

Searching for ESI in the EU:
Some Rules of the Road for the European Union Data Controller

By Marissa L. P. Caylor¹ and Jason R. Baron²

I. Electronically Stored Information and the Information Age

“[I]nformation, as a cultural and technological edifice, has profoundly and irrevocably changed.”³ Electronically stored information (“ESI”) comprises over 90% of records spread throughout the world, while paper comprises less than one-tenth of one percent of all existing information.⁴ In terms of modern day legal practice, the consequences of this new ESI reality are not limited merely to litigation conducted within the United States involving only domestic litigants. As noted in a recent European Union (“EU”) document on pre-trial discovery, the “vast amount” of ESI needing to be managed, and the ease of accessing and transferring it “means that more information is obtainable and discloseable with greater ease.”⁵

¹ Boston University School of Law, J.D. and M.A. in International Relations Candidate 2010; University of Idaho, B.S. *summa cum laude*, 2004.

² Director of Litigation, U.S. National Archives and Records Administration; Boston University School of Law, J.D. 1980.

³ George L. Paul & Jason R. Baron, *Information Inflation: Can the Legal System Adapt?*, 13 RICH. J.L. & TECH. 10 (2007), available at <http://law.richmond.edu/jolt/v13i3/article10.pdf> (arguing that the sheer volume of ESI involved in litigation requires new means of conducting automated searches and a more cooperative discovery model).

⁴ Peter Lyman & Hal R. Varian, *How Much Information?* (2003), http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/printable_report.pdf (“Ninety-two percent of new information is stored on magnetic media, primarily hard disks. Film represents 7% of the total, paper 0.01%, and optical media 0.002%.”). See Mark Levitt et al., *Worldwide Email Usage Forecast, 2002-2006: Know What's Coming Your Way* (International Data Corporations), Sept. 2002, (estimating that over 30 billion emails were sent per day worldwide in 2002 and that the rate would double to 60 billion by 2006); See also Carole Basri & Mary Mack, *EDISCOVERY FOR CORPORATE COUNSEL* § 22:2 (Westlaw Sept. 2008) (reporting that emails typically contain 70% of a company’s digital assets and that Americans alone send more than 50 billion e-mails and two billion instant messages (IMs) every day).

⁵ European Union Article 29 Data Protection Working Party, Working Document 1/2009 on Pre-trial Discovery for Cross Border Civil Litigation, EUR. PARL. DOC. (WP 158) 3 (2009) [hereinafter “Article 29 Working Document”].

Moreover, advances in technology and information management systems impact both private and public sectors in terms of the volume of “personal data”⁶ collected, an issue of particular concern given EU laws and norms regarding protection of said personal data.⁷ Commercial businesses today regularly collect personal data to manage their payroll and human services functions.⁸ Governments, hospitals, and researchers use personal data to facilitate services.⁹ Professionals such as attorneys and accountants use personal data to advise clients.¹⁰ Directive 2006/24/EC of the European Parliament and of the Council on the Retention of Data Generated or Processed in Connection with the Provision of Publicly Available Electronic Communications Services or of Public Communications Networks and Amending Directive 2002/58/EC (“Retention Directive”) requires telecommunication and internet businesses to temporarily retain personal data to assist the criminal justice system.¹¹

⁶ Council Directive 95/46, Protection of Individuals With Regard to the Processing of Personal Data and on the Free Movement of Such Data, 1995 O.J. (L 281) 31 (EC) [hereinafter “Data Directive”], art. 2(b) (Personal data is “any information relating to an identified or identifiable natural person. . . [and] an identifiable person is one who can be identified . . . [by] one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.”); *see generally* Opinion No. 4/2007 on the Concept of Personal Data Issued by the Article 29 Working Party, EUR. PARL. DOC. (WP 136).

⁷ European Union Article 29 Data Protection Working Party, *Eleventh Annual Report of the Article 29 Working Party on Data Protection on the Situation Regarding the Protection of Individuals with Regard to the Processing of Personal Data and Privacy in the European Union and in Third Countries*, Jun. 24, 2008, p. 5 (“Technological and economic developments lead to more and more comprehensive data processing in increasingly complex IT systems.”).

⁸ Masons Solicitors and Privy Council Agents, *Handbook on Cost Effective Compliance with Directive 95/46/EC*, Aug. 1998, p. 6 (“employers need to keep personal information about employees in order to pay wages, make social security contributions and fulfill their other legal and social obligations as Employers”).

