CLAI COMMERCIAL LITIGATION Association of Ireland

Good Practice Discovery Guide

3rd edition - December, 2025

Table of Contents

Foreword		3
Chapter 1	Introduction	6
Chapter 2	Principles	9
Chapter 3	Outline of discovery phases	11
Chapter 4	Preparing for discovery	13
Chapter 5	Identification	24
Chapter 6	Planning to make discovery	31
Chapter 7	Preservation	33
Chapter 8	Collection	36
Chapter 9	Processing	40
Chapter 10	Discovery request	49
Chapter 11	Agreeing the terms of discovery and motions for discovery	55
Chapter 12	Review	62
Chapter 13	Production	67
Chapter 14	Presentation in Court	71
Appendix A	Discovery project checklist	73
Appendix B	Sample legal hold communications	75
Appendix C	Document identification questionnaires	78
Appendix D	Managing audio and video data	87
Appendix E	Understanding deduplication, families, and threads	90
Appendix F	Technology Assisted Review	102
Appendix G	Sample discovery plan	109
Appendix H	Sample review plans	125
Appendix I	Sample request for voluntary discovery	138
Appendix J	Sample affidavit of discovery	141
Appendix K	Overview of legal privilege	147

Foreword

It is hard to believe that it is ten years since the second edition of the Good Practice Discovery Guide (the "Guide") was published by the Commercial Litigation Association of Ireland (the "CLAI") in November 2015. In his Foreword to the first edition of the Guide, published in March 2014, former Chief Justice Frank Clarke said that the recommendations in the Guide *"will come to be seen as the norm which the court will expect parties to comply with in the absence of a good reason for deviation…".* The former Chief Justice also wrote the Foreword to the second edition of the Guide, in which he described the Guide as being of *"significant benefit to litigators, whether involved in small scale cases or major document driven litigation and to the courts"* and as setting the *"industry standard"* for discovery. He observed that the Guide would further the aim of achieving *"the maximum likelihood of producing a fair and just result at the minimum cost to the parties".* I wholeheartedly endorse those comments which apply with even greater force to the latest edition of the Guide.

Both I and other judges of the Superior Courts have had cause to consider the benefits of the last edition of the Guide and all have commented very positively on them. Judges have described the Guide as being *"undoubtedly an excellent document"* (Ní Raifeartaigh J),[1] as providing a *"useful tool in many actions where the discovery of electronically stored information is in issue"* (Keane J),[2] and as being *"invaluable for persons involved in making discovery" (*Meenan J).[3] Judges have also described the Guide as providing *"very clear advice*" (Collins J)[4] which made *"perfect sense*" (Barniville J).[5]

It is fair to say that these sentiments reflect the significant contributions made by the previous editions of the Guide in improving the discovery process and in making it as efficient and effective as it could be, bearing in mind the legal constraints under which courts and practitioners must operate in the discovery area.

This edition of the Guide has been long awaited by practitioners and is clearly the product of an extraordinary amount of work carried out by a dedicated CLAI Sub-Committee, comprising experienced practitioners seasoned in all aspects of discovery. The real value of this edition is not just in the calibre of those involved in its production, but in the superlative quality of the material and the guidance it contains.

[1] *Ryanair Ltd v. Irish Municipal & Ors [2017] IEHC 425, para 28.*
[2] *Gallagher v. Raidió Teilifís Éireann [2017] IEHC 237, para 43.*
[3] *Ryanair Ltd v. Channel 4 Television Corporation & Anor [2017] IEHC 744, para 50.*
[4] *McNulty v. The Governor & Company of the Bank of Ireland [2021] IECA 182, para 68.*
[5] *Victoria Hall Management Ltd & Ors v. Patrick Cox & Ors [2019] IEHC 639, para 73.*

The developments, technological and otherwise, which have taken place since the release of the previous edition of the Guide have been nothing short of ground breaking. As the Guide demonstrates, the whole process of discovery has been radically transformed by the multiplicity of electronic devices now used by people to communicate, including smart phones, tablets and watches, and by the media by which they communicate, including the well known video conferencing platforms. While these have all made communication much easier and faster, they have also greatly increased the quantity of potentially discoverable material in the discovery process and sources of that material. Advancements in Artificial Intelligence ("AI") and, in particular, Generative AI ("GenAI") have brought about enormous advantages in the identification and analysis of vast amounts of data and material with speed and precision. As outlined in the Guide, the use of GenAI in the discovery process can transform the labour intensive task of identifying and reviewing material, bringing it down from weeks to days. Most if not all discovery in litigation involving large amounts of material is made using e-discovery platforms which incorporates GenAI capabilities. While current best practices in respect of these technological advancements are captured in the Guide, it rightly acknowledges that given the pace at which GenAI is evolving, the relevant advice and recommendations will need to be kept under constant review and updated accordingly.

Since the publication of the previous edition of the Guide, there has been an extensive review of the discovery rules in Irish law with significant changes likely to come about in the future. As is noted in the Introduction which follows, the Review Group on the Administration of Civil Justice, chaired by the former President of the High Court, Mr. Justice Peter Kelly, recommended significant changes to the law on discovery in its report published in October 2020 (the "Kelly report"). In May 2022, the then Minister for Justice published a plan for the implementation of the recommendations contained in the Kelly report and it was recently announced by the current Minister that he intends to reform the law on discovery and the production of documents in a Civil Reform bill. The details of that Bill and the precise manner in which the law on discovery will be changed are not known at this time and the legislation, once enacted, may require significant aspects of the Guide to be revised. Notwithstanding that, I am in full agreement with the CLAI that it is right to proceed with the publication of this edition of the Guide and that it would be a lost opportunity not to do so until details of the proposed new legislative regime are announced.

I am pleased to report that the Guide is jam-packed with very useful advice and practical guidance on the nuts-and-bolts of how to handle discovery projects, including those involving significant amounts of electronically stored information to be found in the detailed and informative appendices.

As former Chief Justice Clarke did in relation to the previous editions of the Guide, I unreservedly and wholeheartedly recommend this edition of the Guide to all practitioners involved in the discovery process. This edition of the Guide will, like its predecessors, no doubt become the "go to" source of, and essential reference point for, anyone involved in advising or complying with discovery obligations and for the courts in assessing the compliance with those obligations.

Mr. Justice David Barniville
President of the High Court of Ireland
4 December 2025

Chapter 1 Introduction

This is the 3rd edition of the Commercial Litigation Association of Ireland "Good Practice Discovery Guide". The CLAI is a joint initiative of solicitors and barristers which seeks to support the promotion of best practice in commercial litigation in Ireland and to provide legal education and training to commercial litigation practitioners. The Guide was prepared by a Sub-Committee of the CLAI and contains recommendations and guidance as to best practice in managing discovery projects. The Sub-Committee is made up of legal practitioners with extensive experience in the practice and procedure of discovery and the Guide is intended to be of use to practitioners, as a statement of best practice, but also to the Judiciary to assist in providing insight into the practical challenges that litigants face when making discovery.

The Guide is drafted by reference to the usual life cycle of a case and examines the practical issues that arise in dealing with discovery at each particular stage of litigation. The Guide contains recommendations and guidance which are not prescribed by the Rules of the Superior Courts but which do represent the collective experience of the Sub-Committee as to how best to deal with the day-to-day management of discovery projects.

The 2nd edition of the Guide was published in November, 2015 and the massive technological innovation and change that has occurred in that ten-year period has had a particular impact on making discovery. As our smartphones have become our primary means of communication in both the professional and personal contexts, making discovery now requires that practitioners seek to gain access not just to the servers on which their clients' commercial data is stored, but also to the smart devices which their employees use to communicate. As instant messaging has become an accepted method of communication in the workplace, making discovery requires practitioners to review not just the e-mails of their clients' employees but also their WhatsApp messages, instant messaging and other social media engagement. As smartphones have made communication easier and faster, the number of individuals who may have received relevant documents or data has grown and, because of the growing preference in society to use electronic communication as a substitute for verbal communication, the amount of potentially discoverable data has increased significantly.

Another development in practice over the last decade has been that practitioners have had to begin to negotiate the sensitivities that arise when their obligation to locate potentially discoverable data requires them to review the personal data of their clients' employees. Client employees are frequently very reluctant to

simply allow their employer's legal advisors to download copies of all of their instant messages from their devices out of concerns of personal privacy. Similar concerns can arise with regard to the handling of personal photographs which are stored on their devices. The legal tensions between the obligation of a parties to litigation to make discovery of all documents which are within their power, possession or procurement and the rights of their employees to object to the processing of their personal data under the General Data Protection Regulation and similar legislation remain largely unresolved. Identifying a satisfactory way of "threading the needle" between these two potentially opposing points represents a significant challenge for practitioners and the Courts.

Something which has not changed in the last ten years has been the amount of time that it takes and the level of costs that are incurred in complying with orders for discovery. In the Foreword to the 2nd edition of this Guide in November, 2015, the then Chief Justice, Mr. Justice Frank Clarke, noted that "the burden of discovery obligations has come to represent a significant barrier to access to justice." Notwithstanding the significant developments in technology, including the technology used to actually make discovery (e.g. Continuous Active Learning), the situation has not changed and arguably has got worse. The formation of the Review Group on the Administration of Civil Justice, chaired by the then President of the High Court, Mr. Justice Peter Kelly, was announced in March, 2017 with the objective of seeking to reduce the costs of litigation in Ireland by recommending reforms of the law in including the law of discovery. The Group's Report was published on the 30th of October, 2020, but its recommendations have not yet been implemented. While the Government has recently announced its intention to implement the Report's recommendations on changes to the law on discovery through a planned Civil Reform Bill, the Bill has yet to be published and it is unclear how long the process of passing and enacting the necessary legislation will take. For the time being, it remains the task of practitioners to seek, through co-operation and good practice, to find ways of reducing the burdens that their clients have to meet in making discovery in this jurisdiction. It is the hope of the Sub-Committee that the 3rd edition of the CLAI Good Practice Discovery Guide will contribute to this.

The Sub-Committee wishes to draw attention to two points which should be borne in mind when reading this Guide. Firstly, while Order 31, Rule 12 RSC refers to "documents" and "electronically stored information", for the most part this Guide uses the term "data" to generally refer to the material which practitioners must collect and review for the purposes of making discovery.

Secondly, while the contents of this Guide represent what the Sub-Committee considers to be best practice at the time of publication, practices evolve and practitioners will inevitably identify new and more efficient ways of dealing with the practical issues that arise in making discovery. As was the case with previous versions of this Guide, we ask practitioners to submit any suggestions regarding changes to the practices recommended in this Guide to the CLAI via its website.

Chapter 2 Principles

The following principles have formed the basis of the statements and recommendations regarding good practice in discovery in this Guide.

1. Perfection in the discovery process (i.e. the identification and disclosure of all documents which ever existed and which may possibly be discoverable) is an unreasonable and disproportionate expectation. Save in exceptional circumstances, it will neither be feasible or possible for the party making discovery to guarantee that there have been no inadvertent omissions.

2. All media on which information is stored or recorded are in law "documents" for the purposes of the discovery process. Consequently all data, whether stored electronically or in hard copy, are *prima facie* discoverable and should be considered in every matter.

3. Parties should take all steps necessary to preserve sources of data as soon as they become aware of a matter which is likely to require discovery.

4. Parties should make all reasonable efforts to discuss and reach agreement on the discovery that is to be made in a particular case, including the categories of documents to be discovered, protocols for redacting documents and the parameters for asserting claims of privilege. While *inter partes* communications concerning discovery will entail detailed correspondence, there is no substitute for face-to-face communication between practitioners.

5. In seeking to reach agreement on discovery, parties should strive to ensure that the costs likely to be incurred in making discovery are proportionate to the value and/or importance of the issues in the proceedings, and the likely value which any documents discovered would bring to the matter.

6. Technology should be used to efficiently manage the process of making discovery. This may involve the use of Continuous Active Learning and AI tools to reduce the time and cost of making discovery while improving the accuracy of the production.

7. All data being discovered should be produced in a format which allows the receiving party the same ability to access, search, and review the data as the producing party. Basic metadata should be maintained and produced, including document name, author, recipient, date created, and date last modified.

8. The integrity of data should be maintained, but production of irrelevant material kept to a minimum. Consequently, duplicate families of documents and/or duplicate portions of email threads need not be produced. An audit trail of deduplication should be maintained, should inspection of duplicate documents be required later.

9. The solicitor owes duties to their client and the Court to ensure that discovery is thorough and properly made. Given the volume of data that is usually involved when making discovery in commercial litigation, it is recommended that the discovery process is undertaken with the assistance of a suitably qualified IT specialist. Where such professionals are engaged, it remains the duty of the solicitor to ensure that discovery is properly made.

10. A discovery audit file should be maintained by all parties' legal representatives to record decisions taken in respect of relevance and privilege.

11. Challenges to the adequacy of discovery, if necessary, should be substantive, specific and evidence based.

12. When a project is complete, protocols should be followed to close down the project and to ensure that it can be easily re-activated in the future if required.

Chapter 3 Outline of discovery phases

A discovery project will typically follow the phased approach set out in this guide. The phases involved in a typical discovery project include:

1. Identification Chapter 5	To identify the likely "universe" of relevant data including custodians and sources of data which may contain information which is relevant to the matter.
2. Preservation Chapter 7	To take steps to preserve data where they exist, so that they may not be altered or destroyed in advance of collection. This includes the legal hold process.
3. Collection Chapter 8	To obtain a copy of the data sources identified, so that they can be processed and searched for data of relevance to the matter. It is important to acquire the copy in a manner which does not alter the original data.
4. Processing Chapter 9	To convert the data sources collected into a format which will Facilitate their efficient searching and review. This involves removal of irrelevant data; converting the data into searchable format; deduplication; use of filtering processes to identify documents which may be of relevance (including application of temporal limits/ date ranges and keywords); and use of Technology Assisted Review.
5. Review Chapter 12	To perform a review and determine relevance and the privilege status of data highlighted as potentially relevant. This may be an entirely manual review, or Technology Assisted Review may be utilised.
6. Production Chapter 13	To produce (a.) a schedule listing all the data which have been identified as discoverable through the review phase; and (b.) a copy of the discoverable data.
7. Presentation Chapter 14	To prepare for, and to present, data in Court in an efficient manner.

Discovery projects are usually iterative by nature and it may be necessary for some of the phases to be repeated as more information comes to light. However, given that repetition of phases inevitably increases the time and cost of making discovery, precautionary steps should be taken to minimise the risk of having to repeat phases. As with all projects, detailed advance planning leads to significant efficiencies. Thus, taking care to obtain detailed instructions from the client on likely relevant custodians and data sources (Identification) will reduce the risk of later discovering the existence of relevant custodians and data sources and having to conduct Collection, Processing and Review phases in respect of such additional custodians and sources.

A number of **appendices** have been included in this Guide. The primary focus of these appendices is to provide detailed guidance and template/sample documents which may assist in a discovery project.

A Discovery project checklist	This checklist may be used as an aide-memoire to ensure that all key points of a discovery project are addressed.
B Sample legal hold communications	This provides a set of sample emails/letters which may be used as a template for legal hold communications.
C Document identification questionnaires	This provides a set of questionnaires used at the identification phase to identify sources of potentially relevant documents.
D Managing audio and video data	This provides an overview of how audio and video data may be managed throughout the discovery process.
E Understanding deduplication, families and threads	This provides an overview of what deduplication is and how it is used in discovery; what families of data are and how they impact discovery; and what email threads are and how they can be managed during discovery.
F Technology Assisted Review	This provides an overview of Technology Assisted Review ("TAR"), including Analytics, Continuous Active Learning ("CAL") and Generative AI. It includes common use cases for discovery.
G Sample discovery plan	This is a sample discovery plan which should be used to share information between the parties and the Court.
H Sample review plans	This is a sample review plan which should be used to document and plan the review by the producing party.
I Sample request for voluntary discovery	This is a template/sample letter of request for voluntary discovery.
J Sample affidavit of discovery with sample schedule	This is a template/sample affidavit of discovery with a sample schedule. This may be used as the basis for drafting an affidavit of discovery by parties.
K Overview of legal privilege	This provides an overview of legal privilege and guidance of how it impacts the discovery process.

Chapter 4 Preparing for Discovery

4.1 Introduction

Discovery costs will likely form the most significant element of litigation costs to be borne by the client in terms of financial resources and time inputs. The client should scope and prepare for discovery at the earliest opportunity. This must be a lawyer-led process. While clients may believe they can save costs by taking control of documentation identification and collection, this often results in a higher legal spend if the exercise has not been completed correctly and has to be redone. The client should be made aware at the outset that a senior representative of the client will need to swear the affidavit as to documents, who should be kept informed of the steps taken and decisions made throughout the discovery process.

4.2 Briefing your client on the nature and extent of obligations

Solicitors should advise clients about their discovery obligations and the impact, in time and costs, of the discovery process when first consulted in relation to a potentially contentious matter. In practice (unless the matter has been admitted to the Commercial List) it may be months before requests for discovery are exchanged but once a party becomes aware of a potential dispute it must identify and preserve all documentation that may be potentially relevant to the matters in dispute (See Chapter 5 - "Identification of document 'universe' "; Chapter 7 - "Preservation"; and Appendix B - "Sample legal hold communications"). The obligation to retain all potentially relevant data, including metadata, extends to everything that might contain relevant information. This means all forms of documents and data, including draft documents, handwritten notes, emails, text messages, screenshots, WhatsApp messages, Signal messages, Slack and MS Teams messages and any other form of instant messaging, PowerPoint presentations, Google Docs, back-up data, faxes, spreadsheets, Word documents, photographs and audio/video data, as well as data which may be stored online or contained in social media or in 'internet of things' devices. The obligation also extends to dynamic data, which can change in real time as more information is added to the relevant database, and depending on the nature of the dispute may extend to physical items, such as products in an IP dispute, or tissue samples in a personal injury case.

If a party fails to identify and preserve relevant documents and data, the Court has discretion to impose costs sanctions and may direct that the party bears the additional costs of having to retrieve and/or restore "lost" data. If this is not possible, the Court may draw an inference from the fact that such data no longer exists and cannot be produced at the trial. If the Court forms the view that there has been wilful deletion of potentially relevant data, it could strike out a crucial pleading such as a defendant's defence.

This means it is important to secure information at the outset even before the parameters of what will form the basis of a discovery request are clear. This includes securing information held on devices outside the party's control (e.g. personal devices) or vulnerable devices where data could be on software which is not backed up or saved elsewhere, such as OneNote or recordings of Microsoft Teams meetings that will auto-delete.

> This Guide recommends that at the outset, the solicitor takes instructions as to the methods of communications used within the client organisation, so that steps can be taken to immediately secure and preserve them.

