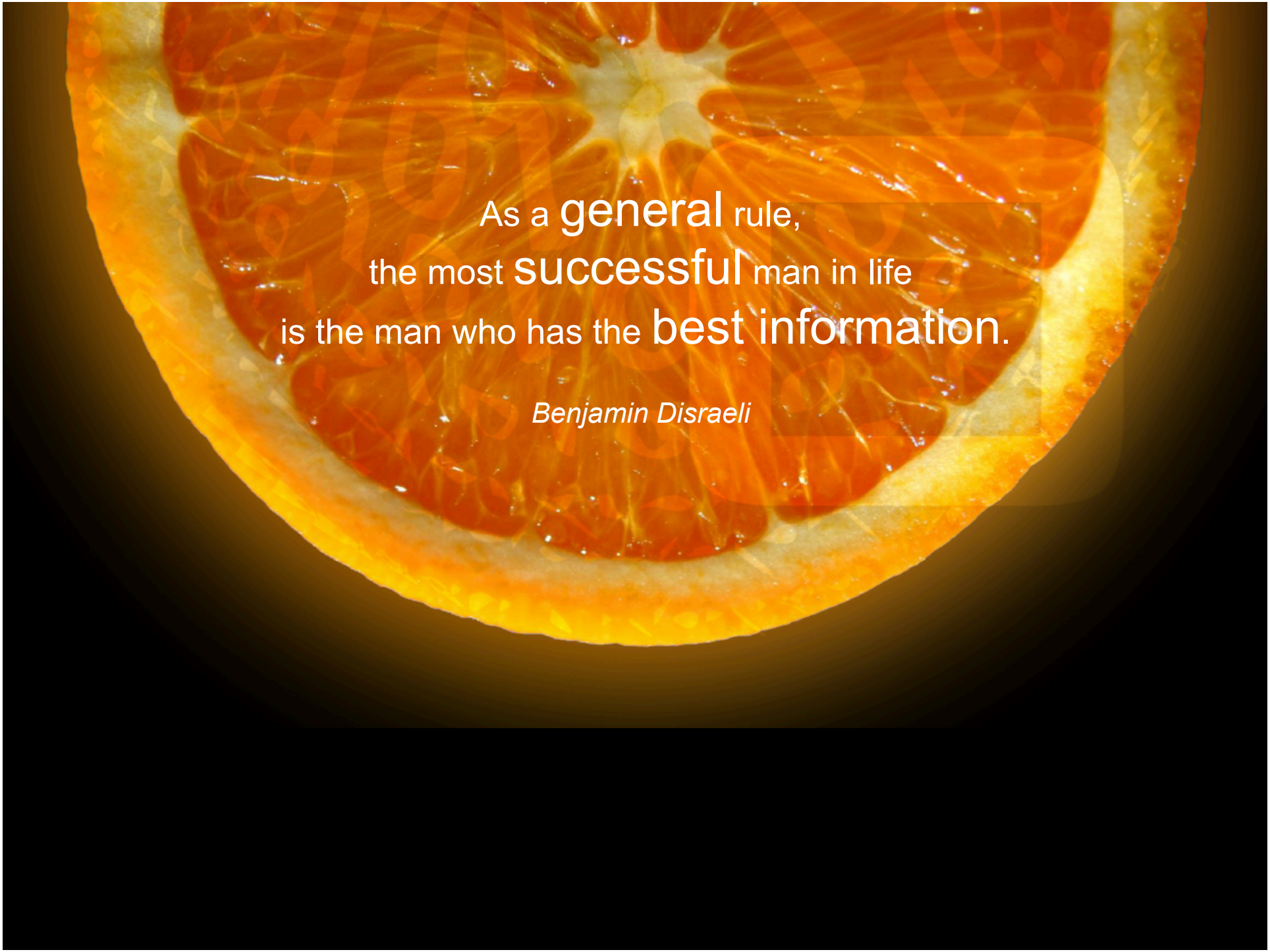
Advantages

- We provide a complete and integrated eDiscovery platform.
- We provide services via a Software-as-a-Service Model.
- We provide a highly competitive pricing structure.
- We have a proven management team.

Results

- eDiscovery in ½ the time of traditional services.
- eDiscovery at ½ the cost of traditional services.
- eDiscovery with lower risk than traditional and advanced services.

ORANGE LEGAL TECHNOLOGIES


As a general rule,
the most successful man in life
is the man who has the best information.
Benjamin Disraeli