⁹ *Id.* at 6 (“hospitals and other health care providers must keep information about their patients including full medical records and any other factors that may have a bearing on patients’ mental and physical well being”).

¹⁰ *Id.* at 7 (“many professional people such as doctors, lawyers and accountants will often need personal data in order to give their clients the correct advice”).

¹¹ European Union Article 29 Data Protection Working Party, Opinion 3/2006 on the Directive 2006/24/EC of the European Parliament and of the Council on the Retention of Data Generated or Processed in Connection with the Provision of Publicly Available Electronic Communications Services or of Public Communications Networks and Amending Directive 2002/58/EC, EUR. PARL. DOC. (WP 119) 3 (2006) (“The data should only be retained for specific purposes.”); European Union Article 29 Data Protection Working Party, Opinion on the Statement of European Data Protection Commissioners at the

When litigation is involved, differences between common law and civil code legal systems further complicate each party's goals, as "the scope of discovery differs greatly between common law and civil code jurisdictions and is seen as a fundamental part of the litigation process of the former."¹² In civil code jurisdictions such as in Spain, France, and Germany, evidence is limited to that which is produced by the party in possession of the data set, and only insofar as it "is needed for the scope of the trial."¹³ The burden on a party seeking evidence from the opposing side is to specifically "be able to know and identify it."¹⁴

As further noted with respect to pre-trial litigation in the EU,

In order for the pre-trial discovery procedure to take place lawfully, the processing of personal data needs to be legitimate and to satisfy one of the grounds set out in Article 7 of the Data Protection Directive. . . . namely, [i] consent of the data subject, [ii] that the compliance with the pre-trial discovery requirements is necessary for compliance with a legal obligation . . . or [iii] [necessary for the] . . . purposes of a legitimate interest pursued by the controller or by the third party to whom the data are disclosed . . . ¹⁵

The Article 29 Working Party Document goes on to set out at length what constitutes meeting one or more of the above conditions. For present purposes, however, it is sufficient to note that in at least some legal situations parties will meet the conditions set out above, and that data controllers (as defined in the next section) must be prepared to retain a large amount of data, including relevant personal data, for the purpose of facilitating access by way of manual and, increasingly, automated searches by various proposed means.

International Conference in Cardiff (9-11 September 2002) on Mandatory Systemic Retention of Telecommunication Traffic Data, EUR. PARL. DOC. (WP 64) 3 (2002) (reviewing proposal requiring telecommunications service providers to retain personal data "for a period of one year or more, in order to permit possible access by law enforcement and security bodies").

¹² Article 29 Working Document, *supra* n.5 at 3.

¹³ *Id.* at 4.

¹⁴ *Id.* at 5.

¹⁵ *Id.* at 8, 9.

“The ultimate goal in discovery is to identify, collect and cull documents and ESI from a larger corporate or institutional data universe and then search for and retrieve the relevant or responsive non-privileged materials with whatever tools or methods have been made available (whether automated, human, or some combination of the two).”¹⁶ Thus, as is increasingly the case, the ever-growing amounts of ESI collected from day-to-day business and personal activities in the EU complicates the search for and identification of relevant documents during e-discovery. Of particular interest is whether, in light of the heightened concerns in the EU regarding the privacy of personal data, there can be said to be a commonality of “best practices” regarding how one specifically conducts searches for whatever relevant documents are preserved by EU data controllers.

II. Data Controllers and the European Union Data Directive

Data Directive 95/46/EC applies to any transfer or processing¹⁷ of personal data containing information that could be used to identify a “natural” person with citizenship in an EU Member State.¹⁸ The Data Directive applies “to [all] processing of personal data wholly or partly by automatic means”¹⁹ Processing of personal data includes alteration, collection, retrieval, e-mailing, or deleting of any data stored in a structured database, whether by manual or automatic means.²⁰

¹⁶ The Sedona Conference® COMMENTARY ON ACHIEVING QUALITY IN THE E-DISCOVERY PROCESS 4 (PUBLIC COMMENT VERSION 2009) [hereinafter “*Sedona Achieving Quality Commentary*”], available at www.thesedonaconference.org,

¹⁷ Data Directive, *supra* note 6, at art. 2(b) (Processing covers “any operation . . . performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction”).

¹⁸ *See generally id.*

¹⁹ *Id.* at art. 3.