Courts have criticised solicitors for not advising clients when first instructed about the need to protect and preserve data that may be relevant to the dispute. Accordingly, solicitors must immediately notify clients of this obligation. A sample notification is included at Appendix B.

4.3 Assembling the discovery project team

Much of the work involved in preparing for discovery can and should be front loaded. This will involve obtaining factual, legal and technical input at the start of the process. No one function within an organisation will have all the necessary skills to deal with a discovery request and one of the first steps should be to establish a Discovery Project Team.
The immediate responsibilities of this team will be to:

(1) work with the solicitor to consider what will be potentially relevant and where it is held, including data held by agents and advisors;

(2) prepare a data map of the potentially relevant material;

(3) identify who the custodians are, including agents and advisors;

(4) issue appropriate hold notices; and

(5) otherwise take steps to ensure that no data is lost, e.g. suspending automated data deletion policies.

At a minimum this team should include the main contact person from the client business/organisation charged with giving instructions and an IT representative. It could also usefully include a Project Manager from the client business/organisation and a Legal Representative. Where hard copy records are involved, a client representative with knowledge of archive procedures should also be included.

Clients should not undertake a discovery exercise without the involvement and oversight of their solicitor, who owes duties to the client and the Court to determine what is relevant and privileged and ensure that the client makes full and frank discovery. It may be a breach of the solicitor's professional obligations to facilitate a client preparing an Affidavit of Discovery without legal input or analysis.

In choosing the appropriate team members, regard should be had to seniority and experience. For example, the legal representative on the team should be (or have access to) a lawyer with previous discovery experience. This individual may have to swear an affidavit or give evidence in relation to the approach taken if something goes wrong. Similarly, the client's IT representative should be a senior IT person with universal access privileges. They may also have to swear an affidavit if issues arise at a later stage and defend the discovery process.

Given the volume of data that is usually involved when making discovery in commercial litigation, it is recommended that the discovery process is undertaken with the assistance of a suitably qualified IT specialist. The lawyer should work with the client to identify a suitable provider, taking into account any particular technological requirements. For example, if there is a need to access legacy data on old backup systems or carry out a forensic collection of devices the provider must have the capability to do this. Equally a sensitive matter might require a specialist who can operate a confidentiality ring.

If it is clear that expert evidence will be required for the purposes of the proceedings, an independent subject matter expert may assist in framing and responding to a discovery request, advising (1) what they require from the opposition to form a view as to the merits or otherwise of the case; (2) on the potential impact of particular documents on the case

(which may not be immediately apparent); and (3) what technical documents should be collated.

This Guide recommends that if the initial review of data is being conducted by a third party provider rather than by a law firm, the solicitor on record should prepare a briefing document with instructions for the external review team, including clear directions on the approach to be taken to relevance in the initial review. Frequent and robust sampling and quality control should be undertaken by solicitor's team, including sampling different reviewers' "not relevant" decisions to ensure that the review and analysis of the data is accurate and complies with the discovery order/agreement for voluntary discovery, and that categories are correctly assigned. Privilege should be assessed by the solicitors on record, who may need to stand over the process in the event of a challenge.

4.4 Commencing an audit file

It is recommended that a discovery audit file is opened as soon as instructions are received and maintained and updated throughout the proceedings. All steps undertaken and decisions made in the discovery process should be recorded in this file, including:

- how the list of potential custodians and data sources was compiled;
- all information regarding the legal hold, when it was put in place, and the timing of scoping and collation;
- any considered decisions arrived at, e.g. about excluding time periods, sources or parties from the discovery exercise;
- any filters agreed and subsequent changes to them;
- the logic behind decisions on privilege or relevance; and
- any significant issues that arose at any stage and the steps taken to deal with them.

Maintaining an audit file makes it easier to identify why decisions were taken if a decision becomes the subject of scrutiny in court, e.g. if the other party challenges a claim to privilege, why certain documents have not been discovered, the omission of a potential custodian, the sources of data searched or the parameters placed on the collation exercise.

The audit file will enable the party who made discovery to provide a detailed explanation to the Court and the opposing party as to why such an approach was adopted and will assist in evidencing the reasoning behind decisions and the timing of them. Although a court may not ultimately endorse the decisions made, if it is convinced that there was a logical and reasoned motivation for them, it will be less likely to impose sanctions. The audit file will also serve as a record of the work completed should costs be disputed.

Helpfully, most discovery software also creates an audit trail, with tailored "tags" to track what members of the team and the client have reviewed, queries on documents and decisions reached and by whom.

There is no entitlement for the other parties to the matter to have access to a solicitor's discovery audit file, which will be protected by litigation privilege and, where applicable, legal advice privilege.

4.5 Preparing discovery plan and budget

A discovery plan should be prepared in conjunction with the discovery project team[6]. To assist in this, detailed instructions should be taken from the client to obtain an understanding of the client's data storage systems. Preparing the discovery plan will also assist in preparing a budget for making discovery.

The cost of retrieving documents may vary depending on the class of data and/or the complexity of the system. For example, it is usually more expensive to retrieve inactive, residual or legacy data and consideration should also be given to whether the documents are available from a more accessible active source, such as substantially duplicative backup.

The discovery project team should be aware that even the most sophisticated organisation can use systems that have come into existence piecemeal and an information system may not be used uniformly across the organisation. This is especially true where the organisation is international, with other sites and subsidiaries with different systems in different countries. Acquired subsidiaries may have legacy systems that may also be in scope.

The extent to which any given data source may contain relevant documents is a question of judgment to be considered in conjunction with the client. When scoping discovery with their clients, whether it is

[6] *A template/sample plan has been provided at Appendix G. See also Chapter 6 - "Planning to make discovery".*

necessary, reasonable or proportionate to collect each data source will depend on the likelihood that the data source contains documents relevant to the issues in dispute. It may be advisable to run sample checks where it is unclear whether a given data source is required. All decisions regarding the collection and/or exclusion of specific data sources should be appropriately recorded in the audit file.

This Guide recommends that if any organisation is engaged in a "data cleanse" or "data governance programme" in the course of its business it is good data management to keep a contemporaneous record of the steps taken, details of the nature of any data deleted and the reasons why. This will be useful to refer to if such data would be relevant in litigation that arises later.

Documents may be stored in a number of formats:

- **Native documents** are those in the electronic or hardcopy format in which they were created and maintained.
- **Near-native documents** are those which need to be converted to a different electronic format to allow them to be managed as individual documents. For example, emails stored in a database or mailbox are typically extracted and converted into individual documents for each message (e.g. MSG, MHT or MHTML files). This would also include WhatsApp exports.
- **Near-paper or image format** where a native document is rendered into a picture of itself in a non-editable electronic file. Essentially a "picture" of the document is taken as it would exist if it were printed to paper. Depending on the print settings of the document or computer, and on the nature of the file type (e.g. tif, pdf) information can be lost or altered through the process.
- **Paper documents** are those originally created (e.g. handwritten) on paper, or electronic documents printed to paper.
- **Dynamic data** which is an organised collection of structured information or data which will change in real time as more information is added to the relevant database (e.g. financial data logistics information, flight tracking data).

It is important to note that the variety of data types and potential locations is constantly changing. Practitioners should keep up to date with how their

clients tend to store data. Appendix C contains sample questionnaires which may be used to identify potentially relevant data types and sources.

It should be established early on whether the potentially relevant data set will contain foreign language documents as this must be factored into decisions regarding how this data will be dealt with in the discovery (e.g. if specialist translation technology or services are required for the purposes of reviewing for relevance).

It should also be established whether there are likely to be large volumes of highly technical documents that may need to be reviewed by relevant experts.

If there is a large volume of data containing relevant information, which may require redaction, this will increase the costs of making discovery and should be factored into budgeting.

4.6 Good Document Management

Organisations that manage data in an organised and efficient manner will find it easier to deal with the burden of making discovery, particularly in the early stages of identification, preservation and collection. If an organisation has a good understanding of what data it holds and where, potentially relevant information will be easier, and less costly, to identify and retrieve. This is the essence of good information governance and the foundation of efficient data management.

The broader topic of data management is beyond the scope of this Guide but there are procedures an organisation may establish and maintain to improve its data management in relation to discovery:

- a data classification policy and process, whereby different types of data are managed based on their priority to the organisation
- only retain data for as long as it is required by the organisation and any regulations, and then dispose of it at the end of its useful life
- a data map which details the different types of data managed by the organisation, to list all systems, the type and classification of the data they hold and where, ownership and access requirements
- a process for retaining and accessing data from historical systems and previous employees, including access to encrypted data

- a standardised format for naming conventions for all documents which should be implemented across the organisation

- a procedure for labelling potentially privileged communications when generated, which will greatly assist with the identification and ring-fencing of potentially privileged data.

There are a number of stakeholders that are key to achieving efficient data management and information governance in every organisation. They will typically include the Head of IT and Chief Information Officer or members of their teams. Where no formal information governance policy is in place within an organisation the completion of a discovery exercise can be a useful starting point. When considering good document management, parties should be conscious that additional costs incurred during discovery due to prior or existing poor document management may not be recoverable as costs in the matter.

4.7 GDPR

It is likely that any party to litigation who holds data of relevance to a discovery request will be a data controller for the purposes of data protection law. Data controllers have a range of obligations under data protection law, and in particular must comply with the principles of data protection, as found in Article 5 of the GDPR, ensuring personal data are: processed lawfully, fairly and transparently; processed for specific purposes; limited to what is necessary; kept accurate and up to date; stored for no longer than necessary; and protected against unauthorised or unlawful processing, accidental loss, destruction, or damage. Controllers must also be able to demonstrate compliance with these principles, under the principle of accountability.

In addition, controllers must have a legal basis to process personal data. Article 6 of the GDPR provides six possible legal bases for processing personal data, including, under Article 6(1)(c), where processing is necessary for compliance with a legal obligation to which the data controller is subject, or Article 6(1)(f), where the processing is necessary for the legitimate interests of the data controller or a third party. In addition, under s.41(c) of the Data Protection Act 2018, personal data can be lawfully processed for a purpose other than that for which it was collected, including where "necessary and proportionate" for the purposes of legal proceedings.

As such, while personal data can be processed and disclosed for discovery purposes and in the context of legal proceedings under data protection law, it

must be done lawfully, fairly, and transparently, with a valid legal basis, such as compliance with a legal obligation or for the legitimate interests of the company or a third party. The specific legal basis for processing the data should be identified, and the processing must adhere to the principles of data minimisation, accuracy, and purpose limitation.

Recent case law from the CJEU (C-268/21 - *Norra Stockholm Bygg*) establishes that when assessing whether to order production of a document containing personal data in the context of court proceedings, the national court must have regard to the interests of the data subjects concerned and must balance those interests according to the circumstances of each case, the type of proceedings at issue and consider also the principle of proportionality and the principle of data minimisation (as referred to in Article 5(1)(c) of the GDPR). Consequently, when considering a request for discovery, a party should consider the extent to which the request will encompass personal data and have regard to the interests and rights of the data subjects concerned when deciding how to properly respond to that request.

It will also frequently be the case that a client which has received a request for discovery from an opposing party in litigation may also have previously received from that same party a Data Subject Access Request under Article 15 of the GDPR and/or a Freedom of Information Request under the Freedom of Information Act, 2014. In such circumstances, the client should ensure consistency between its responses to such previous requests and its response to the request for discovery.

It also may be that in accordance with its obligations with the GDPR's storage limitation principle (Article 5(1)(e)), data is no longer held which may be of relevance to a dispute. If so and it is clear that the data was deleted prior to litigation being threatened or contemplated this should be set out and explained with reference to the "deleted" data, in the Second Schedule of the Affidavit of Discovery.

4.8 Personal devices

An order for discovery encompasses data which is stored on the smartphones and other devices which the employees of the party against which the order is made use to communicate. Even where the employer has provided its employees with work specific devices, there will be circumstances where the employees have used their personal devices for work purposes (e.g. power outage on the work device) and compliance

with an order for discovery requires that data on employee personal devices be reviewed to identify potentially responsive material.

Prudent employers will have a policy dealing with the use of personal devices in connection with the employer's activities. Employees, agents or ex-employees of the client should be contacted and asked to provide access to potentially relevant data held on any personal devices. If an independent e-discovery contractor has been engaged they should be in a position to search the device, copy the relevant material and the device can then be returned to the impacted individual. This Guide recommends that if an employee refuses or cannot surrender his/her personal device he/she will meet in person with a representative from the legal team (and if appropriate someone from the external forensic provider) to go through the device and identify and obtain a copy of relevant data for inclusion in the discovery.

The Guide recommends that where possible, only work devices should be used for work purposes. However, if personal devices are used for work related purposes there needs to be a clear and understood policy in place for the preservation and collection of relevant data from personal devices. This policy will need to be continued for a number of months post any employment period.

4.9 Document Security

Parties taking part in the discovery process, and especially legal advisors and third party providers engaged to manage data on behalf of their clients, should ensure that appropriate security controls are in place at all times to protect data. This should include, as appropriate, access controls, encryption while in transit (and sometimes when at rest), and the secure disposition of data as soon as it is no longer required.

4.10 Discovery project plan

For parties and legal advisors who are frequently required to make discovery, it is prudent to prepare a response plan for the management of discovery projects. The following areas should be addressed in a tailored discovery response plan:

- Appointment of a Discovery Team
- Notification and initial assessment – This details how a new matter might be notified to a party and what information needs to be gathered in order to make an initial assessment. It also includes details of who will make the initial assessment of the matter and sets out reporting lines.
- Approval process – This details the approval process required within the organisation to decide upon and approve next steps, which will usually include approving the litigation hold.
- Tailored litigation hold process – This should contain a tailored version of the litigation hold process outlined in Chapter 7 – "Preservation", including common technical measures used by the organisation.
- Data map – An up-to-date data map for the organisation should also be maintained.

Chapter 5 Identification of the likely scope of the "universe" of relevant documents for the purposes of preserving and collating relevant documents

5.1 Introduction:

Where the nature of the dispute between the parties is such that the parties can reasonably anticipate that they will receive requests for discovery of categories of relevant documents, they are required to take steps to put themselves in a position where they will be able to comply with their discovery obligations within a reasonable period after the terms of their discovery obligations have been determined, whether by court order or agreement (*Thema International Fund plc v HSBC Institutional Trust Services* [2011] IEHC 496).

Thus, where practitioners can reasonably anticipate that their clients will face requests for discovery, they should take steps to advise their clients on the scope of the "universe" of relevant documents which their opponent's request for discovery will likely cover. Generally, this will involve identifying the relevant periods of time covered by the material facts of the dispute; the likely relevant custodians; and the likely sources of relevant documents in the client's own records. The obligation is on the parties to preserve such documents.

5.2 Identification of relevant time period

Parties are only entitled to request discovery of, and by extension are only obliged to disclose, documents which are relevant to the matters at issue in a case. Whether a document is relevant to a case will be determined by the nature of the disputed issues in the case. Once the disputed issues are known, it should be possible to identify the time period during which the facts giving rise to the dispute took place. Consequently, it will be necessary for the party making discovery to determine the time periods it will apply when searching for documents potentially responsive to the discovery request. From the perspective of the requesting party, and with a view to avoiding disputes later as to the completeness of the searches undertaken, it can be helpful (where possible) to specify, for each category of documents of which discovery is sought, the time period which should apply.

In the initial stages following instruction, a cautious approach should be adopted in assessing the date range to be applied to the preservation and collection exercises. Parties should be made aware that while a dispute may have taken place at a specific point in time, how parties subsequently addressed it could equally form part of a discovery category request. All documents touching upon the matter which is now the subject of the dispute should be preserved.

It will always be possible at a later stage to refine the date range, particularly following receipt of further particulars of the claim. It may also be possible to reach an agreement with the other party regarding the time period to be applied to the discovery categories. However, it may also transpire that at a later stage in the discovery process, the relevant time period may need to be expanded. This might occur in circumstances where it becomes clear from the discovery provided that there are documents of assistance and relevance which pre-date the date upon which disclosure was agreed and could be the subject of further discovery. Therefore a party should be very slow to dispose of documents which could potentially fall within the scope of the litigation unless and until the litigation concludes.

Through the negotiation of the scope of categories of discovery, it may be possible to agree an appropriate end date for the discovery. Given that the dispute the subject of the case will have crystalised prior to the commencement of proceedings, it may not be necessary to individually collate and list documents which came into existence after the proceedings issued. However, this should be clarified with the other party as unless a specific order is made, party is required to individually list all responsive documents which were generated prior to the date on which discovery is made. Relatedly, absent a specific request and reason, there is generally no requirement to continue to make discovery after the filing of an Affidavit of Discovery[7]. Such an obligation does arise however, in circumstances where relevant documents were only located after the Affidavit of Discovery was sworn. Such documents must be disclosed in a Supplemental Affidavit of Discovery.

When considering preservation obligations, it is important to remind your client that even if it is possible to maintain a claim of privilege over a document, this does not relieve it of the obligation to preserve and ultimately discover the document by listing it in the relevant schedule.

[7] *Where a party is ordered to make discovery, Order 31, Rule 13 RSC requires that this is done on affidavit made out in the format required by Form 10, Appendix C RSC. The proper title of this Affidavit, as set out in Form 10, Appendix C is "Affidavit as to Documents", although in practice is almost universally referred to as an "Affidavit of Discovery".*

5.3 Identification of relevant custodians

For the purposes of preservation and ultimately compliance with the discovery request, clients should be asked at an early stage to identify and prepare a list of custodians (individuals) who may hold (or did hold) data relevant to the matters in dispute. Such a list can be created by speaking to key witnesses and others who may have an involvement in the matter. It is also essential to determinate how and where these individuals stored this data e.g. on personal devices, mobile phones, laptops, cloud servers, etc. A sample custodian-data source map is contained as Attachment One to Appendix G.

It is important to remind parties that they should consider existing and former employees when drafting a list of custodians. When seeking to identify relevant custodians, in addition to those custodians that engaged directly with the other party to the dispute, their co-workers or members of the team that worked on the matter, along with any persons they reported to or who reported to them, should also be considered. Furthermore, personal assistants and secretaries or other support staff to key custodians are often involved in issuing and amending documents and such personnel should also be considered when compiling the list of custodians.

Even if an individual is no longer working for the organisation it may be that a laptop/desktop/other device used by him/her is still in use within the organisation. In such circumstances steps should be taken with the IT department to identify and/or establish if any data of potential relevance to the dispute can be retrieved, such as ex-employees' email accounts.

Furthermore, if there are grounds for believing that a former employee may hold documentation of relevance on their personal devices or computers, or in their files or papers, he/she should be contacted and requested to preserve those documents and to provide them for the purposes of making discovery. It is also important to clarify whether employees, even if contrary to company policy, use personal email addresses or computers to store documentation of relevance to the organisation and this should be clarified in early course.[8]

Second, steps should be taken to identify whether data could have been stored under a username other than those used by the custodians who have been named on the list. For example, data could be located in mailboxes set up for a

[8] *See Chapter 4, para. 4.7 "GDPR", for discussion on necessity of considering whether a request for discovery will encompass personal data and to have regard to rights of the data subjects concerned when deciding how to properly respond to that request.*

particular purpose (e.g. to deal with a specific project, to address complaints, or to deal with invoices) which were used by or were accessible to relevant custodians.

Third, anything which is held by the servants or agents of a party which could be of relevance to the dispute is within the power, possession and procurement of that party. Accordingly the servants or agents should be requested at the earliest possible opportunity to retain all of the documentation held by them which could be of relevance to the subject matters in dispute. This would involve contacting professional advisors or service providers who were engaged by the party and depending on the corporate structure of the group may also require affiliated or connected companies to be contacted.

It may be helpful to prepare interview question templates with matter-specific questions when interviewing potential custodians to ensure all relevant issues are dealt with and addressed. This is particularly important if you are dealing with an ex-employee who may not appreciate multiple engagements.

It is likely that the identification of custodians, the identification of document sources and the legal hold process will have to be carried out in parallel and/or in a number of different iterations, as new custodians and data sources are identified which need to be included in the process.

If the document identification questionnaire (Appendix C) is used and responses received, this would be of benefit in demonstrating the efforts made to properly identify all relevant documents and for audit purposes should an issue arise as to the completeness/integrity of the discovery made.

5.4 Identification of likely types and sources of data

Once custodians have been identified, steps should be taken to identify the likely types and sources of data which will likely be responsive to a request for discovery.

Data, whether in electronic or hard copy form, can be stored in any number of locations. To avoid disputes arising in relation to the completeness of a party's preservation efforts, it is critical that parties are advised that they must understand how and where the data is located in their organisation.

In collating information, steps should be taken to ensure that data is identified in each potential source which is utilised by each custodian who may hold data of relevance to a dispute. While most organisations store their data centrally on

servers (cloud/ on-site), employees may also have saved down data to their specific devices which will include desktop computers, laptops, and mobile devices such as smartphones, tablets and PDAs.

Custodians may also have interacted with an AI tool in respect of matters relevant to the dispute and all inputs to and outputs from that AI tool should be identified.

Investigations should be undertaken with the IT function within the organisation as to what the position is before directing retrieval from individual document sources, though at preservation stage it is advisable to adopt as comprehensive an approach as possible. An explanation as to the configuration of the system should be sought so that it can be averred to as necessary in an Affidavit of Discovery should it be appropriate or required.

Furthermore it may be necessary to restore data which has been archived either manually or by way of backup media.

It is extremely important that at an early stage in the process and prior to any data being destroyed and/or deleted and/or overwritten that steps are taken to identify all potentially relevant data in dispute and to preserve same.

Recordings of relevant custodians' audio/video calls may exist, in which case these also need to be scoped for potential relevance. (See Appendix D for suggested approaches to managing audio and video data through the discovery process.)

While almost all business information is now stored in electronic format, most discovery projects will involve at least a small element of traditional hardcopy or paper documents. For example, participants in meetings may have taken handwritten notes of the discussions. These will typically be lower in volume than electronically stored information.

If there is a large volume of hard copy data, it may be more efficient to conduct a first pass hard copy review at the relevant storage location in order to identify which documents may contain relevant information.

These relevant documents can then be scanned into electronic format and then included in the review process alongside the electronically stored information. This should include making the scanned data searchable (e.g. through Optical Character Recognition/ OCR), and may also involve manually extracting

information regarding the contents of the data manually for the purposes of creating appropriate schedules to the Affidavit of Discovery.

When managing hardcopy data, it is important to ensure that the family relationship between data is retained, in addition to the ability to sort the data in its original order. This can be vital to the review process where data does not have a date. The suggested identification checklist at Appendix C, with sample client letter, can be used as a guide to identifying potential sources of data with clients at an early stage in a matter.

Where it is unclear if a data source, such as backup tapes, may contain data of relevance to the matter, it may be prudent to undertake sampling of a statistically relevant portion of the data source in order to identify the volume (if any) of relevant documents contained therein. The results of such a sampling process can be used to determine the likelihood of uncovering further relevant documents should the full document source be included in the process.

Data may be archived by way of technology which becomes obsolete and it can be expensive and difficult to retrieve such ”inactive” data. It may not be necessary to retrieve all of this information and ultimately it will come down to a cost benefit analysis which weighs the likelihood of relevant data being retrieved against the costs that would likely be incurred in the retrieval process.

Consideration should also be given at this early stage as to whether foreign languages will play a significant part in the documents subject to discovery as it may be necessary to include foreign language terms in key word searches for the purposes of identifying potentially relevant documents at source.

5.5 Specific issues including accessibility/retrieval (cross-border or technical)

In making discovery, a party is obliged to discover all data, which is in its power, possession or procurement and which is within the scope of the discovery categories. As stated above, in certain circumstances this could include data held either by companies affiliated to or associated with the party making discovery which is a party to the matter. It might also require documents which are held by agents and/or representatives of that company who are based in different jurisdictions. Therefore although an Order for discovery does not have extraterritorial effect, in that it is not legally enforceable outside the jurisdiction, if a party to litigation in Ireland is ordered to make discovery and if that entity is legally entitled to require another entity which is based outside the jurisdiction

to provide certain documents to it, these documents must be discovered by the Irish litigant. The order for discovery is enforceable in Ireland against the party which was ordered to make discovery.

An issue may arise where such information is held in a jurisdiction in which the data protection or national privacy laws intentionally or inadvertently constrain such disclosure. Accordingly, where data is held in other jurisdictions, care needs to be taken to ensure a party would not contravene local legislation by releasing such documents to comply with Irish discovery obligations. This is not to say that such an impediment could be relied upon by a litigant as a means of shielding the production of such data. Courts will look very carefully at any such claims and will seek to identify, at a minimum, that no data was transferred to such locations after the parties became aware of the potential dispute.

Chapter 6 Planning to make discovery

6.1 Introduction

At the same time as the scope of the likely universe of relevant documents is being determined, a full discovery plan should be drafted. In complex cases, the discovery plan should be prepared before work begins on identifying the relevant document universe.

A template/sample plan has been provided at Appendix G. It is important to note that this is a document that will evolve throughout the project. At its inception, it will be likely to only contain the steps taken in the identification and preservation phases, and the remainder of the document will set out the intended steps for the remainder of the project. As each phase of the project is complete, the plan should be updated to reflect what was actually completed. At the conclusion of the project, the plan should be updated to reflect all steps actually taken during the project.

In drafting the plan, it will be necessary to assess the time periods potentially required to complete the project, prepare a budget, and start to consider the use of technology to facilitate the process.

6.2. Assessing the time required for retrieval and review

It is essential that an accurate estimate of the time it will take to make discovery is established to inform the overall timetable for discovery and to seek a realistic period for completion from the Court. In arriving at a reasoned estimate of the time it will take to complete the discovery project, practitioners should have regard to each of the phases that are involved in a discovery project (e.g. Identification, Preservation, Collection, Processing, Review etc) and consider how long it is likely to take to complete each of those phases[9]

The time involved for review will depend on the amount of data identified as being potentially in scope and the extent to which Technology Assisted Review ("TAR") is used for the purposes of reviewing potentially responsive data.[10] However, even where TAR is used, it will still be necessary to establish a review team to carry out manual reviews of the data. The size and scale of the review team (and the technology used) will dictate the amount of time that will be

[9] *A summary of the phases that make up a typical discovery project is contained in Chapter 3 - "Outline of discovery phases). See also Appendix A - "Discovery project checklist".*
[10] *See Appendix F - Technology Assisted Review for overview of TAR.*

involved in the review process. It is difficult, if not impossible, to accurately estimate how long a discovery project may take in advance of the Identification (Chapter 5) phase being completed. Until the completion of the Processing phase (Chapter 9), it will be very difficult to give a precise estimate of the actual volume of data for review.

6.3 Budgets

Preparing and updating a detailed budget is essential to monitor costs and ensure they stay within limits. Budgets should cover each stage of the discovery process and include provisions for diverse disciplines involved. A global budget, refined as the process progresses, helps manage overall legal expenses and ensures proportionality.

It is important to have an accurate assessment of the costs that will likely be involved in addressing each phase of the discovery process as this will enable the party to have a proper understanding of the extent of the burden that complying with each phase will impose and, potentially, to make arguments to the Court that proportionality requires that a particular approach be taken. Regular monitoring and client approval for additional costs are recommended.

When seeking eDiscovery services, obtain comparative quotes. Experienced providers can offer accurate estimates for typical projects. These estimates should cover the number and location of custodians, data volume, collection costs, processing time, review access, quality controls, ongoing hosting, and archiving costs. Identify all project aspects in the estimates and monitor for potential additional costs.

Chapter 7 Preservation

The objective of the preservation phase is to take steps to preserve data where it exists, so that it may not be altered or destroyed in advance of collection. This includes the legal hold process.

7.1 Legal hold process

As stated in para. 4.2, clients should be advised at an early stage of the duty to preserve data which may be of relevance to the dispute and to suspend routine or automatic data destruction processes. This is best achieved by putting in place a "legal hold", i.e. informing all of the relevant personnel, in writing, of their obligation to preserve all data that may be relevant to the dispute. All actions taken to preserve data, and actions not taken, should be fully documented, along with the reasons why.

A legal hold should describe the nature of the dispute and identify the types of data, including all electronically stored information and hard copy documents which may potentially contain information relevant to the issues in dispute and therefore would potentially fall within the scope of a subsequent discovery request. The legal hold will also assist in preserving documents that the client may wish to rely on.

It should be clear from the legal hold that it applies to all data, of whatever type and wherever stored. A legal hold should specify that it includes all data, including all electronically stored information and hard copy documents. A non-exhaustive list of examples of electronically stored information and documents which should be covered includes e-mails, WhatsApp message, Teams Messages, Word documents, spreadsheets, pdfs, voice mails, recordings of telephone conversations, calendar entries, handwritten notes, and all inputs to and outputs from AI tools. All such data, whether stored locally on personal or work devices, on office servers or on personal or work cloud-based storage facilities, or in hard copy files or notebooks, should be covered.

The legal hold should be addressed to personnel involved in the activities that are relevant to the dispute and also to the IT personnel or service providers of an organisation. Each party should be directed to suspend the destruction of and hold related data until such time as the legal hold has been lifted.

A record should be kept of the individuals to whom the legal hold has been sent and each party who has received the legal hold should be requested to

send either an email or to sign a document, confirming receipt and acknowledging that he/she has reviewed the legal hold, understands it and agrees to comply with it.

Given the likely duration of litigation, it may be advisable to issue periodic reminders of the legal hold and/or to modify the hold if it becomes apparent that the scope of the proceedings and/or all relevant information has expanded or indeed narrowed, (though any narrowing should be done with extreme caution).

People will join or leave an organisation during the lifetime of proceedings. Clients must be advised of the need to inform new arrivals of the presence of the legal hold. In respect of leavers, consideration should be given to the likelihood of potentially relevant information being held on the leaver's personal device(s) and if appropriate steps should be taken to secure that data before their departure.[11] At a minimum, contact details for leavers should be obtained to ensure enquiries can be made of them should the need arise.

It is only when the proceedings are at an end and all potential avenues of appeal have been exhausted that consideration should be given to releasing the legal hold.

The format of a legal hold notice will be fact specific; however suggested draft legal hold notices are appended to this guide at Appendix B.

Note: In some matters, it may not be possible to immediately identify the relevant parties to include in the legal hold process. Where there is a risk that data could be lost before the identification of the specific parties is complete, consideration should be given to issuing a broad legal hold notification to all personnel in the organisation alerting them to the need to preserve data until specific custodians have been identified.

7.2 Technical preservation

The legal hold process outlined above implements a 'soft' control in that it relies on custodians and other individuals not taking actions to alter or destroy data. It is therefore prudent to take additional steps to preserve data where they reside in the event that one or more custodians fail to act on the legal hold instruction.

[11] *See Chapter 4, para. 4.7 "GDPR", for discussion on necessity of considering whether a request for discovery will encompass personal data and to have regard to rights of the data subjects concerned when deciding how to properly respond to that request.*

In the first instance, relevant systems which automatically destroy data should be suspended.

Technical preservation steps will be highly dependent on the data sources identified, their location, accessibility, and the capability of the available technology in place in the organisation. Steps may include:

- The most recent backups of email servers/file servers/application servers may be removed from backup rotation and stored securely.
- Technical controls may be implemented which prevent custodians from altering or deleting historical data.
- Access to hardcopy documents may be restricted to the discovery/litigation team only.
- Documents that are routinely updated, such as excel documents, should be identified so that they are preserved as at a relevant date and time, such that they are not overwritten at a later date.
- Certain types of data should be collected at an early stage to ensure it is preserved e.g. data from portable devices such as mobile phones that are not otherwise backed up. Such devices can either be lost or data on them may not be backed up and therefore the data itself is more likely to be lost. Further, it can be difficult to obtain a copy of such data where an employee leaves the company, therefore, it is best to move quickly to preserve such data.

As with all steps taken in the discovery process, they should be fully documented in the audit file.

Note: Throughout the preservation phase parties and custodians must be instructed not to search for, access, or move any original data. Such access will likely alter the original data (and underlying metadata) and may cause the reliability of such data to be disputed at a later date. Any collection and searching of data should be carried out by appropriately trained personnel who will employ methods to protect the original data throughout the process.

Chapter 8 Collection

Once the data sources have been identified, and preservation steps have been taken, the next step is to obtain a copy of the data sources, or selections of data from each source, so that they can be further processed and reviewed.

This typically involves working with the custodians, their IT provider, and any third parties who maintain custody of the data. As data sources are collected, the custodian-data map should be updated with the collection status.

8.1 Practical considerations

8.1.1 Where will the copying take place and by whom?

In general, it is possible to extract electronic data sources remotely without the need to attend onsite. In certain instances, it may however be appropriate to attend on site to collect certain categories of documents, such as potentially relevant data held on personal devices (e.g. WhatsApp messages), data stored locally on devices and hard copy documentation.

Who completes the collection is another important decision. The entire process of data collection should be supervised by the solicitor and, generally, the actual collection itself should be conducted by the solicitor, with the assistance of a suitably qualified IT specialist. If self-collection by custodians is under consideration, three key risks should be addressed. The first is whether the custodians themselves can be relied upon to complete an accurate unbiased collection. The second is whether the custodians and/or their IT team have the technical skills and tools to complete a collection without altering the original data. Thirdly, there is also the risk of varying opinions as to what is potentially relevant. For these reasons, in Commercial Court litigation it is very rare for clients to carry out their own document collection process and practitioners should take great care in allowing the custodians, and/or those close to the matter, to undertake the collection process. It may be more cost effective to engage the services of a data collections specialist and/or an IT team (who may be internal to the organisation or external specialists) with the necessary tools and skills to complete the collections independently of the custodians. In complex matters, the option of having the process carried out by an independent specialist should be given due consideration.

This Guide recommends that collection should be completed by or with the assistance of a suitably qualified IT specialist.

8.1.2 What is the scope of the collection?

In general terms, it may be more cost effective and less disruptive to a client's business if the whole data source is collected and then filtered later for potentially relevant data.

However, given the increasingly large volume of data held by organisations, it is often more efficient and practical to adopt a focused approach which involves extracting specific sub-sets of data sources. For example, if there is a single project folder containing all data related to the disputed project, then it may be far more efficient to only take a copy of this folder rather than the whole data source (full computer server) which may hold folders for thousands of irrelevant projects. However, caution must be used when narrowing the focus of a collection at an early stage to take into consideration the probability that data may have been stored at other locations. Decisions as to which sub-sets of data sources are relevant to the dispute must be taken in conjunction with the legal advisor who should supervise the process.

Generally, it is common practice for custodian-based data sources to be copied in full (such as email mailboxes, private network folders, and laptop/desktop computers), whereas non-custodian-based data sources tend to undergo a focused collection (such as large server computers).[12]

8.1.3 What type of collection is required?

In collecting the client's data, one of the principal objectives is to secure a forensically sound copy of the data as it was stored on the date of collection. This is necessary for the purposes of ensuring the integrity of the discovery process as a whole and may also assist in if the admissibility or validity of the data is later questioned.

There are a number of industry standard tools which are freely available which are capable of collecting data in a forensically sound manner. As with all tools however, there is a level of expertise required to use them effectively and help ensure that the original data and metadata is preserved throughout the copying

[12] *See Chapter 4, para. 4.7 "GDPR", for discussion on necessity of considering whether a request for discovery will encompass personal data and to have regard to rights of the data subjects concerned when deciding how to properly respond to that request.*

process. Consideration should also be given to the use of encryption to secure copies of data. As outlined at 8.1.1 above, it is recommended that appropriately skilled personnel be engaged to complete this process. Such personnel will be adept in verifying the accuracy and completeness of documents which have been collected.

8.1.4 Hardcopy documents

The collection of hardcopy documents differs from that of electronically stored information in one major aspect. The very physical effort of identifying and preserving hardcopy documents usually requires that an element of review be performed to determine if the information contained in the documents is likely to be of relevance to the matter. Because the moving of hardcopy documents for scanning and coding requires significant effort, a level of effort is required to filter the documents prior to this.

As such, hardcopy document collections tend to be focused and reduce the volume of irrelevant documents up front. This is the opposite of how electronically stored information is managed, which is typically more efficient to collect en masse and then filter later.

Once the potentially relevant hard copy documents have been identified, they should be scanned where possible into electronic format, made searchable where possible through an Optical Character Recognition/ OCR process and the metadata extracted through a manual coding process. They should then be included alongside any electronically stored information throughout the discovery process.

It is important when managing hardcopy documents that the ability to reconstruct their original order and family groupings is maintained throughout the process. In particular it is important to preserve the source information, including the custodian, the box description (if applicable), any box ID and any details as to the office from which the box or files came.

As with electronically stored information, the process of identifying and collating hardcopy data should be supervised or led by a legal advisor.

8.1.5 Social/professional media

It is becoming more frequent that data and/or messages of relevance are stored on social media websites platforms, such as LinkedIn. This data is typically

accessible by viewing it on the platform in question, however it can require certain permissions to access, for example if it has been posted in a private group. There are three common approaches to address this:

- Apply for a court order to require the platform provider to preserve and produce the data. This can be prohibitively time-consuming and expensive particularly where the platform provider is domiciled in a different jurisdiction. While parties may request that the provider voluntarily preserve and disclose the data, in practice providers rarely act upon such requests given the risks of adverse action for breach of data protection rights, breach of confidence, etc.
- Assuming the content is accessible without specific permissions or where access has been procured, utilise the services of a specialist data collections provider who can acquire a copy of the content and verify its authenticity as a copy in Court if required.
- Again, assuming the content is accessible without specific permissions or where access has been procured, simply take a screenshot of what is displayed, print and sign it in the presence of a witness (preferably a solicitor), and present it as a verified copy of what was viewed at a point in time.