²⁰ *Id.* at 2(b) (Processing “covers any organization or storage of e-mails, electronic data, or scanning of paper documents.”); *see also* Article 29 Working Document, *supra* note 5 at 8 (“[A]ny retention, preservation, or archiving of data [is] processing.”).

Processing of personal data does not violate the Data Directive if it is part of a legal obligation, if it is necessary to protect the interests of the public or the data subject, or if the subject of the personal data consents.²¹ Processing may also be appropriate if it is “necessary for the purposes of legitimate interests pursued by the controller . . . except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject.”²²

Any organization, agency, or individual is a “data controller” if it is responsible for determining whether personal data falls under the Data Directive, and whether the purpose of processing might violate the Data Directive.²³ Entities or people who process the information but do not determine the type of data collected or the purpose for which is collected, are data processors and not data controllers.²⁴

Data controllers must ensure that any processing of personal data is fair and lawful, and done for a legitimate purpose.²⁵ They must notify data subjects before processing data for any new purpose, even if the new purpose is legitimate.²⁶ Data controllers must also ensure

²¹ Data Directive, *supra* n.6 at art. 2(h) (consent means “any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed”).

²² *Id.* at art. 7.

²³ *Id.* at art. 2(d) (“‘controller’ shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by national or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law”).

²⁴ Data Directive, *supra* n.6, at art. 2(e) (a processor is “a natural or legal person . . . [who] processes personal data on behalf of the controller”).

²⁵ *Id.* at art. 2(d) (defining data controller as a “natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data”); *see also id.* at art. 6 (“personal data must be: (a) processed fairly and lawfully; (b) collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes . . . ; (c) adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed; (d) accurate . . .”).

²⁶ *Id.* at 18-19 (“These laws provide for a series of rights for individuals such as the right to receive certain information whenever data are collected, the right of access to the data and, if necessary, the right to have the data corrected, and the right to object to certain types of data processing.”).

“proportionality and the balance of the rights of the different interests,”²⁷ and must not transfer personal data to third countries without ensuring that the third country will adequately protect the data, unless the necessity of the transfer outweighs the interests of the data subject.²⁸

Any processing of the data must comply with national laws, regardless of where the data subject is located.²⁹ Under national data protection laws that implement the Data Directive, data controllers are responsible for using information management processes to avoid collecting personal data that is not necessary or relevant.³⁰ Data controllers must filter out any irrelevant

²⁷ Article 29 Working Document, *supra* n.5, at 10.

²⁸ *See generally* Data Directive, *supra* n.6, at art. 25 (“The Member States shall provide that the transfer to a third country of personal data which are undergoing processing or are intended for processing after transfer may take place only if, without prejudice to compliance with the national provisions adopted pursuant to the other provisions of this Directive, the third country in question ensures an adequate level of protection”); *see generally id.* at art. 26 (“By way of derogation from Article 25 and save where otherwise provided by domestic law governing particular cases, Member States shall provide that a transfer or a set of transfers of personal data to a third country which does not ensure an adequate level of protection within the meaning of Article 25 (2) may take place on condition that: (a) the data subject has given his consent unambiguously to the proposed transfer; or (b) the transfer is necessary for the performance of a contract between the data subject and the controller or the implementation of precontractual measures taken in response to the data subject’s request; or (c) the transfer is necessary for the conclusion or performance of a contract concluded in the interest of the data subject between the controller and a third party; or (d) the transfer is necessary or legally required on important public interest grounds, or for the establishment, exercise or defence of legal claims; or (e) the transfer is necessary in order to protect the vital interests of the data subject”).

²⁹ Data Protection Unit of the Directorate-General for Justice, Freedom and Security, *Frequently Asked Questions Relating to Transfers of Personal Data from the EU/EEA to Third Countries*, p. 19 (“Each controller has to comply with the provisions of the Member State where he or she is established even if the personal data relate to data subjects established in other Member States.”).