Access to view and copy information from websites, unless publicly available, should not be procured through deception (such as posing as a colleague or friend in order to gain access).

8.1.6 AI tools

Where client employees have used AI tools in connection with the matters the subject of the dispute, all relevant inputs to and outputs from the AI tool in question should be collected.

8.1.7 Chain of custody

It is recommended that a record of the chain of custody for all data sources should be kept during the collection process. A chain of custody record can be vital in helping to demonstrate that no unauthorised access was made to the original verified copies of data. In addition, a chain of custody record can assist from a security perspective as well as an evidential perspective. Completing a chain of custody should add very little in terms of time and cost to the collections process.

Chapter 9 Processing

9.1 Introduction

Processing is the process of removing clearly irrelevant data types from the data set collected and converting the remaining data into a format which will facilitate efficient searching and review. The effective processing of a suitably tailored dataset is crucial to running an effective and cost-efficient discovery process.

Tailoring the dataset will entail the application of filters albeit in a manner that does not unduly narrow the data being collected. The filters will include date range(s), keywords, data source(s) (e.g., email, Google Workspace, Slack, Microsoft Teams, Box/Dropbox, SharePoint/OneDrive, Google Drive, SMS, WhatsApp, AI tools, etc.), custodians, file types, etc. This also provides an opportunity to get a high-level view of the data, and to search for key documents.

The document sources will then usually be prepared for the use of Technology Assisted Review (TAR), including Continuous Active Learning (CAL).

It is generally advisable, both from a cost and burden perspective, to wait until categories of discovery have been agreed or an order for discovery made before commencing processing. This enables the disclosing party to run definitive date filters and search terms on the custodian data pre-collection. Guidance on the appropriate application of filters is addressed below in section 9.5. The risk of processing at an earlier stage is that the disclosing party either collects an excess of data which leads to its incurring excessive processing and storage costs, or collects an unduly narrow data set such that the collection process will have to be re-run at a later stage resulting in potentially significant additional cost and delay.

The extent and nature of processing required in any given discovery project will depend on the nature of the data collected, the technology being used, and the expected review process. The processing phase typically consists of the steps outlined below. Throughout, parties should be mindful of data protection obligations and proportionality, documenting decisions and statistics to ensure that they are in a position to defend their decision-making at a later stage, if necessary.

9.1 Initial Processing: convert data into searchable format

The data set is first loaded onto an eDiscovery platform. Once this has been done, a number of processes are run over the data set.

9.1.1 Removal of irrelevant document types

All non-user created data is removed during the processing stage. While, in general, terms, user-created data is easily identifiable and readily distinguished from non-user created data, unusual categories of user-created data, such as CAD files, which are created by the design software often used by architects, engineers, etc., may need to be identified by custodians in order not to be excluded during processing. It is important therefore to engage early with custodians to discern whether such unusual user-created documents may form part of the dataset.

Where a focused collection of user-created data has been acquired, for example a single folder from a network share, a data type filter will typically not be required. Such focused collections will by their very nature likely only contain user-created data.

For instant messages, chat and other cloud collaboration data, it is necessary to ensure that the context and content are captured and this may require the chats to be sliced into 24 hour periods. The advice of a technical expert should be sought in this regard, where required.

Note: Where not-readily accessible data types have been agreed for inclusion, such as deleted files, these should be recovered at this time and included in the process.

9.1.2 Conversion into searchable format

The discovering party's chosen eDiscovery platform extracts the content of the data, including all associated metadata. This metadata is used to populate much of the detail required in the First Schedule to the Affidavit of Discovery. This is a key benefit to processing as it reduces the manual input that would otherwise be required of the disclosing party and its legal advisors

9.1.3 Deduplication

In parallel with the steps described above, a family-level deduplication process will be run against all data in the set. This suppresses any duplicate families of data while leaving one copy of each unique family of data for further processing. While duplicate families are suppressed, the list of custodians who hold a duplicate family which was suppressed is recorded and included in the remainder of the process. This allows only one copy of the family to be considered, while also allowing the reviewer to quickly understand who held duplicates of the family. The result of this deduplication process is that the volume of data is typically reduced. Statistics as to the number and type of documents in this new data set are typically recorded. Appendix E contains a detailed overview of deduplication and families of data.

Note: It is common and also good practice to deduplicate data families and only produce one copy of each unique data family which is of relevance to the matter. Should a receiving party wish to inspect duplicate data, this may be requested after receipt of production and would typically require justification before the additional cost of inspecting/producing duplicate data would be considered.

9.1.4 Optical Character Recognition ("OCR")

The volume and type of non-searchable data is identified. An OCR process is then run against these documents in order to convert them to a searchable format.

9.2 Thread deduplication

Once the initial processing is complete, email thread deduplication is then run against all emails and their attachments. This process identifies the inclusive portions of each email thread along with the non-inclusive (or duplicative) portions of the email thread. The non-inclusive portions of the email threads are then suppressed and excluded from the documents otherwise presenting for review. Email threads which are unable to be subjected to the email threading process should not be suppressed and should be included in further processing. Appendix E contains a detailed overview of email threads and thread deduplication.

9.3 Manage problem documents

Quality controls at the processing phase must identify problem data which cannot be processed by whichever system is in use. These include corrupt files and encrypted files, amongst many others. Therefore, this data will not be accessible for keyword searching. (Note: their location and names, including metadata would however likely be searchable.)

9.3.1 Encrypted data

When managing encrypted data, there are a number of options. The route to take will depend on the method which has been used to manage the encryption system in place. Where an enterprise-wide encryption system has been used, it is typically possible for the IT manager of the system to provide a mechanism to unlock the data. Where individuals have set the encryption using a stand-alone system, such as simply applying a password to a spreadsheet, then it will be necessary to request the password from them.

In the event that the password/decryption key is not available, then a decision will have to be made on whether to attempt to remove the encryption by other technical means. This typically involves using specialist software. Such software can have a varying success rate, and while it is typically not a costly exercise to undertake, it can take a long time, with very little indication as to when, if ever, it will be successful.

It is therefore best to reduce the number of encrypted files to attempt to decrypt. One approach to this is only attempting to decrypt files which have been highlighted as potentially relevant either through their name, or through association with another file. For example, if an email is deemed to be relevant, however it has an attachment which is encrypted, then it may be useful to attempt to decrypt the attachment. However, if it is just an encrypted attachment to a non-relevant email, in a universe of many thousands of emails, then it may not be justified to incur the time and cost of attempting decryption.

The number of deduplicated encrypted/password protected data in the set should be recorded.

Where encrypted/password protected data are located in a data source which will not be subject to filtering, such as those in a shared project folder (i.e. likely to be relevant), all such data should have decryption attempted. Attempts to access encrypted or password-protected files should be documented. It should

be noted that, given the sophistication of modern encryption, efforts to decrypt are likely to be unsuccessful.

It is generally appropriate for the methodology used to be recorded in the Affidavit of Discovery. If decryption is successful after the delivery of that Affidavit, any discoverable decrypted data will then need to be the subject of a supplemental affidavit of discovery. That is the appropriate course of action rather than delay substantive compliance within the original discovery Order or agreement to make voluntary discovery.

Other classes of problem data are wide ranging. These may include very large spreadsheets which will not be viewable using the proposed review system, or complex technical drawings which may need to be converted to a different format to allow viewing. The approach to managing such data will be heavily dependent on the technology in use. Of importance is to identify the problem data and not inadvertently miss them, and then decide if they should be subject to filtering (which may not work) or brought straight for review. Where documents are stored in collaborative platforms (such as Microsoft Sharepoint or Google Docs), any available versions, comments, and tracked changes should be preserved and reviewed.

9.3.2 Foreign language data

Foreign language data, while not necessarily "problem data" must be identified and a strategy for addressing them established at this time. This is due to the fact that filtering and other areas of the process, such as Technology Assisted Review/ TAR and/or review, will be impacted by the presence of foreign language data (or data which has mixed languages). Keywords will need to be devised in the relevant language (having regard to considerations such as English language words which may have a small number of synonyms but which may have a multitude of synonyms in a foreign language), and reviewers with appropriate language skills will need to be engaged.

Solicitors should identify documents in foreign languages and arrange for translation where necessary. Advances in technology including the rapid development of Generative AI (GenAI) means that there is an abundance of technology-based translation tools available. The use of these tools on foreign language documents as an initial step can be an effective way to assess their potential relevance.

9.4 Apply filters and perform early data assessment

Filters, such as search terms and date parameters, are frequently used as part of the extraction process, prior to processing. Care should be taken to ensure that the terms used are sufficiently broad at this stage to capture potentially relevant documents to the categories, as the extracted data set can be further refined post-processing.

Such further refinement may result in additional search terms being applied, exclusionary search terms being applied to remove false positive documentation, and other steps undertaken to target a rich data set, i.e. one that is designed to return documents responsive to the categories that have been agreed or ordered. Throughout this process and the data review, the adequacy and appropriateness of any filtering criteria used to extract the data ought to be kept under continual review. Further extractions should be conducted at source if it becomes apparent during the review that there are reasonable grounds for believing that additional search terms, custodians, and/or date ranges ought to be applied to extract relevant data, or the review indicates that a further relevant document location exists.

Using the discovery plan at Appendix G, the filters which were used can be listed in Attachment Three and the total volume of data for initial manual review recorded.

It is important to test the processed data to ensure that relevant data has been captured, and forms part of the review set. The objective of this testing is to ensure that the filtering criteria will highlight for review the data which is likely to be of relevance to the matter, whilst also managing the volume of irrelevant data for manual review. Testing can be done in a variety of ways – from reviewing metadata schedules, using analytics tools such as cluster wheels (which conceptually group the documents), in addition to sampling the results of documents responsive to particular search terms. Timeline analysis charts should be used to identify any significant gaps in the data collated. Testing can assist in ensuring that the review seeks to target the most relevant data from what has been processed in a structured manner.

A Search Term Report is often produced which contains details of the combined number of documents which are responsive to one or more of the search terms. To understand the totality of what may require review arising from documents responsive to search terms, it is also important to understand the total responsive documents together with their families. This information is required in order to fully understand the volume of data for review and will be an

important factor in determining how the review will be approached and managed, and if Technology Assisted Review/ TAR (including Computer Assisted Learning/ CAL) is an appropriate technology to employ. It may also be appropriate at this point to conduct sampling of included and excluded documents to ensure that the filtering process is adequate.

9.4.1 Using dates as filtering criteria

It may be possible to identify and exclude wholly irrelevant data types, or to confine the discovery review to a specific date range and thus exclude data falling outside the date range. Where e-mails are concerned, the relevant date is the date of the e-mail family (i.e. the date it was sent/ received). This filtering should be done at the outset. Care should be taken to ensure that document dates have not been corrupted in any way in the course of the processing, which can occur and may require specific IT input to correct.

9.4.2 Use of key words

As noted above, in practice, there are likely to be two stages during which keywords are applied, a broader set of keywords applied at collection and a more refined set (or combinations) of keywords applied to the processed dataset in order to exclude false positives and prioritise key documents for review early on. Developing effective keywords is complex and requires specialist input, from a legal professional and/or an IT specialist experienced in developing and applying keywords. It is an iterative process, so that the results should be monitored to identify and exclude any obvious false positives, while ensuring that care is taken to identify potential gaps. Overbroad keywords will result in a very large review set with a significant proportion of false positives, while overly narrow keywords risk missing relevant data. Keywords also need to anticipate potential spelling errors. The legal adviser must take care to ensure that keywords are not misspelt when entered.

One approach is to split keywords into those which identify parties and those which identify issues. 'Parties-based' keywords may be used to identify specific parties in a data set. For example, the producing party may search for all data which reference the requesting party. This would return all data within the producing party which reference the requesting party. This is useful if the only communications the parties in question ever had was regarding the matter in dispute.

Where parties-based keywords return a large volume of irrelevant data (typically due to the parties conversing/transacting on a number of matters), then it is

necessary to use issues-based keywords to narrow the search. Issues-based keywords are typically combined with parties-based keywords and focus on the specific topics which are the subject of the dispute. For example, the name of the project, account numbers or project/property addresses/locations.

By applying parties-based keywords first and then narrowing using issues-based keywords, the producing party can refine keywords on a step-by-step basis, testing each iteration.

There are a variety of ways to combine parties and issues-based keywords (and also parties and parties keywords, and issues and issues keywords, or any combination possible). These include Boolean operators such as AND and OR, and for excluding false positives, NOT. For example, searching for "John Smith" AND "Central Bank" may bring back a vast number of false positives whereby John Smith's email signature contains the text "Regulated by the Central Bank". Proximity searching may be employed to search for "John Smith" AND "Central Bank", but not where "Central Bank" is located within three words of "Regulated by".

In matters where it is unclear if spelling is correct, it is possible to use wildcards in place of letters. For example, "John Smith" might return both "Smith" and "Smyth". A concept referred to as fuzzy searching can also be employed to account for spelling variances. There are a wide variety of searching techniques available. It is recommended to engage with the technology provider and/or vendor for the systems in use as they will typically be able to provide expertise in this area.

Where a dispute regarding keywords arises, it can be helpful to outline the approach taken in developing the keywords in any correspondence and/or affidavit evidence submitted to the Court.

One key rule when searching documents is that it is generally not productive to search one's own data for one's own name, business name, or email address. Such keywords will almost certainly result in all data being responsive. It is also usually not productive to search for the word "privileged" which tends to appear in footers to emails sent by organisations.

It can be very useful to sample the outcome of keyword searching to identify any false positives before commencing the review, as this can often dramatically reduce the volume of wholly irrelevant documents requiring review and may reduce costs.

9.4.3 Technology Assisted Review/ TAR

When all of the data and its detailed statistics have been collated, the effectiveness of any proposed filters will be understood. It is very difficult to accurately decide on the use of TAR before this information is available.

While TAR is deployed during the Review phase (Chapter 12 below), it is at the Processing phase that the decision is made as to which particular data sub-sets will be subjected to TAR. Appendix G contains a detailed overview of TAR, including CAL workflows.

9.5 Bring forward for review

Data that is responsive to the filtering criteria, together with those sources in respect of which a determination has been made should be brought forward for review. Their families should also be brought forward for review (i.e. where an attachment is responsive, then its parent email should also be included). Where duplicates of responsive data exist within another unique family of data, this other unique family of data should also be brought forward for review (e.g. where the same attachment is attached to two different emails, both emails and two copies of the attachment should be included). This approach allows decisions to be made about how duplicates and families of data should be managed throughout the review phase.

In the event that TAR is to be used at the initial review phase, then all unique families of data (after email thread deduplication) should be brought forward for review. For instant/short message data, a decision should be made as to whether the review unit should be a message, a conversation slice, or a channel/thread segment, and steps should be taken to ensure necessary context (e.g. prior messages, reactions, edits) is included. For example, WhatsApp/SMS/direct messages will often be sliced into separate 24 hour periods (i.e. 12 hours either side of the relevant message(s)) in order to allow sufficient context to show the start and end of a relevant exchanges of messages. Where emails reference cloud-hosted documents via links, practitioners should ensure that linked items identified for review are collected and associated back to the email in which they were referenced. Where decisions have been made to exclude for review certain documents (e.g. in efforts to cull false positives and other wholly irrelevant docs from the processed set), then these should be retained and audit evidence retained which explains how they were excluded and any sampling conducted around that.

Chapter 10 Discovery request

10.1 Introduction

In Commercial Court litigation, the general practice is for parties to exchange requests for discovery following the close of pleadings and the Court normally makes directions on that basis.[13]

Order 31 rule 12(6)(a)) RSC requires that a party seeking discovery must specify precise documents or categories of documents of which discovery is sought and set out the particular reasons why discovery of each document is required. (A template/sample of a request for discovery is at Appendix J.)

In providing the reasons for seeking discovery, the discovery request should explain how each category of document is relevant to a material issue in the case and why discovery is necessary. In cases where it is known from an early stage that expert evidence will be required, it may be helpful for the expert witnesses who have been engaged to advise on what categories of documents it will be necessary for them to review in order to prepare their opinions for the Court.

10.2 Focus of request

In preparing the request for discovery, a balance needs to be drawn between seeking every possible type of document that might potentially be relevant and limiting the scope of the categories so narrowly as to overlook essential documents. Parties should avoid formulating discovery requests which contain an excessive number of categories.

Discovery requests should generally: (a) seek only those documents which are relevant to the issues in dispute as appears from the pleadings; (b) articulate clearly for each category why the documents are relevant and necessary; and (c) not present unreasonable difficulty or disproportionate expense for the recipient to comply with.

Broadly phrased terminology such as "all documents relevant to" or "all documents on which the plaintiff intends to rely" should generally be avoided as it can lead to excessively broad categories of discovery which in turn increase the time and expense of making discovery. Parties may come to an agreement whereby additional documents which have not been produced, but which a party intends to rely, will be provided as soon as is practical in advance of trial.

[13] *It is permissible in exceptional circumstances for a party to seek discovery in advance of delivering a pleading but this occurs only very rarely. See Law Society of Ireland v Rawlinson [1997] 3 IR 592; Craddock v Raidió Teilifís Éireann [2014] IESC 32.*

There are particular cases in which the broader framing language of "all documents relevant to" will be appropriate. Generally, these cases involve allegations of behaviour which has been concealed or hidden from the plaintiff, (e.g. conspiracy) with the result that the plaintiff is only broadly aware of the details of the conduct and necessarily requires the categories to be phrased widely so as to capture all documents which relate to that conduct.

Well drafted discovery requests generally have the characteristics listed below.

10.2.1 Clearly defined categories tailored to the pleadings

All categories of documents sought must be relevant to material issues in the case as they are defined in the pleadings. Poorly drafted pleadings tend to give rise to poorly drafted discovery requests which are unfocused, imprecise, open-ended, generalised, difficult and expensive to comply with.

Practitioners should consider whether the issues in the case can be refined through the use of targeted and precise notices for particulars or interrogatories. This may be a far more effective and less costly exercise than seeking overly broad discovery in response to vague or imprecise pleadings. In addition, practitioners should avoid using generic phrases such as "all documents relevant to..." and seek instead to define and be as prescriptive as possible as to the nature of the data that is being requested, e.g. "invoices" or "correspondence between A and B".

The request for discovery should explain why: (a) each document category which is sought is relevant to the matters at issue in the case; and (b) discovery of the document category is necessary for the fair disposal of the proceedings or to save costs.

The relevance of each category must be separately explained by reference to specific paragraphs of the pleadings and the particulars. Parties should not seek discovery of documents where there is merely a possibility that they will be relevant to the issues in the case. In order to be deemed relevant it must be reasonable to suppose that the category of documents, when discovered, will either directly or indirectly enable the party seeking discovery to advance its own case or to damage the case of the opposing party.
In demonstrating the necessity for discovery of each category, the request must explain how discovery of each category is necessary for the fair disposal of the case or to save costs. This involves more than merely stating this in the request, and a statement that discovery is necessary, without any explanation as to why this is so, should be avoided. For example, it may be necessary to obtain

discovery of a particular category of documents because the information, which it is anticipated will be revealed by those documents, is not available from any other source.

Where a category is likely to contain documents which are confidential, the discovery request should explain why it is necessary that the documents should be discovered notwithstanding their apparent confidentiality. Subject to arrangement between the parties, it may be permissible for the producing party to redact certain confidential or commercially sensitive data – but typically only where this data is not relevant. This point is further explained at para 12.3.

The necessity for discovery will be considered having regard to all the relevant circumstances including the burden, scale and cost of the discovery sought. Categories of discovery sought should be confined to what is genuinely necessary for the fairness of the litigation. Practitioners should avoid merely asserting necessity without supporting facts.

While there is no explicit reference in Order 31 Rule 12 RSC to the concept of proportionality, it is closely aligned to an assessment as to the requirement of "necessity on the facts of a particular case". Practitioners should ensure that there is proportionality between the extent or volume of the documents to be discovered and the degree to which the documents are likely to advance the case of the applicant or damage the case of his opponent, in addition to ensuring that no party is taken by surprise by the production of documents at trial.

10.2.2 Limit timescale involved

Practitioners should ensure that, as far as possible, the applicable timescale for every category of discovery is clearly defined so that, for example, only documents generated or coming into the party's possession over a particular period of time are captured by the request. The timescale should be limited so as only to encapsulate those documents which are both relevant and necessary. Timescales should be referenced specifically to the pleadings and fully explained. The exercise of limiting timescales may significantly reduce the amount of discovery captured by a particular category and thus the expense and burdensome nature of it, therefore avoiding an argument that it may be disproportionate or unduly oppressive. Different timescale limitations may be sought in respect of different categories of discovery, depending on the matters at issue.

10.2.3 Avoid duplication between categories

A well drafted discovery request should contain little or no duplication between categories of discovery requested. The duplication of categories of documents not only extends the scope of discovery to be made but it is also problematic insofar as the categorisation of discovered documentation is concerned.

10.2.4 Seek documentation in searchable format

Order 31, Rule 12(1)(c) RSC provides that where the discovery sought includes electronically stored information then the party seeking discovery should specify in their discovery request whether they seek the production of any documents in searchable form (or a format which allows the receiving party the same ability to access, search, and review the documents as the producing party) and, if so, whether that party seeks the provision of inspection and searching facilities using any IT system owned or operated by the party to whom the request is directed.

In practice, given that almost all categories of documents mentioned in a discovery request will capture electronically stored information, parties proceed on the assumption that the discovered documents will be produced in searchable form. If however, the party seeking discovery requires that the documents be produced "in the searchable form in which they are held by the party ordered to make discovery" (Order 31, Rule 12 (2)(c)(i) RSC), or that "the party ordered to make discovery make available inspection and searching facilities using its own information and communications technology system" (Order 31, Rule 12(2)(c)(ii) RSC), then specific reasons for such requests must be given.

The only express restriction in the Rules on the entitlement to obtain the documents "in the searchable form in which they are held by the party ordered to make discovery" is where this can be done without significant cost to that party. Where the Court is satisfied that the party seeking discovery would not be able to search the discovered documents electronically "without incurring unreasonable expense", the Court can order that the party providing the discovery should make available inspection and searching facilities using its own IT system, so as to allow a party seeking discovery to avail of any search functionality available to the party ordered to make discovery. In reality however it is generally more expensive (for all parties) to review discovery documents using the other party's IT system. In practice, an actual request for interrogation of the other party's IT system is likely to arise only in very particular circumstances, such as where the recipient cannot review and understand the data (where, for example, it exists on a bespoke software platform); where there is a concern about the integrity or completeness of the discovery; or where access to particular categories of metadata are required for a specific reason.

10.3 Agreement to make voluntary discovery

An agreement by a party to make voluntary discovery has the same effect as if a court order in those terms had been made (Order 31 Rule 12(7) RSC), provided that the party requested to make voluntary discovery was informed at the time of the request that:

a) Voluntary Discovery was being sought pursuant to Order 31 Rule 12.

b) An Agreement to make discovery would require discovery to be made in like manner and form and would have such effect as if directed by court order.

c) Failure to make discovery might result in an application to court to penalise the default.

Therefore, where an agreement to make discovery is reached, the party who has agreed to make discovery is obliged to produce an Affidavit of Discovery[14] in the proper form and is liable to the same remedies for default in making discovery as apply for breach of a court order for discovery.

Every request for voluntary discovery should address points (a) to (c) above. The request should also confirm the time limit for response and for making discovery.

10.4 Considering how to respond to a request for discovery

Practitioners receiving discovery requests should be thorough in examining the full extent of the discovery sought against their clients and be particularly live to the concepts of relevance, necessity and proportionality. Specific instructions should be sought from clients in relation to agreements to make discovery with the practitioner having clear instructions and an understanding of the practical issues which compliance with each category would raise.

Where the party from whom discovery has been requested considers that it would be unduly burdensome to make discovery in the form sought, a detailed description of the nature of the burden which will arise should be given. In most cases, it is also preferrable for the party to propose a more limited category of documents to be discovered and to provide an explanation as to why it would be less burdensome for this category of documents to be discovered.

10.5 Discovery requests against non-parties

[14] *Where a party is ordered to make discovery, Order 31, Rule 13 RSC requires that this is done on affidavit made out in the format required by Form 10, Appendix C RSC. The proper title of this Affidavit, as set out in Form 10, Appendix C is "Affidavit as to Documents", although in practice is almost universally referred to as an "Affidavit of Discovery".*

Where it appears to the Court that any person or entity who is not a party to an action is likely to have or have had in their possession, custody or power any documents which are relevant to an issue arising or likely to arise in the action, the Court may order discovery or inspection of such documents, or may give leave to deliver interrogatories. The Rules regarding *inter partes* discovery apply equally to non-party discovery.

A party who seeks non-party discovery must request specific categories of documents and give reasons why each category is relevant and necessary. A request for discovery from a non-party must be proportionate and if it is oppressive may be resisted on that basis. The Court must be satisfied that the documents are not available to the applicant from another source.

The party seeking discovery from the non-party must indemnify the non-party in respect of all costs reasonably incurred in making discovery. Discovery is made on oath in the usual way and the applicant is entitled to seek inspection of the discovered documents under Order 31 Rule 29. The obligation to provide the documents arises where the applicant undertakes to indemnify the non-party in respect of the costs of making discovery, although in practice a non-party may be reluctant to produce the documents before its costs are discharged and the party seeking discovery may be content to discharge the costs prior to receipt of the discovery. This is, however, a matter for agreement between the non-party and the applicant, as there is no entitlement under the Rules to resist producing the discovered documents pending payment.

Chapter 11 Agreeing terms of discovery and motions for discovery

11.1 Agreeing the terms of discovery

11.1.1 Meetings between representatives to negotiate terms of discovery

In the Commercial List, the Court typically makes directions for the parties to exchange requests for discovery and then to exchange responses to those requests. Correspondence regarding the categories of documents to be discovered can quickly become voluminous and time consuming to deal with and it is generally more efficient, once initial responses to discovery requests have been exchanged, for the parties' representatives to meet to seek to agree the scope of the categories to be discovered by each side.

Indeed, the current practice in the Commercial List is that the Court will generally decline to assign a hearing date to motions for discovery until it is satisfied that the parties have made adequate efforts to meet to attempt to negotiate the terms of the discovery to be made. The basis of this practice is that as the legal principles regarding the test for discoverability are so well established, practitioners are best placed to agree on the terms of discovery. Court time should only be taken up with resolving discovery motions as a matter of last resort. The practice also serves the interests of the parties as the preparation, running and determination of discovery motions can add considerably to the cost and duration of litigation. While it is frequently the case that it is necessary for a party to issue a motion for discovery in order to seek to persuade the opposing party to alter a position it has adopted, every effort should be made to agree the parameters of the discovery to be made before bringing a discovery motion before the Court.

Meetings between parties' representatives on the parameters of discovery should be held on a "without prejudice" basis so that the discussion can be as candid as possible. It is generally helpful for tables to be prepared which record the original wording of each category sought and the iterations of those categories as they evolve over the course of the discussions.

It is also recommended that in the course of these discussions, parties also attempt to agree the methodology of the discovery process, including what date ranges will apply to searches to be carried out in respect of particular categories;

the circumstances in which portions of relevant documents may be redacted; and a cut-off date where all documents generated after that date are deemed to be privileged without the necessity for them to be individually specified.

In some circumstances, it may be possible go further and to agree the identity of the custodians within and outside of a party's organisation whose documents will be discovered or even the data sources to be searched. However, parties may feel that they do not have sufficient information about their opponent's organization to enable them to agree that a particular custodian or data source may be omitted until the discovery exercise is further advanced.

11.1.2 Documenting an agreement on terms of discovery

Where the parties reach an agreement on the categories of documents which they are each to discover, and on any other aspect of the discovery to be made, the agreement should be documented. The form of the agreement may be in a letter, a set of agreed terms or in a consent order.

Irrespective of the form, in substance, the agreement should include at a minimum, terms which address the following:

- The wording of the categories to be discovered. The agreed wording should be as specific as possible and ensure that the party making discovery will be able to quickly identify whether any data falls within the wording of that category and ought therefore to be discovered. Parties should avoid using generic phrases when describing the categories.
- The deponent who will swear the affidavit of discovery.
- The period of time from the conclusion of the agreement to when the discovery will be delivered.
- Details of any further aspects of discovery on which agreement has been reached (e.g. date ranges, approach to metadata custodians or data sources to be searched) should also be included.

While Form 10 of Appendix C RSC requires that the party making discovery lists the documents in a manner corresponding with the categories in the agreement, documents may correspond to more than one category. It is best practice to list each document only once under the category to which it is corresponds most closely and this approach should be reflected in the averments made in the relevant affidavit of discovery, which lists these documents.

11.2 Motions for discovery

In the event that the parties do not reach agreement on the terms of discovery, it will be necessary for the party seeking discovery to issue an application to the Court seeking an order requiring the other party to make discovery in the terms sought.

As with all interlocutory applications to a court, a party bringing a motion for discovery should issue a notice of motion and ground the application on affidavit.

The notice of motion will set out each separate form of relief which the applicant seeks from a court. Primarily, in the High Court this will usually involve seeking an order for discovery pursuant to Order 31, Rule 12 RSC and the categories of documents for which the order is sought should be listed in a Schedule to the Notice of Motion.

The affidavit grounding the discovery motion should be sworn by the solicitor who acts for the party seeking discovery. It is generally preferable for affidavits used in interlocutory applications to be sworn by the parties themselves and not by the solicitors acting on their behalf. However in the case of applications for discovery, it is more usual for the solicitor to be the deponent as he or she will have more detailed knowledge of: (i) the material issues in the case; (ii) the reasons for which the categories of documents sought are relevant to those issues; and (iii) the necessity of those documents for the fair disposal of the case and/or the saving of costs.

The grounding affidavit sworn on behalf of the party seeking discovery should include averments dealing with the following:

- A brief background to the history of the dispute and a summary of the material issues in the case. In summarising the positions adopted by the parties on each material issue, the deponent should identify the relevant paragraphs of the pleadings in which the parties engage on the issue in question and exhibit any further particulars of pleading which have been delivered which illustrate the position adopted by either party on that issue.
- Where interrogatories have been delivered, it may also help to refer to any parts of the sworn replies which supplement the position taken by either party on the material issues.
- The categories of documents in dispute should be listed and the request for discovery by which they were originally sought should be exhibited. Order 31, Rule 12(1)(b) RSC requires that the affidavit grounding the motion furnish the reasons for which discovery of each category is sought. While it

is common for parties to simply exhibit the request for discovery and allow the letter to speak for itself, it is better practice to recite the reasons for which each category is sought in the body of the affidavit itself and to supplement those reasons with any relevant evidence.

- The positions respectively adopted by each party on the disputed categories and/or approach should be summarised with reference to the correspondence exchanged between them. All inter partes correspondence dealing with discovery should be exhibited to the affidavit.
- Additionally, the deponent should exhibit any other documentary material which illustrates the reasons for which discovery is sought. Alternatively, any pre-action correspondence which suggests that the respondent possesses, or has the right to possess, documents of the sort of which discovery is sought should be exhibited.
- Where several categories are in dispute, it may be clearer to separately address each category and to summarise the reasons given in correspondence by the respondent for its refusal to make discovery in the terms sought in respect of each category and to set out the applicant's response. Any revised wording to the categories which has been proposed by the respondent should also be set out, together with the applicant's response to such proposals.
- In some cases, the grounding affidavit may be accompanied by additional affidavits sworn to illustrate why discovery of some or all of the disputed categories is sought. For example, where an expert witness considers that it is necessary to review the documents specified in a particular category in order to proffer an opinion in the case, an affidavit from that expert should be delivered explaining why the document is required. The expert should not however offer an opinion on the question of whether a particular category of documents is "relevant to" a particular issue raised on the pleadings as this is a legal question.

A respondent to a discovery motion may deliver a replying affidavit in order to set out the reasons for which he or she objects to making discovery in the terms sought by the applicant and/or why a more limited form of discovery is appropriate. A party usually declines to make discovery of a particular category because it believes that: (i) the category sought is not relevant to a material issue in the case; or (ii) discovery is not necessary for the fair disposal of the case or to save costs. The replying affidavit sworn on behalf of the party resisting discovery should include averments dealing with the following:

- Where the respondent believes that averments made in the applicant's grounding affidavit have incorrectly described the material issues in the case or have misstated the position adopted by the respondent either on the pleadings or in correspondence, the position should be corrected by the deponent.
- Where the respondent claims that a particular category or approach is not relevant to the material issues in the case, the deponent should identify any particular part of the pleadings or concessions made in replies to interrogatories or correspondence which demonstrate either that the alleged issue is not actually a matter in dispute or if it is, that the applicant has not correctly described the parameters of that dispute.
- Where a respondent claims that discovery of a particular category of documents is not necessary for the fair disposal of the case or for the saving of costs, the Court will generally balance the litigious advantage which the applicant claims discovery of the documents will confer against the prejudice that the respondent claims will be caused if an order for discovery of that category is made.
- Objections as to necessity frequently involve contentions that: (i) the quantity of documents likely to fall within the category are so voluminous that an oppressive burden would be imposed on the respondent were an order for discovery to be made; (ii) the information which the applicant seeks to obtain from the category in dispute could be adequately obtained through a more narrowly framed category, which would reduce the burden on the respondent, or from a wholly separate source, e.g. by expert evidence; or (iii) the documents likely to fall within the category to be discovered will contain irrelevant confidential information the disclosure of which will harm either the respondent or some other connected party.
- Where the respondent asserts that an order for discovery of the particular category would be oppressive and would require the review and collation of voluminous quantities of documents, details of the material to be reviewed should be given. This should include information as to the range of likely custodians, their geographical locations, and the range of document sources in which the documents are likely to be stored.
- An objection on the grounds of oppression must be supported by evidence and it will therefore also be necessary to put before the Court an affidavit from an eDiscovery/IT/document retrieval expert averring as to the number of documents that are likely to be stored within the document universe, the detail of the automated retrieval and searching techniques that will be required to collate the documents and the likely man hours and costs

involved in such a process. This affidavit should also identify what reduction will be achieved in the number of documents to be reviewed and in the man hours and costs if discovery of a more limited category is ordered.

- Where the respondent asserts that an order for discovery of the particular category would be likely to lead to the disclosure of confidential information, the deponent should set out the basis on which a duty of confidentiality is said to arise, the identities of the persons to whom the duty is owed and why disclosure of the category sought is likely to impinge on that duty. If it is alleged that the category sought will encompass personal data the disclosure of which will likely infringe data subjects' rights under the GDPR, details of this should be set out in the affidavit.

Once the affidavits have closed and in advance of the hearing of the discovery motion, it is generally helpful for the parties and the Court for tables to be prepared which record the original wording of each category sought and any proposals and counter-proposals to modify those categories since.

On the hearing of the discovery motion, the Court will review the parties' affidavits and hear oral submissions on their behalf. Where an order for discovery is made, it is usual for the applicant to take a note of the wording of the order and then to email a draft of the order to the relevant court registrar so that the order may be perfected.

The court will also deal with the question of costs. Generally, where an applicant succeeds in obtaining an order for discovery for all or most of the categories in dispute, he or she will be entitled to an order for costs. Conversely, where the Court has declined to make an order for discovery or has made only a limited order requiring a small portion of the categories in dispute to be discovered, costs will be awarded to the respondent. If it is not possible to determine which of the parties has been successful having regard to the scope of the discovery which has been ordered, the costs will generally be made costs in the cause.

A factor which will affect the exercise of the Court's discretion is whether both parties were prepared to meet to discuss the parameters of discovery prior to issuing the discovery motion. Where one party was prepared to meet but the other was not, and the Court is satisfied that had such a meeting been arranged it is likely that the discovery motion would not have been brought or if a motion had been brought, there would have been fewer issues in dispute, the Court may order that the party who declined to meet should bear the costs of the motion. Courts typically look unfavourably on parties who refuse or fail to engage constructively in the process and particularly those parties who refuse to meet and confer on discovery issues.

11.4 After discovery is ordered.

In the Commercial List, when making an Order for discovery, the Court will also set a date on which the proceedings will next be listed for directions in the Commercial List after the parties have made discovery and completed the production of the documents in question.

It is commonly the case that on receipt of discovery documentation, the party to whom discovery has been made will have queries or concerns regarding the adequacy of the discovery which has been made which may include that the party making discovery has: (a.) omitted documents which other evidence suggests must exist; (b.) failed to search the documents held by particular custodians or stored in particular data sources; (c.) incorrectly redacted parts of the text of relevant documents; or (d.) wrongfully asserted claims of privilege.

The Court will usually set directions to allow time for such queries/ concerns to be addressed. Where the parties cannot resolve these issues, the party who has raised the concerns may then apply to the Court to resolve the issues in dispute. However, the Court will generally encourage the parties to resolve these issues themselves, in order to avoid Court time being taken up. Parties should also be aware that a party seeking further and better discovery in respect of the discovery process undertaken should be able to show how the approach undertaken has resulted in relevant documents being omitted from the discovery. This is best demonstrated by showing actual gaps in the discovery provided in terms of timeframes and/or the existence of documents (or portions of email conversations) which the responding party ought to have discovered but has not.