³⁰ Data Protection Unit of the Directorate-General for Justice, Freedom and Security, *Frequently Asked Questions Relating to Transfers of Personal Data from the EU/EEA to Third Countries*, p. 18 (“Data protection laws generally demand good data management practices on the part of the entities that process data . . . and include a series of obligations and rights for data subjects.”).

data and disclose only relevant data.³¹ Even if the data is relevant, the data controller may have to anonymize it if the identity of the data subject is not necessary or relevant.³²

Despite the wealth of responsibilities and the ease of violating the Data Directive, there is currently not much guidance to aid data controllers in carrying out their important search and information retrieval tasks.³³

III. Data Directive Compliance and Search and Retrieval During E-Discovery

“The selection, organization and filtering of ESI through the use of a search protocol is a critical element in reducing the volume of information to be collected and thus the time and cost of collection.”³⁴ Data controllers may use any number of common search strategies³⁵ to process personal data, from common forms of keyword searching, to more complex automatic algorithmic searches.³⁶ In litigation, parties use keyword searches to identify specified words or phrases within documents, and courts in both the U.S. and now in the U.K. often use them “to define discovery parameters and resolve discovery disputes.”³⁷

³¹ See Article 29 Working Document, *supra* n.5, at 10 (“As a first step controllers should restrict disclosure if possible to anonymised or at least pseudonymised data. After filtering (‘culling’) the irrelevant data – possibly by a trusted third party in the European Union – a much more limited set of personal data may be disclosed as a second step.”).

³² See *id.* at 11 (“Once personal data has been identified, the data controller would need to consider whether it is necessary for all of the personal data to be processed, or for example, could it be produced in a more anonymised or redacted form.”).

³³ See generally Regulation (EC) No. 45/2001 of the European Parliament and of the Council of 18 December 2000 on the Protection of Individuals with Regard to the Processing of Personal Data by the Community Institutions and Bodies and on the Free Movement of Such Data, O.J. L. 8, Jan. 12, 2001.

³⁴ *Sedona Achieving Quality Commentary*, *supra* n.16, at 15.

³⁵ *Id.* at 17 (“(e.g., Boolean searches, concept searches, metadata filters, language-based approaches using taxonomies and ontologies, statistical clustering techniques, or other proprietary strategies”).

³⁶ *Id.* at 15.

³⁷ The Sedona Conference®, *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*, 8 THE SEDONA CONFERENCE JOURNAL 199 (2007) [hereinafter *Sedona Search Commentary*]; see, e.g., *FTC v. Ameridebt, Inc.*, 2006 WL 6188563 (N.D. Cal. Mar. 13, 2006) (“e-mail could likely be screened efficiently through the use of electronic search terms that the parties agree upon”); Jason R. Baron, “Toward A New Jurisprudence of Information Retrieval: What constitutes a ‘reasonable’ search for digital evidence when using keywords?,” 5 DIGITAL EVIDENCE AND ELECTRONIC SIGNATURE REVIEW 173 (2008). See also *Digicel (St. Lucia) Ltd. & Ors. v. Cable &*

However, “[s]imple keyword searches end up being both over- and under-inclusive in light of the inherent malleability and ambiguity of spoken and written English (as well as all other languages).”³⁸ Since keyword searches identify specific words regardless of context, “the relative percentage of ‘false positive’ hits or noise in the data is potentially huge, amounting in some cases to huge numbers of files which must be searched to find responsive documents.”³⁹ Keyword searches may also fail to find relevant documents due to misspellings and variations on search terms.⁴⁰

Searches that are over-inclusive can increase the risk of a violation of data protection laws by bringing up personal data that is irrelevant to the purpose of the search, even if the purpose of the search is legitimate, since “[k]eyword searches identify all documents containing a specified term regardless of context, and [can] capture many documents irrelevant to the user’s query.”⁴¹ Cautionary tales based on recent U.S. litigation highlight the potential for such over-inclusive searching and production, based on inadequacies in how keywords were utilized.⁴²

“Electronically stored information contains human language, which challenges computer search tools. These challenges lie in the ambiguity inherent in human language and tendency of

Wireless & Ors., [2008] EWHC 2522 (Ch.) (U.K. case addressing at length the propriety of the parties’ proposed keywords).

³⁸ *Sedona Search Commentary*, *supra* n.37, at 200; *see also* *Quinby v. WestLB*, AG, 2006 WL 2597900 (S.D.N.Y. Sept. 5, 2006) (narrowing demand for 170 proposed search terms due to use of common words).

³⁹ *Sedona Search Commentary*, *supra* n.37, at 200.

⁴⁰ *Id.* (“Keyword searches can also exclude common or inadvertently misspelled instances of the term (e.g., “Phillip” for “Philip,” or “strik” for “strike”) or variations on “stems” of words (e.g. “striking”).”

⁴¹ *Id.* at 200 (“Some case law has held that keyword searches were either incomplete or overinclusive . . .”); *see* *Quinby v. WestLB*, AG, *supra* n.38 (narrowing demand for 170 proposed common search terms as over-inclusive).

⁴² *See, e.g.,* *Walter Gross Construction Associates Inc., v. Am. Mfrs. Mutual Ins. Co.*, 2009 WL 724954 (S.D.N.Y. March 19, 2009) (court criticizing party for proposing 1,000 search terms); *In re Fannie Mae Litigation*, 552 F.3d 814 (D.C. Cir. 2009) (appellate court upholds finding of contempt against federal agency in failing to review a production set consisting of 660,000 recovered documents, where 400 keywords had been proposed and agreed to). *See also* *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251 (D. Md. 2008) (party deemed to have waived attorney client privilege based on inadequate keyword methodology resulting in overproduction of privileged documents to opposing counsel).

people within organizations or networks to invent their own words or communicate in code.”⁴³

Placed in an international litigation context, foreign language materials make this challenge twofold, as each language contains its own unique characters.⁴⁴

IV. Best Practices for Data Directive Compliance During E-Discovery

A. Structure Searches to Avoid Over-inclusiveness and Document Searches to Ensure Quality

Search and retrieval methods can affect the cost of e-discovery whether they are under-inclusive and fail to produce necessary relevant information, or over-inclusive and produce too many irrelevant results.⁴⁵ In either case, additional processing is necessary, raising the already high cost of e-discovery and document review.⁴⁶ “Thus, the identification and use of best practices in collection, review and production are essential.”⁴⁷

One of the factors included in Principle 3 of the *Sedona Achieving Quality Commentary* that identifies “a well thought out e-discovery ‘process’” is whether the process “enhance[s] the overall quality of the production in the form of . . . reducing cost”⁴⁸ “The first step, then, is the development of a well thought-out process in which the applicable review method can be applied.”⁴⁹

⁴³ *Sedona Search Commentary, supra*, n.37, at 194.

⁴⁴ Jaculin Aaron & Laura J. Lattman, *Electronic Discovery in Europe: A Different Story*, THE NATIONAL LAW JOURNAL, Dec. 11, 2007, <http://www.law.com/jsp/legaltechnology/pubArticleLT.jsp?id=1197281078728> (“The best format for supporting a multilingual review is Unicode, as it will process diacritical marks (such as accents, umlauts, diereses and cedillas) and special characters in Latin alphabets (such as Eszett (ß), Thorn (þ), and Æthel (œ)), as well as characters from languages not written in the Latin alphabet, e.g., Chinese and Arabic.”).

⁴⁵ *Sedona Achieving Quality Commentary, supra* n.16, at 1 (“Failure to employ a quality e-discovery process can result in failure to uncover or disclose relevant evidence which can affect the outcome of litigation.”).

⁴⁶ *Id.* at 1 (“A poorly planned effort can also cost more money in the long run if the deficiencies ultimately require that e-discovery must be redone.”).

⁴⁷ *Id.* at 2.

⁴⁸ *Id.* at 2.

⁴⁹ *Id.* at 6.

Commentary to Principle 11 of the original *Sedona Principles*⁵⁰ suggests that discovery procedures, “whether manual or automated,” must consistently comply with all discovery obligations.⁵¹ Consistent documentation of the process also ensures consistency among searches, and can help identify when a search has been over- or under-inclusive, and if over-inclusive, when Data Directive issues arise.⁵² Such documentation would keep track of what is being searched for, how searches are being conducted, as well as the overall process employed. This would appear to be especially important in light of the increased need to conduct “federated” searches across data sets residing in a variety of EU jurisdictions, in an effort to comply with cross-border discovery demands.

Practice Point 7 of the *Sedona Search Commentary* recommends that “[p]arties should expect that their choice of search methodology will need to be explained, either formally or informally, in subsequent legal contexts (including in depositions, evidentiary proceedings, and trials).”⁵³

This may also include making the algorithm available to data subjects in the European Union upon request. Time should also be set aside to obtain permission from a data subject to use the data for a secondary purpose.

⁵⁰ See The Sedona Conference®, THE SEDONA PRINCIPLES, SECOND EDITION: BEST PRACTICES, RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION (2007), available at http://www.thesedonaconference.org/content/miscFiles/TSC_PRINCP_2nd_ed_607.pdf (“A responding party may satisfy its good faith obligation to preserve and produce relevant electronically stored information by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data reasonably likely to contain relevant information.”).