Parties should also be aware of the possibility under Order 31, Rule 12(11) RSC of applying for a variation of an order or agreement for discovery, if the discovery originally ordered or agreed proves to be unreasonable having regard to the costs or other burden of providing discovery – or conversely where further discovery is necessary. In considering an application of this type, the Court will usually require to be satisfied that the reasons for seeking the variation could not have been known at the time that the original application for discovery was brought. It is advisable to make such an application as the earliest opportunity after a party determines that an application seeking to vary will be necessary.

Chapter 12 Review

12.1 Introduction

This phase in the discovery process involves a manual review of the potentially relevant data to determine relevance by reference to the categories of discovery agreed and/or ordered. This may be an entirely manual review. Alternatively, in suitable cases all unique families of data may be brought forward (with or without filters such as keywords previously applied) and predictive coding may be used to filter the documents to those likely to be relevant, after which a manual review is typically performed.

Reviewers should be able to search, annotate, redact, flag, and bookmark individual documents or collections of data by reference to project specific issues (e.g. the various categories, privilege options, commercially sensitive options, hot documents, and any other relevant tags.)

Prior to the commencement of the review, the following steps should be taken:

(1.) Deduplication of documents.

(2.) Decision as to how documents will be batched for ease of review.

(3.) Guidance on the application of tags, including privilege tags.

(4.) Whether emails will be threaded from the start of the review.

(5.) Whether there is any category of documents that can be bulk tagged

The review phase is overseen by a solicitor who is familiar with the proceedings, who is in a position to answer queries raised by the document reviewers and liaise with the client to address any queries. That solicitor must be able to stand over the quality and completeness of the discovery when ultimately produced. A detailed sample review plan is attached at Appendix H. This contains a number of typical approaches to review and guidance of when each approach might be appropriate.

12.2 Categories

Each relevant document must be categorised in line with the categories of discovery as agreed and/or ordered. Depending on the volume of potentially discoverable documents, this can be a lengthy and costly process. It is common practice that practitioners agree to list the documentation by reference to the most relevant category but include an averment in the

Affidavit of Discovery that this may not be the only category to which the document is relevant.

A sample averment in this regard is:

> The [**INSERT NAME OF PARTY**] in making discovery has for ease of reference listed each document being discovered under one of the [**INSERT NUMBER**] categories of documents within the [Agreed Discovery **OR** Ordered Discovery]. It is the case, however, that there is an overlap between various categories in the [Agreed Discovery **OR** Ordered Discovery] in that many of the documents being discovered might reasonably be considered to fall within the terms of a number of categories. The [**INSERT NAME OF PARTY**] has been advised that it is not obliged, and could not reasonably be expected, to identify every category under which a particular document might be listed. Accordingly, while the documentation listed in the First Schedule under the [**INSERT NUMBER**] categories comprises the totality of documentation which the [**INSERT NAME OF PARTY**] is discovering pursuant to the [Agreed Discovery **OR** Ordered Discovery], the [**INSERT NAME OF PARTY**] does not thereby suggest nor is it the case that all the documents listed under any particular category are the only documents being discovered which are within the scope of that category.

Where parties assign multiple categories to a document, the document should be listed and provided only once, with the categories it is responsive to listed together in the schedule.

12.3 Privilege

Where a document contains information which is privileged from disclosure (e.g. because the data encompasses communications between a client and solicitor for the purposes of discussing legal advice), the party making discovery must still discover the document, in the sense that it must be listed in Schedule 1, Part 2 of the Affidavit of Discovery, but it may decline to produce the document to the other party.

Privilege may be claimed over a document and not the fact of its existence. All relevant documents over which privilege is claimed in full must be listed in Schedule 1 Part 2 of the Affidavit of Discovery. Where only part of a document is privileged, the document should be listed in the First Schedule, Second Part. Please see Appendix K for an Overview of legal privilege.

The documents should be individually listed and the type of privilege being claimed (whether in whole or part) should be specified. Using the metadata associated with privileged documents to populate the descriptor columns that typically form part of the First Schedule Second Part (as set out in the Sample Affidavit of Discovery at Appendix J) will ordinarily disclose sufficient information (through the inclusion of its name, author, date, doc type, sender/recipient details etc) to enable the receiving party to assess whether a claim of privilege is appropriate. Making use of document metadata in this way also represents a time and cost-efficient solution to the individual listing of documents.

If there is a risk that disclosing the name of the document, subject bar of an email or name of communication thread, may disclose the privileged content that is sought to be protected, it may be appropriate to provide in the alternative, a meaningful narrative of the document over which privilege is being claimed. This circumstance was addressed in Quinn v IBRC [2015] IECA 84, which required the defendants to provide "*a meaningful narrative*" with a sufficient description of the document to allow the receiving party to make a reasoned judgment as to whether the document had been validly categorised as privileged. Where the underlying metadata is not being relied upon and a meaningful narrative is being provided, this should be addressed in the body of the affidavit of discovery and made clearly identifiable in the schedule of the affidavit of discovery.

Separately, where the basis for the privilege claim it is not clear from the name of the document, subject bar of an email or name of communication thread, e.g. where it simply states "Letter" without more, or is a scan of a document with an alpha-numerical name, or is simply entitled "Privileged & Confidential" (at a point in time which would not attract litigation privilege), a further narrative description may provide context to the document and reduce the potential for challenge.

Separately, to ease the administrative burden of the review exercise, it is common for parties to agree that documents created after the commencement of proceedings, over which a claim of privilege is being asserted, do not need to be listed in the Affidavit of Discovery. The assessment of whether documents are privileged is made by a solicitor familiar with the facts of the case and the discovery process.

12.4 Redactions

There is no express entitlement in the Rules to redact information. Redaction should be used as sparingly as possible and only where it can be justified. This is particularly so as the time involved in reviewing multiple chains of email containing data to be redacted can become an inordinately expensive exercise.

Where discovered documents are redacted, the Affidavit of Discovery as to documents should contain a suitable averment setting out the basis for those redactions.

The main reasons why data may be redacted is because the data are:

- Privileged communications.
- Not within the scope of the discovery categories.
- Commercially sensitive and/or confidential information.
- Personal data relating to other parties who are not party to the proceedings

Where data is within the scope of the categories of discovery sought it must be discovered whether or not it is personal data, commercially sensitive and/or confidential, save where leave of the Court is obtained. In certain circumstances it may be permissible to redact commercially sensitive information in a discoverable document where the information concerned is not within the scope of the categories of discovery or does not advance the opposing party's case.

While it is not usual to redact non-relevant portions of an otherwise discoverable document, personal data relating to specific employees, customers or other third parties should be redacted to the extent that it is neither necessary, proportionate nor relevant to the legal proceedings.[15] The decision in Farrell v Everyday Finance DAC [2024] IECA 16 §132 recognises that redactions "to protect the data rights of all third parties" are appropriate in cases where such information is not relevant to the proceeding.

Given the proliferation of email as a form of communication and the resultant chains, where a redaction is applied, care should be taken to ensure that redactions are consistent across duplicate or near duplicate documents. The use of email threading reduces the risk of inconsistent redaction across emails chains.

When documents have been redacted, the document name and other metadata should be reviewed to ensure that they do not disclose the content that has been redacted.

The deponent of the Affidavit as to Documents may have to explain the basis for redactions at trial and it is recommended to note in the schedule the reason for redactions by way of acronym – e.g. 'RR' = redacted for relevance; 'RP' = redacted for privilege, etc.

[15] *Court approval may be required in respect of any such redactions to ensure they are necessary and proportionate.*

12.5 Dealing with problem documents

Section 9.3 outlines some approaches to dealing with problem documents at the processing phase. However, many problem documents cannot be easily identified until manual review has been completed. Where a reviewer identifies a problem with a document they should tag it as a "tech issue" and those documents can then be escalated to identify and resolve the issue.

12.6 Client review

It is recommended that throughout the discovery process, the proposed deponent of the Affidavit of Discovery is kept up to date on the progress. It is vital at the review stage that the deponent of the Affidavit of Discovery takes time to review the draft affidavit and schedule, and the documentation listed in the schedule.

Chapter 13 Production

13.1 Introduction

The output of the review phase will be a set of documents deemed discoverable, some of which will also be marked privileged or partly privileged and some may have redactions applied.

The objective of the production phase is to generate a copy of the documents that respond to categories of discovery.

Adequate time should be allowed for the production phase to be completed. Producing the documentation can be complex as comprehensive quality checks are required and the production of electronic schedules, amongst many other tasks. While it may be possible to complete a small production in a number of hours, in a large discovery most productions take a number of days to complete to an appropriate level of quality. It is therefore recommended that in large scale discovery projects parties ensure that an appropriate amount of time is incorporated into the timetable to cater for production.

13.2 Families of documents

Consideration for how families of documents (see Appendix E for a detailed description) are managed at production will have been included in the review planning and strategy. It may be helpful to include a schedule of irrelevant family members which have not been produced, and/or include a slip sheet for each document which has not been produced. This can assist in demonstrating that the document has been withheld intentionally, rather than due to a technical issue or oversight.

It should not be the case that wholly irrelevant documents are produced just because they are associated with a family where only one member is relevant. Neither should it be the case that documents should be redacted in full in these circumstances. If a document is redacted in full, it would simply be more cost effective not to produce it in the first place. While it is prudent to only produce the relevant portions of document families, care should be given to ensure documents are not produced as orphans, and are produced with their parent, e.g. cover email (as more particularly described in Appendix E).

13.3 Production format

It is common (and usually the most cost effective) for discoverable documents listed in Schedule 1, Part 1 of the Affidavit of Discovery, to be produced in their native format, along with a schedule listing their original metadata details, categories, etc. in a "load file". Documents which have been redacted are produced alongside non-redacted documents, but in a redacted (e.g. PDF/TIFF) format. Documents which have been marked as privileged are not produced (save where redacted as part privileged as referred to in Schedule 1, Part 1), and a schedule of such documents is produced.

The RSC state that, if requested, production should be in a searchable format and in the format which they are held by the party making discovery. This is often referred to as native format. In most cases, this is the most efficient way to produce electronically stored information, as it does not require the producing party to incur the cost of converting it to a different format. If a party decides not to produce documents in native format the reasons should be clearly explained and agreed before the documents are produced. Unless requested, documents should not be converted into a less accessible format (such as electronic images or to paper) for production purposes.

Where an Optical Character Recognition/ OCR process has been completed to convert non-searchable documents to a searchable format, the results of this process may also be provided to the requesting party. Given the nature of OCR technology, such text should be provided on an "as is" basis, with no assurances that the technology has rendered complete and accurate text

In the event that a party converts electronically stored information into a different format, steps should be taken to ensure that elements of the electronically stored information, such as metadata, are not unintentionally lost or obscured in the process.

Note: The actual documents produced are often renamed as their production number, with their original electronic file name being included in the schedule instead. This is helpful as often electronic file names are too long to be easily moved between disparate systems, so using the document production number as the name (typically a short alphanumeric string) avoids such compatibility issues between systems.

13.4 Schedule

A sample Affidavit of Discovery and schedule is attached at Appendix J. The sample schedule includes suggested standard fields and format. Further detail is included in Attachment Four of the Discovery Plan at Attachment G.

13.5 Inadvertent disclosure of privileged and/or other documents

Even with comprehensive quality controls in place, the complex nature of discovery projects can result in data which should not be disclosed, accidentally being disclosed. Parties should attempt to agree in advance of making discovery how to deal with situations in which data is inadvertently disclosed. This is often referred to as a "clawback agreement" and accounts for how a party might notify the other party of which data should not have been disclosed and what steps might be taken to remedy the situation. This may be in the case of privileged documents and/or documents subject to data protection restrictions. The absence of a claw back agreement does not dilute the obligation on a solicitor not knowingly to read or deploy an obviously privileged document belonging to another party and to notify the other solicitor of the receipt of the document promptly. See Appendix K for Overview of legal privilege.

13.6 Inspection

Once the Affidavit of Discovery has been served the opposing party is entitled to inspect the documentation listed in Schedule 1, Part I of the Affidavit of Discovery.

It is more efficient if the non-privileged data which is being discovered in Schedule 1, Part I of the Affidavit of Discovery is also produced at the same time as the Affidavit of Discovery is provided to the other party. Indeed, the data being discovered will usually be electronically linked to the schedule.

Production of documents does not prevent a party seeking to inspect original documentation, which can be requested if necessary. Such inspection may be required in order to verify original signatures on handwritten documents, or more frequently when the documents produced cannot be rendered into a readable or understandable format without the use of specific technology which only the producing party has access to (and would be disproportionately expensive or impossible for the receiving party to obtain access to a similar system to read the documents). Examples of this include bespoke accounting and auditing systems, whereby the information is meaningless outside the original system used to

generate and store it or complex imaging or mapping systems whereby it is not possible to view the images outside the original system.

Where additional relevant documents are identified after production and inspection has been completed, a party has an obligation to produce these with a supplemental affidavit of discovery.

Chapter 14 Presentation in court

14.1 General

The objective of the presentation phase is to prepare for, and to present, documents in Court in a manner which facilitates their efficient presentation and the running of the matter.

Often, the most efficient method to present a document at a formal hearing is electronically on-screen in its native format. This saves considerable cost in printing bundles of documents, and time in leafing through large bundles of documents to find the one under discussion at a certain point in proceedings. However, in some cases, it may actually be more efficient or cost effective to print key documents for presentation.

Adequate preparation and planning should be undertaken when deciding on the best method to present the data at a formal hearing. As the technology equipment in court rooms can vary, it is vital to liaise with the Court's Registrar in advance of trial to see what can be arranged, and to ensure that the trial judge is happy to review documents electronically rather than in hardcopy.

In cases involving large volumes of core documents, it is preferable for the parties to employ trial management technology to avoid the proliferation of hardcopy files in court. This requires early liaison between the parties, the Court and the Courts Service.

It is incumbent upon the parties to agree on the use of technology, select a single technology provider, and agree how costs will be managed (usually split between the parties based on the number of users each party has). There are a number of technology providers in this area who can work with the Irish Courts Service, should the parties and the Court think it helpful. Experience shows that significant time and cost savings can be achieved at trial stage when such systems are deployed.

The court will expect technology to be used in a manner which ensures equality of arms as between the parties as regards access to the technology and training required for its use. Parties should therefore seek to agree a suitable platform for document management during the trial as well as an appropriate timetable for testing and for training for the Court, counsel and parties to the litigation. For such technology to work effectively at trial it is important that the parties agree a common document identifier convention for documents. Counsel should be prepared to identify the relevant document identifier and page or paragraph

in respect of each document being produced at trial so that it can be produced on screen promptly. Where possible parties should provide the technology service provider with the document identifiers required for the following day, confidentially if necessary, to ensure the smooth running of the trial, although in practice this may be difficult. Most technology providers providing these services provide a person skilled in the use of the system for the duration of the trial.

14.2 Co-operation in relation to preparation of core books

In advance of the start of a trial, all parties to litigation which has been admitted to the Commercial List must agree on the content of the core books of documents which are likely to be frequently referred to by the Court, Counsel and witnesses in the course of the trial. It is also recommended best practice to try to do the same in all other divisions of the High Court.

Where electronic trial management technology is not being used, it may help if core books are produced in hardcopy format and all other information is kept in electronic format (assuming it can be easily accessed at trial).

A simple method of arranging core books is often to agree that the documentation on which each party proposes to rely is presented in strict chronological order rather than by category of discovery or theme. This way, the books can be easily supplemented at trial if parties wish to add additional documentation.

As a general rule of thumb, key documents referred to in the pleadings or replies to particulars should be included in the core books. Further, if there is a direction for the service of witness statements then any documents referred to in the witness statements should also be included and ideally list the document number or identifier ascribed to the document in the Schedule to the Affidavit as to Documents. Parties should wherever possible limit the amount of documentation contained in the core books to that documentation which is truly "core".

Appendix A Discovery project checklist

Preparing	1. Brief client. 2. Assemble discovery project team. 3. Commence audit file. 4. Draft discovery plan. 5. Draft budget.
One – Identification	1. Identify relevant time period of dispute. 2. Identify custodians. 3. Identify types and sources of data.
Two – Preservation	1. Complete legal hold process. 2. Complete technical preservation steps.
Three – Collection	1. Decide on where copying and extraction of data will take place and by whom. 2. Decide on the scope of the data collection. 3. Decide on the type of data collection. 4. Consider specific data types, including data stored on social media, AI tool inputs and outputs, and hard copy documents.
Four – Processing	1. Remove irrelevant document types. 2. Convert into searchable format and load into database. 3. Deduplicate to unique families only. 4. OCR. 5. Thread deduplication. 6. Manage problem documents. 7. Apply filters, including date based filters and keywords. 8. Decide whether to apply Technology Assisted Review in Review Phase (Phase 5). 9. Bring forward for Review (Phase 5).
Discovery requests and motions	1. Prepare request for discovery to be sent to each party. 2. Meet with other side's legal representatives, prior to issuing motion for discovery, to seek to agree the terms of the discovery to be made. 3. If agreement is not possible, a notice of motion seeking an order for discovery should be issued and an application should be made to the Court.
Five – Review	1. Establish review team 2. Develop approach to review and document review plan (two-pass, single-pass, use of Technology Assisted Review) 3. Consider use of sample discovery plan at Appendix G.

Six – Production	1. Produce documents and schedules 2. Arrange inspection, if required
Seven – Presentation	1. Agree presentation format and use of technology. 2. Agree format and content of core books. 3. In cases involving large volumes of documents, consider use of trial management technology.

Appendix B Sample legal hold communications

Sample text for legal hold emails can be seen below. There are a number of steps in the process which should be considered:

- The communication is typically sent by email
- It is distributed to all custodians, including IT custodians, and 3rd parties who may have documents relevant to the matter
- While it should be broad in nature at the outset, every effort should be made to focus the request so as to not to overburden custodians
- *Where individual custodians have been identified, responses* must be tracked, including **written confirmations from each custodian that they have received, read, and understood the notice**
- Periodic reminder notices should be sent for the duration of the legal hold
- There should be a mechanism for releasing the legal hold when it is no longer required

B.1 Template legal hold email

To: [Potential Custodians]
From: [External/General Counsel]
Subject: [Matter reference] – Legal hold – Preservation of relevant information

Dear [Custodian name],

As you may be aware, we have been notified of a potential [litigation/regulatory review/complaint] regarding [the work we completed for [insert client or project name]/the product we supplied to [insert client name]]. We intend to vigorously [pursue/defend] these [proceedings/review/complaint].

During the course of this [litigation/review/investigation] we will need to make our paper files and any electronically stored information (ESI), which could be of relevance, available to our legal team. Also, if discovery requests are made in the course of the [litigation/review/investigation] we will need to make them available to lawyers representing [**complainant/investigator**]. It is therefore essential that you take immediate and affirmative steps to preserve all paper documents and ESI which may be of relevance to this matter which are in your custody or control.

There is a duty to preserve potentially relevant documents as soon as litigation is reasonably anticipated. Immediate compliance with this notice is therefore essential.

In the context of a potential discovery exercise documents and data means anything in which information of any description is recorded by any means, including without limitation: writings, communications, pictures, drawings, programmes and data of any kind, whether recorded or maintained on paper, electronically, audio, visual or other means and also includes all types of

electronic and printed communications, including emails, memoranda, letters, reports, presentations, spreadsheets, charts, handwritten notes, diary entries in written, printed, photographic or electronic form (including electronic information stored on cloud infrastructure), SMS text messages, instant messages, (including WhatsApp messages / Slack messages), voice recordings, voice notes and video recordings. In addition, all drafts and versions of documents must be preserved, If multiple copies exist with annotations or differences, each version should be retained.

Please note that this **obligation includes** all potentially relevant documentation and information stored on your work laptop, mobile phone, **any personal devices** and any other portable devices, such as USB keys, etc. in addition to information stored within our [project/engagement] files and our shared servers. It also includes all forms of documentation such as correspondence, diaries (electronic or hardcopy) and instant messages**. [INSERT DETAILS HERE OF THE TYPES OF COMMUNICATION USED IN THE ORGANISATION, e.g. MS Teams Messaging, Slack etc.]**

The above list is intended to give examples of the types of information/records you should preserve, but is not exhaustive. If you have any queries as to whether you should be retaining something, please do not hesitate to raise them directly with [me] or [secondary contact].

Where you are unclear as to whether a document may be potentially relevant, you should preserve that document for a more detailed review by our legal advisors at a later stage. The only preservation steps required are to not access, alter, or delete, any potentially relevant information. If you fail to preserve these materials it could be detrimental to our position in the [litigation/review/investigation].

In summary;

Step 1: What should I preserve?

Preserve all documents and data (previously defined) that relate to the matter, regardless of format or location.

Step 2: What should I do with the documents/data?

Maintain them in their current form and location. Do not alter, delete, or move them.

Step 3: What else do I need to do?

Suspend any automatic deletion policies and notify relevant departments and third parties.

Our IT team has been notified of this legal hold. They will be working with us to help ensure that we implement the legal hold effectively. We will follow-up with more information as the [litigation/review/investigation] proceeds, including advising you as to when the legal hold is no longer required.

I would be grateful if you would please reply to this email to acknowledge receipt of it and that you understand and will comply with the preservation obligations outlined above. it

If you have any questions please contact [me] or [secondary contact] at [insert contact details].

Regards, [External/General Counsel]

B.2 Template legal hold reminder email

To: [All custodians only]
From: [External/General Counsel]
Subject: [Matter reference] – Legal hold – Reminder

[Forward original full legal hold email]

Dear All,

Please be reminded that the legal hold is still in place until further notice. We will follow-up with more information as the [litigation/review/investigation] proceeds, including advising you as to when the legal hold is no longer required. In the interim, please respond to this email and confirm:

1. That you have reviewed this reminder and the original notice below
2. That you understand the notice and agree to comply with it

If you have any questions please contact [me] or [secondary contact] at [insert contact details].

Regards, [External/General Counsel]

B.3 Template legal hold release email

To: [Each custodian who data collection has been completed fully]
From: [External/General Counsel]
Subject: [Matter reference] – Legal hold – Release

[Forward original full legal hold email]

Dear [Custodian name],

The legal hold referred to in the email below no longer applies to you. If you believe that you still have potentially relevant information which has not been collected to date, then please contact me immediately.

Many thanks for your assistance in this matter.

Regards, [External/General Counsel]

Appendix C Document identification questionnaires

There are two sections to the document identification questionnaire. The first is the custodian questionnaire which can be used to gather information as to the documentation of relevance to the matter which the custodian has knowledge of. The second is the IT questionnaire, which can be used to gain a deeper understanding of the IT systems in place which may contain Electronically Stored Information ("ESI") of relevance to the matter. This additional technical information will unlikely be known by the custodians themselves and typically require IT management's knowledge of the systems in place. A sample cover letter is included in this Appendix. It is followed by the two questionnaires. Before using them please carefully review them and ensure that they are amended in order to be appropriate to the particular circumstances of the case or the client. Also please ensure that they are addressed to the correct people within the organization.

C.1 Sample cover letter

Dear Sirs,

We refer to our recent instructions in respect of the dispute between [X] and [Y]. As you are aware through earlier advice and discussions, [if proceedings are issued and] as these proceedings progress through the Courts you will be required to provide discovery of documents relating to the issues in dispute according to particular categories that will be agreed or ordered by the Court in due course.

To prepare for this, it is our practice to request all clients to complete the attached identification questionnaires to help us ensure that all documentation relevant to the proceedings is identified and preserved and to assist us in assessing the volume of documents which [client] holds regarding the issues in dispute.

We will first need to identify all potential custodians of documents and speak with them to identify all potential sources of documents which they hold in relation to the [(contemplated) litigation]. We will also need to confirm the date range of relevance to the matter. These consultations should take place as soon as possible and all relevant document sources should be identified so that the retrieval of relevant documents can be commenced. You should contact each of the individuals likely to hold documents, to clarify their individual document retention practices. We have enclosed a custodian document identification questionnaire to assist in this regard. Once the discovery process for these individuals has been completed, we can assess together whether the process should be repeated for any other employees.

It may be necessary to obtain information regarding your IT systems. The IT document identification questionnaire enclosed is designed to stimulate discussion between you, your IT department, and us. It does not present an exhaustive list of document sources that you must consider, nor do we imply that all its terms and/or sections apply to you. The purpose of the questionnaire is to identify the entire universe of documents which may have to be considered and to gain an understanding of the amount of documentation that might potentially have to be discovered as matters progress. Your IT department

may be able to assist in clarifying the rules for storing information on computers at [Client] so as to identify the universe of documents that are potentially relevant to the issues in dispute.

When reviewing and completing the attached identification questionnaires we recommend that you consider who has access to documentation stored not only in your business premises but also those employees and service providers who use their personal devices for business purposes. It is extremely important that you fully understand your obligations to retain all relevant/potentially relevant documentation, and we will discuss this issue with you to ensure you fully understand your obligations.

We recommend that once you have reviewed both questionnaires and considered them with your IT department, we arrange to meet to discuss your findings. In the meantime, if you have any queries or comments regarding the above or the attached questionnaire, please do not hesitate to contact me.

Yours faithfully, [Counsel]

C.2 Sample custodian document identification questionnaire

This questionnaire may be used to assist in the determination of the hardcopy document and ESI sources which a custodian has access to. The custodian may then be asked to provide information as to which of the sources may contain documents relevant to the matter, given the background to the matter and the relevant time periods, etc. Further, for all sections below, it may be necessary to determine what was in place during the time under review, and what has become of those systems and ESI if they are no longer in use.

It will also be helpful to provide a short general summary of the particular dispute.

A) General

1. How long have you been employed with [client], and what roles have you held and for which periods? What physical locations have you been located at, and what address are you located at now?
2. Considering the issues in the matter, covering the time between [X] and [Y] approximately, what documents, both hardcopy and electronic, might you have or had, which are relevant to the matter?
3. Where do you keep hardcopy documents which may be of relevance to this matter?
4. Are there any relevant hardcopy documents or ESI that once existed but are no longer held by you? If so, please provide full details of the documents, together with what became of them.
5. Do you have a policy of archiving your hard or soft copy documents? If so, please give full details of this policy.

B) Custodian-based document sources

1. What desktop and laptop computers do you use? Please detail each of them with reference to asset numbers if they have such numbers allocated to them within the company.
2. What mobile devices do you use? e.g., Android, iPhone, iPad, Tablet PC, and mobile phones. For each and any devices please confirm if they are personal or work devices.
3. Have you, since [x] had any other devices which are no longer in use (e.g. a mobile phone which has since been replaced. If so, please confirm where these devices are located and whether or not you have access to these devices.
4. What portable storage devices do you use? e.g. USB keys, floppy disks, CD/DVD's, ZIP disks. If so please confirm where these are located.
5. What email accounts do you use? Do you have more than one account? Please list all relevant accounts. Do you use any personal email accounts that you may have used purposively or inadvertently in relation to the matters in dispute.

6. Do you have a private folder which only you have access to on a network server? If so, what is its name, and what drive letter do you use to access it?
7. Do you use instant messaging? e.g., SMS text, WhatsApp, Teams chat, Zoom chat, Slack, Skype, voice notes, Telegram, etc. Internal and external messages/voice notes, group chats, individual messages, archived messages and deleted messages are all in scope so if you have such messages, please provide details of same and whether you use such messaging to communicate in relation to this dispute.
8. Do you ever use AI tools for work related purposes (e.g. Copilot, ChatGPT, Gemini or other AI platforms? If so, please provide details of the tools you use, how you use these tools and whether you use such AI tools to generate content in relation to the dispute. Please outline details of the content generated by the AI tools. Please note this includes both inputs/prompts into an AI tool and any output. It can also include for example, Q&A sessions, video or audio recordings, voice recordings, transcription of online meetings, creation of draft memos, emails, documents etc.
9. How do you remotely access your corporate IT systems? Do you use any personal computers/devices at home for your work?
10. Do use any externally hosted networking websites? e.g., LinkedIn, Facebook, Twitter. If so, are there any potentially relevant communications in respect of the dispute on these websites?
11. Do you have colleagues or an assistant who would have access to your documents?
12. Are there any other locations where ESI may be stored that you are aware of? e.g. voicemails, video conferencing systems. Please note this includes any transcription or recordings of Teams, Zoom, Skype calls
13. Do you use any form of encryption and/or password protection on the devices you use and/or on individual documents?
14. Do you have ESI which would be considered personal data under the Data Protection Acts

C) Non-custodian document sources

1. Do you use shared folders located on a server computer to which others have access ? If so, what is its name, and what drive letter do you use to access it?
2. What transactional systems do you have access to? e.g. accounting, Payroll, HR, manufacturing, funds transfer, etc.
3. Do you have access to any internally hosted websites and/or collaboration sites? e.g., Internal file sharing websites, SharePoint, eRoom, etc.
4. Do you have ESI which is hosted on the Internet? e.g., externally hosted websites, file-sharing websites, Google Docs, etc.
5. Are there any other systems which you use to access and/or store ESI? e.g., fax, scanning,

etc.

6. Do you have a policy of archiving your soft copy documents? If so please give full details of the policy.

7. Are there any other systems which you use to access and/or store electronic data related to your work or the business of

8. Are there any potentially relevant softcopy documents that once existed but are no longer held by you? This may include, for example, chats on WhatsApp where the disappearing messages setting is on? / data which been transferred to a new device and therefore lost. If so, please provide full details of the documents, together with what became of them.

C.3 Sample IT document identification questionnaire

In parallel or following the receipt of the responses from the individual custodian document questionnaires, this IT questionnaire may be used to further explore the document sources and gain more detailed information as to the underlying IT systems in place. It may be necessary to determine what was in place during the time under review, and what has become of those systems and ESI if they are no longer in use.

A) General

1. Who manages the IT infrastructure in the organisation? Please provide contact details.
2. Who supports the IT infrastructure in the organisation? Please provide contact details.
3. Are there any 3rd parties that either manage or support the IT environment? Please provide contact details.
4. Are there any 3rd parties who process or host ESI on behalf of the organisation? Please provide contact details.
5. How many IT users are there in the organisation?
6. What are the primary technologies in use? i.e., Windows, Linux, desktops, laptops, etc.
7. What standard applications are in use? e.g., Word processing, spreadsheets, Internet access, etc.
8. Do you have access to any Discovery tools, for example Microsoft Purview. If so, what Licence do you hold?
9. What encryption technologies are deployed?
10. Are there any in-house or industry-specific software programs deployed?
11. Is there a legal/regulatory requirement that requires the organisation to retain ESI?
12. At what physical locations/addresses does [client] have employees and IT systems located?
13. Is the network connected to any form of off-site storage? This could be an alternative business site, a specialist data facility, or an online provider of storage space. Please give details:

B) Policy

1. Do you have a document management policy that applies to the creation, modification, retention and destruction of hardcopy documents? If so, please give details:
2. Are there, to the best of your knowledge, any policies that apply to the use of re-usable media (floppy discs, CDs, DVDs etc.). If so, please give details.
3. Is there any policy relating to the procurement, replacement, or disposal of hardware, in particular relating to hard disk drives? If so, please give details.

4. Is there a leaver's policy? What becomes of laptops or desktop PCs used by someone who has left the organisation and who might have documentation stored on their computer? What becomes of their e-mail accounts and any of their documents on shared drives etc.?
5. Is there a leaver's policy in relation to mobile devices? Are all mobile devices wiped? Is the wiped device reallocated to another staff member or does the leaver have the option of keeping the device?
6. Is there, to the best of your knowledge, any policy relating to the creation, modification, retention, and destruction of files produced by recording telephone calls? If so, please give details:
7. Are there compliance steps in place to audit any named document policies? If so, please give details:
8. Are there any policies in relation to the use of AI, in particular the use of GenAI e.g. Copilot, ChatGPT, Gemini etc.
9. To the best of your knowledge, have any relevant or potentially relevant documents been affected by any named document policies? If so, please give details:

C) ESI and Associated IT Systems/Protocols

1. What standard applications are available to you (Microsoft Office Suite, Word, Excel)?
2. Was a different email system used during the period in question? If so, please give details:
3. Was a different domain ever in use, if so, please give details?
4. Was a different instant messaging system used during the relevant period? If so, please give details:
5. Were any different software applications used during the relevant period? Please give details:
6. Does your organisation employ any Early Case Assessment or eDiscovery software and / or methodology to preserve, collect and cull data which may be relevant to this matter? e.g. Microsoft Purview. If so, please describe the software and / or methodology and identify the person(s) who operate or manage them.
7. Please provide us with an Information or Data Map that shows the entire network and computer architecture for your organisation. Please explain how your employees and (more importantly) custodians access the network and computer architecture of your organisation.
8. Please identify any ESI that has been destroyed since [insert earliest date relevant to the case]. With respect to such ESI, please indicate whether such ESI was within the intended scope of the litigation hold, as described above.

9. To the best of your knowledge, were there any major changes to your IT architecture during the period relevant to this matter? This is principally in relation to how potentially relevant documents were created and stored. Please give details:
10. Is the IT equipment used during the relevant period still available? Please give details:
11. During this time period, would relevant electronic documents relating to this matter have been deleted? Please give details:

D) Email and Other Electronic Communications

1. What email system do you use? Is there a central server such as Microsoft Exchange? Are there any restrictions on the size of a users' mailbox and are there any policies in place which automatically delete items from folders? Please give details**:**
2. Do you or have you ever had a document management system? If so, please specify the name and manufacturer and how documents are profiled in all such systems:
3. Please identify the individual currently employed by your organisation who is most qualified to explain to us the operation behind your email data storage, backup and retention policies during the relevant period.
4. During the relevant period, was your organisation's email software or system upgraded? If so, what migration protocol(s) was in place and implemented during the upgrade(s)?
5. List the email servers and repositories in use by your organisation and provide hardware type, operating system name and version, email (client and server-side) application name and version, number of users per server and physical locations.
6. Do you use instant messaging and, if so, which software is used?
7. Do you use any specialist software specific to the industry which may contain relevant documents / data? If so, please give details:
8. Is your telephone system IT based? Are voicemails and telephone calls recorded? Is there any sort of retention policy? This includes the transcription or recording of Teams, Zoom, Skype calls. Please give details:

E) Archiving and Back Ups

1. Is there any kind of electronic archiving system? If so, please explain how the process works, how it is triggered and if duplicates are stored:
2. To the best of your knowledge, have potentially relevant documents (paper or electronic) been archived? Please give details.
3. If potentially relevant documents have been archived, in your opinion what is the best method of retrieval? Please give details:

4. During the relevant period, was any sort of backup system in place? If so, please give full details; what media was used, how often was a backup carried out, how large was an average backup and are they currently available? Was the archived data backed up a to a cloud server. Were backup tapes ever used to archive data, if so, are these fully catalogued and can they be restored?

Consideration should be given to having the both the custodian and the person completing the IT questionnaire acknowledge the completeness of the information provided by way of a signature.

Appendix D Managing audio and video data

D.1 Background

The ease of generating and sending audio and video content on smart phones and other devices means that audio and video files now often feature in discovery exercises. Care should be taken to check for potentially relevant audio and video files, which will require specific treatment throughout the discovery process to ensure that they are properly identified, captured, processed, reviewed, discovered (if relevant) and produced in an accessible format.

In a media dispute, care should be taken to ensure that all potentially relevant rushes[16] are captured for searching and review.

Audio data refers to the recording of sound only. Video data refers to the recording of both sound and picture, which may require different treatment, depending on the nature and volume of data.

D.2 Identification

Potentially relevant audio and/or video data should be identified as part of the standard identification phase. For example, practitioners should be careful to ascertain whether the party making discovery or its custodians are in the habit of leaving or receiving voice notes, taking recordings of meetings or calls, sending gifs or creating video content. Such data may be held on work issued or personal devices. If the data are stored amongst personal image files, it may be necessary to make special arrangements for identifying and capturing potentially relevant audio and video files.

Audio and video data generally adds considerably to both the volume of data being sent for review and the duration of the review, with a consequent impact on discovery costs.

It may be possible to identify and exclude obviously irrelevant audio and video files prior to commence the review phase to narrow down the material being pushed forward for review.

It will be important to confirm at this stage whether the data files are held in analogue or digital format.

It is important to do this in collaboration with an expert to ensure that steps taken do not alter the integrity of the underlying data.

Analogue files are not searchable and review of them entails listening to them or viewing them. They can be converted to digital format to make their content searchable, but this may entail substantial costs and the systems on which they are held may no longer be supported. If not converted, the review of analogue files requires full 'manual' review, either by watching or listening to the content end to end.

[16] *Draft film footage used when compiling a film or television programme.*

In a regulated entity where phone lines are recorded, each individual audio file may contain several different calls, which should be borne in mind if estimating data volumes. Where there is a very large volume of analogue audio files, it may be reasonable to extract audio/video data for a specified channel/phone number and between defined date ranges to reduce the volume of data for processing and review, provided that a reasonable, diligent search is conducted: *Quinn and Ors -v- Irish Bank Resolution Corporation (In Special Liquidation) and Kieran Wallace* [2014] IEHC 577.