⁵¹ *Id.* at 58 (“Regardless of the method chosen, consistency across the production can help ensure that responsive documents have been produced as appropriate.”).

⁵² See *Sedona Achieving Quality Commentary*, *supra* n.16, at 6 (“Documentation. The overall process should be documented to ensure coordination and communication within the discovery team and to increase the defensibility of the process.”).

⁵³ See *Sedona Search Commentary*, *supra* n.37, at 195.

B. Ensure Someone Is In Charge of Monitoring Compliance

Once a search has been structured to avoid over-inclusiveness and potential violations of the Data Directive, the data controller should monitor the process for consistency.⁵⁴ “This individual or the team he or she leads should have sufficient experience in the various phases of e-discovery to effectively execute the management duties.”⁵⁵ “[C]are must be taken to obtain regular status updates, maintain frequent contact with the team at all levels, and ensure the effective and appropriate dissemination of information to and from team members.”⁵⁶ A consistent and monitored process is vital to the discovery process, as “failure to employ a quality e-discovery process can result in failure to uncover or disclose key evidence . . . and a poorly conceived or managed e-discovery process may allow privileged or confidential information to be inadvertently produced.”⁵⁷ When necessary, the individual in charge could utilize these best practices by delegating the monitoring and response to potential Data Directive violations to someone with knowledge of EU law or cross-border litigation.

“Measuring the quality of discovery processes as they occur, or reasonably soon thereafter, allows the attorney(s) functioning as Team Leader to determine if the discovery processes are working in the manner intended, or if there is a systemic or systematic error that is biasing or corrupting the results, and that necessitates some kind of modification.”⁵⁸

⁵⁴ *Sedona Achieving Quality Commentary*, *supra* n.16, at 2 (“Principle 1. In cases involving ESI of increasing scope and complexity, the attorney in charge should utilize project management and exercise leadership to ensure that a reasonable process has been followed by his or her legal team to locate responsive material.”).

⁵⁵ *Id.*

⁵⁶ *Id.*

⁵⁷ *Id.* at 8.

⁵⁸ *Id.*

Having individuals review and monitoring a process for effectiveness and potential violations of data protection laws makes the quality of the process “more defensible.”⁵⁹ “[P]oor quality can also cost a party more money in the long run if deficiencies are noted in the documents produced, motion practice ensues and, if in the end e-discovery must be redone.”⁶⁰ “It is often less expensive to engineer quality into the process than to add it on after the fact.”⁶¹

Most e-discovery processes designate someone responsible for overseeing compliance with discovery rules.⁶² In e-discovery processes involving material stored or processed within the EU, someone should be charged with ensuring compliance with the Data Directive.⁶³ Team leads within data controllers who are involved in litigation should be familiar with their responsibilities and structure their search and retrieval process to avoid potential violations of the Data Directive.⁶⁴ As noted above, the process needs to include consistent documentation of the search and information retrieval methods used, as data subjects have a right to obtain algorithms used to process their personal data and to maintain legitimacy of processing.⁶⁵

V. Conclusion

Data controllers in the EU will, like their counterparts in the United States, be increasingly faced with pressing litigation demands that will require increasingly sophisticated means of searching large amounts of ESI, some of which will be personal data that is stored

⁵⁹ *Id.*

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² *See id.* at 6 (“An effective process will usually include most, if not all of these elements. Leadership. The process should be led by a person who will be responsible for assuring that a discovery process reflects a reasonable good-faith effort to be complete and accurate.”).

⁶³ *See id.* (“The discovery process should be tailored to the specific size, risks, needs and circumstances of the case or investigation that is the occasion for the retrieval effort.”).

⁶⁴ *See id.* (“The effort should incorporate and draw on the appropriate range of expertise required to meet and accomplish the goals set for it in a timely and cost-effective manner.”).

⁶⁵ *Id.* (“The overall process should be documented to ensure coordination and communication within the discovery team and to increase the defensibility of the process.”).

within European databases and information repositories of all kinds. Notwithstanding the differences in approaches to e-discovery in common law and civil code jurisdictions, there is a shared interest in reducing the volume of ESI – and personal data -- to be searched using automated means of filtering and more sophisticated search methods and techniques. We believe that the EU data controller would benefit from close study of emerging best practices in this area.