With digital format, a variety of options available are discussed below.

D.3 Audio data, including sound on video files

There are four primary approaches to managing audio files during discovery. All have associated risks and benefits, and as in all cases the most appropriate approach should be chosen taking proportionality into account.

D.3.1 Computer transcribe and search, then review

Generative AI tools are now available within many eDiscovery platforms that enable auto-transcription to text, which can then be searched and reviewed, or categorized using data analytics tools. AI tools are also coming on stream that allow for simultaneous auto-redaction of both the audio or video data and the transcript.

Once files have been auto-transcribed there are various options, such as applying theme based 'clustering' to the dataset, enabling reviewers to prioritise likely relevant sub-sets, set aside irrelevant material and focus the review; or applying computer assisted learning to predict the likely relevance of individual audio and video files.

The standard of accuracy of auto-transcription will depend on the quality of the recordings involved, whether strong accents are featured, and so on. It follows that the accuracy of the review will depend on the accuracy of the transcription, and the accuracy of the computer assisted learning will also depend on the quality of the transcript. Care should always be taken to go back to the source file if an auto-transcription appears to be of poor quality or contains substantive errors. As with all AI generated outputs, practitioners should be on inquiry for the potential for the system to get things wrong.

Practitioners should therefore ensure that appropriate quality control sampling is conducted against review outputs, particularly against a sample of audio and video files determined not to be relevant, to guard against the risk of errors and log all such checks in the audit document.

D.3.2 Manual review

The review team manually reviews (listens to and/or views) the files and identifies any relevant content, while excluding files with no relevant content from further review.

This more traditional approach to review can be expensive, taking an average of 3 hours to review each hour of audio, given multiple review passes. Time consuming and costly for large volumes of audio data, this method works well for small volumes of audio and video data.

D.3.3 Manual view/listen, transcribe and search, then review

The data is manually transcribed to a searchable text format, by teams of transcribers who listen to the audio data and type out what they hear. This can take up to 2 hours for each hour of audio as it often has to be re-reviewed for accuracy due to issues such as complex multi-voice recordings or audio quality. Once transcribed the data still has to be searched and reviewed, and video content still has to be viewed for relevant imagery.

Slow and costly for large volumes of audio and video data, this method can work well for small volumes of data.

D.3.4 Direct computer index and phonetic search

Specialist technology is available which allows digital audio data to be searched phonetically. This is referred to as 'phonetic searching', where the audio data is searched for sounds rather than text, and the results then reviewed for relevance. Keywords are converted to sounds, which the audio data is then searched against. Considerable effort is required to devise and refine the keywords through an iterative test process to manage error rates and achieve a proportionate degree of quality. These systems are often used in risk monitoring technology for global organisations and are said to provide time and cost savings where there is a high volume of data.

D.3.5 Video files

There are no widely used technologies available for reliably filtering and searching the images in video data at the time of writing, although it can reasonably be expected that the technology will continue to improve during the currency of this Guide. In general therefore it should be assumed that video content has to be manually viewed, although as described above it may be possible to auto-transcribe the audio portion.

It is possible to use technology to identify movement in the images in video data, such as CCTV recordings which record no movement for long periods may be filtered using technology. More advanced automated video filtering technology can recognise patterns in footage, such as number plates or human faces, skin or documents. Such technologies can be considered if video data has to be searched for specific patterns.

D.5 Other data reduction options

In many cases, audio data will contain long periods of non-voice activity, or silence, and video data will contain long periods of inactivity. There are technologies available which can identify these silent portions of data in digital files and suppress them from further searching and/or review, which may help to save costs.

Appendix E Understanding deduplication, families, and threads

Electronically Stored Information (or 'ESI') has brought with it a number of additional relationships between documents which are not so prevalent in hardcopy documents. These concepts play an important factor in managing complex sets of documents through the discovery process. This section describes examples of acceptable methods of addressing these issues but is not an exhaustive statement of the only methods that may be followed.

E.1 Deduplication

One feature of ESI is the level of duplicate information that is generated and stored. For example, if Custodian One sends an email to Custodian Two and both their emails are collected during the discovery process, then both duplicate emails will be present in the dataset. Managing and/or reviewing duplicate documents will generally be a waste of time and money, therefore it is most efficient to suppress duplicates, insofar as is possible, early on in the process. It is important to note that, for the reasons described below, there is no automated means of removing all duplicates from a dataset. Reviewers will inevitably encounter the same, or what they perceive to be the same, documents during their review. The purpose of the deduplication methods described below is to minimise the number of duplicates in a manner that is robust and defensible.

E.1.1 Hash Values

It is possible to identify a digital fingerprint, or 'hash value', for any electronic document. Like a traditional fingerprint, this value (which is just a relatively short alphanumeric code) can be used to uniquely identify a document and thus be used to identify two identical documents (where they have the same value).

One of the first steps (usually completed automatically) during the processing phase is to identify the digital fingerprint or hash value for every document imported into the processing system. Most systems designate the first time it encounters a document as the primary copy and then designates each subsequent copy a duplicate. So, if Custodian One's emails are processed into an eDiscovery system first, their emails will be designated as the primary copy. If Custodian Two's emails are subsequently processed in, their emails will be designated as duplicates. If the emails of multiple custodians are processed at the same time, it is not normally possible to decide whose emails will be designated as the primary copy, so the system may identify the version of an email held by a key custodian as a duplicate while designating the same email held by a more peripheral custodian as the primary copy.

An important feature of any system that performs deduplication for discovery is that they will keep a record within the system of all those custodians who held a duplicate of the document which has been

deduplicated. This information is recorded in a metadata field typically called 'duplicate custodian' and it can be presented with the document for review, allowing the reviewer to see which custodian had the copy of the document they are reviewing, as well as a list of other custodians who held an exact copy of it. This can be significant when a party's state of knowledge or awareness of a particular matter is a relevant issue in the proceedings. In such cases, it may be appropriate to include the 'duplicate custodian' field in the schedule of to the discovery affidavit and/or in the loadfile.

There are a number of different methods and algorithms available for identifying a document's hash value. The most common algorithms are called MD5 and SHA. What is likely to be more important is what portion of the document is used to identified the hash value.

E.1.2 Identifying the hash value of documents

In the case of documents, the most common approach employed by eDiscovery systems is to identify the hash value based on the contents of the document only. For example, the hash value of a spreadsheet would be identified from the contents within the document and not take into account the document metadata, which accompanies the document. This would ignore the fact that the duplicate of this document had a different name and was found in a different location. This is generally most efficient from a discovery perspective, as a detailed analysis of the document's metadata can be performed during the analysis phase if required.

E.1.3 Identifying the hash value of emails

In the case of emails, eDiscovery systems normally identify the hash value based on the contents of the email and the recipient, attachment and header information.

The header information includes sender, email subject and time sent. In each eDiscovery system, the time sent information is formatted in a particular way (e.g. mm/dd/yyyy hh:mm:ss AM/PM) and only those emails with precisely the same time sent information will have the same hash value. Where emails are sent across different time zones, it is possible that two identical emails will have different time sent details, as a result they will not have the same hash value and will not be deduplicated by the eDiscovery system.

Similarly, where an entity has an automated warning that appears on external emails, e.g. "This email originated from outside of the organization", this will alter the contents of the email for hash value purposes. This means that where two copies of the same email have been collected, one from inside an organization with automated warnings and the other from outside, these emails may not have the same hash value and may not be deduplicated.

Some deduplication systems exclude the Bcc address from the hash value identification. It is important to know if this is the case because while it will still be possible to know who had a copy of the email from the duplicate custodians list (assuming their documents were collected), the version of the email which actually showed them in the Bcc address field may have been suppressed through

deduplication. This is significant in cases where the state of awareness or knowledge of specific individuals about various topics is a relevant issue in the case.

E.1.3 Deduplication using hash values

Once the hash value has been identified it is then possible to deduplicate individual documents across the entire dataset. However, unless the entire dataset is comprised of individual documents, and no document families, then this simple deduplication may not be sufficient. The section immediately below describes the more common approach of 'family-level deduplication'.

Note: While duplicate hardcopy documents might be scanned into electronic format, it is almost impossible to apply a traditional electronic deduplication to them. This is due to the fact that each scanned copy will have a different hash value. At best, some of the analytics technologies described in Appendix F might be used to identify similar documents.

E.2 Document families

E.2.1 What are document families?

A document family refers to a set of documents which have a relationship. An example of this would be the contents of a letter and its enclosure. The most common example in discovery is that of emails and attachments. The email is the parent and the attachment is the child. (To add complexity, sometimes the child attachment can be an email, which has its own attachment, referred to as the grandchild.)

By default, in discovery, documents should be considered in the context of their families. One of the key reasons for this is that reviewing a document in isolation may not disclose its true meaning, which may only be apparent when viewed in the context of its family. For example, in a data theft case, a customer list may have one meaning when reviewed in isolation, however when reviewed attached to an email sent by an employee to their personal email address, the meaning of that customer list might change significantly.

It is important to note that when we refer to families in an electronic document perspective, there can only be one parent.

E.2.2. Review of document families

When filters, such as keywords, are applied to a document set, they will return a number of individual documents which are responsive to those keywords. These individual documents may be standalone documents or they may be part of a multiple-document family. If they are part of a multiple-document family, they can either be the parent or the child (or one of multiple children). Therefore, when an individual document is responsive to filtering criteria, it is good practice to consider it in the context of its family for review purposes.

One approach is to review the individual document and determine its relevance before bringing the rest of its family for review. This is efficient in cases in which the individual document is deemed to be irrelevant, negating the requirement to review the rest of the family (in circumstances where no other document in the family is separately responsive to keywords). Only if it is found to be relevant are the other family members brought into the review. An alternative approach is to review the whole family at the same time where one or more of the family members are responsive to the filtering criteria. This can be efficient if the individual responsive family member is relevant, but not so efficient if it is not.

E.2.3 Document families and deduplication

Document families in which the parent and the children are identical will have the same hash value and only one copy of the document family will present for review following deduplication.

However, the same document can be attached to multiple document families, for example, the same spreadsheet may be attached to two different emails. These two emails form two unique document families, in that as a family unit each is unique, but they do contain duplicate children. It is important to consider both families of documents for discovery. As such, both unique parents and two copies of the same attachment would be included for review. It may even be that the same document is determined to be relevant in one instance and not relevant in another, due to its family relationship.

By way of illustration, figure 1 below shows three emails collected from two custodians. The first parent (Parent 1) has two children (Child 1 and Child 2). The same parent (Parent 1) is also found in the second custodian's mailbox, along with the same attachments. This occurs when both custodians are parties to the same email. As such, the copy of this document family in Custodian Two's mailbox would be deduplicated as it is an exact duplicate of the document family in Custodian One's mailbox. The second parent (Parent 2) has two attachments, Child 1 and Child 3. Child 1 is a duplicate of the attachment in the first family; however it is attached to a different parent and a new attachment, Child 3, is in the family as well. As such, both the Parent 1 and Parent 2 document families would be included for review, with Child 1 being included twice so that it could be evaluated in the context of each of its families.

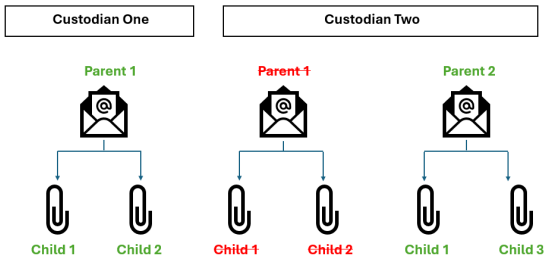

Figure 1 – Two unique families containing duplicate attachments and one duplicate family.

This is referred to as family-level deduplication. The hash value of each combined family is identified and deduplication is completed at a family level. The occurrence of duplicate attachments in this example can be frustrating for reviewers who may have to review the same document multiple times but this type of duplication is a feature of almost every dataset.

E.2.3 Document families and review

When determining the approach to review it is important to decide how document families are going to be treated at production stage[17] as this will determine how they are considered at review.

In the case of standalone documents, such as emails with no attachments and/or loose files, these are simply considered in isolation and marked as relevant or not relevant to the issues in the matter (see Figure 2 below).

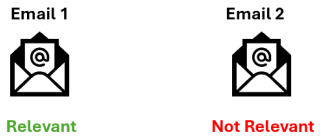

Figure 2 – Standalone documents marked as either relevant or not relevant.

It is good practice not to produce irrelevant documents during discovery. Doing so is generally a waste of time and money. As such, a balance must be struck between presenting documents in the context of their families and not producing irrelevant documents.

This Guide recommends that parent documents are produced when the parent document OR ANY of its attachments are relevant. The production of 'orphan' children (i.e. producing just the relevant attachment without its parent) is not recommended. Irrelevant attachments need not be produced.

Take for example (see figure 3 below) where we have an email (Parent 1) with six attachments comprising five spreadsheets containing customer records (Child 1- Child 5). One of the attachments, Child 1, is responsive to the categories of discovery (it refers to the customer in question), while the other five attachments refer to other customers and contain sensitive personal data. It is recommended in this scenario that the parent (Parent 1) be produced (with references to other unrelated customers within the parent email redacted) along with the single relevant attachment (Child 1). The other five attachments (Child 2- Child 6) should not be produced. While it would be possible to redact these

[17] *This will also impact the format in which the data is processed into the eDiscovery system, e.g. it is not possible to supress irrelevant attachments in a native production if emails have been processed into the eDiscovery system in .msg format.*

attachments in full to protect the irrelevant sensitive personal data of unrelated parties; not producing them on the grounds of irrelevance would be significantly more efficient.

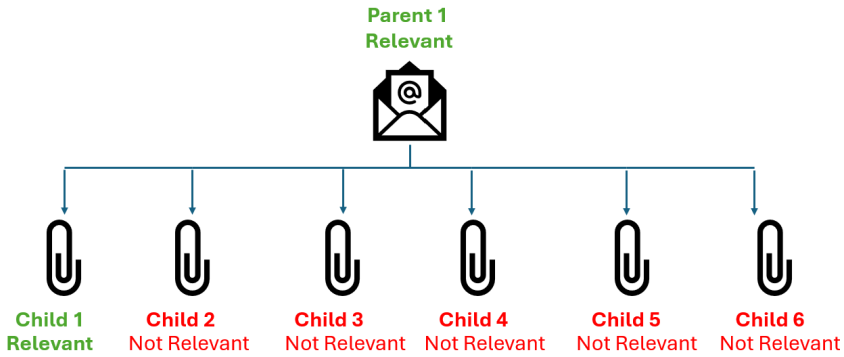

Figure 3 – Parent 1 and Child 1 are marked relevant and for production, whereas five other attachments (Child 2 – Child 6) are marked as not relevant and will not be produced.

Where a parent is found to be relevant, but its children are found to be not relevant, then the parent may be produced in isolation (see figure 4 below).

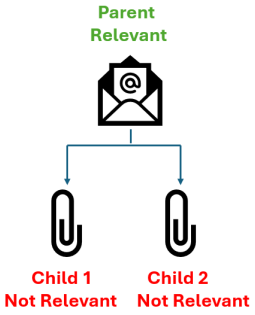

Figure 4 – Relevant parent with non-relevant attachments.

Where a parent email is found to be not relevant, but one or more of its attachments are found to be relevant, then the parent should be produced as well in order to show the context of the child.

Tagging non-relevant parents as 'relevant' can lead to confusion among reviewers, CAL models and those ultimately in receipt of the discovery about the parameters of relevance. Therefore, it is preferable to tag non-relevant parents with a distinct tag that recognizes their status as irrelevant documents which are being produced to provide context.

This Guide recommends that the basis for the inclusion of non-relevant parent documents in a discovery is made clear by tagging such documents as 'Parent for Context' or 'Non-relevant parent'.

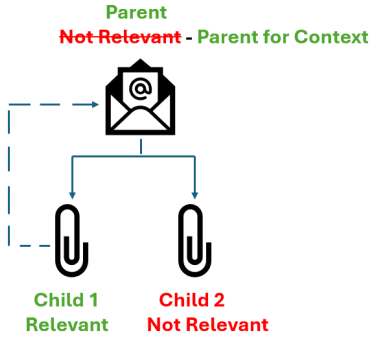

Figure 5 – Non-r ion with child.

It is important therefore, that when reviewing and marking documents in the context of their families that these principles are applied. For example, in figure 5 above, it is often the case that the parent will be reviewed first and marked as not relevant, then Child 1 will be marked as relevant and the Child 2 as not relevant. It is then necessary to go back to the parent and change its tagging to 'Parent for Context' or 'Non-relevant parent' to ensure that it is included in the discovery schedule and in the production.

To summarise, this Guide recommends that document families are produced as follows:

- If a parent is relevant, but its children are not, then only the parent would be produced.
- If a child is relevant, then its parent would also be produced. Orphan children would not be produced in isolation.
- If there are multiple children, not all of which are relevant, then only the relevant children and the parent would be produced. Irrelevant attachments would not be produced.

Further, checks should be performed before production to ensure that no irrelevant and/or orphan children are produced. It may be helpful to include a schedule of irrelevant family members which have not been produced, and/or include a slip sheet for each document which has not been produced. This can assist is demonstrating that the document has not been produced intentionally, rather than due to a technical issue or oversight.

E.3 Email threads

E.3.1 What are email threads?

When an email is sent (from Custodian One to Two) it is stored as an individual message (Email 1). If Email 1 has an attachment (Child 1); this is sent attached to the message [**Email 1**/Child 1]. The person who receives

the message, (Custodian Two) receives it as [**Email 1**/Child 1]. Both the original message and attachment are stored as two related documents on the sender's (Custodian One's) system and on the recipients (Custodian Two's) system. As the hash values would be the same for both, this document family would be deduplicated as outlined above.

If Custodian Two replies to the message, this generates a new email thread that includes the reply plus the original message, Email 1. This new email thread is [**Email 2** + Email 1]. Typically, the attachment, Child 1 is dropped when the email thread [**Email 2** + Email 1] created and it is not included in the new email chain.

Custodian One receives the reply from Custodian Two, so now Custodian One has [**Email 1**/Child 1] in their sent items and [**Email 2** + Email 1] in their inbox, while Custodian Two has [**Email 1**/Child 1] in their inbox and [**Email 2** + Email 1] in their sent items.

In this simple exchange, the document family [**Email 1**/Child 1] needs to be reviewed because it contains the original attachment, Child 1. Email thread [**Email 2** + Email 1] also needs to be reviewed because it contains the reply. While [**Email 2** + Email 1] will contain the original text from Email 1, it does not include the attachment, Child 1.

If the example is extended so that Custodian One replies to [**Email 2** + Email 1], a new longer email thread [**Email 3** + Email 2 + Email 1] is created. This thread contains the new reply plus the original text of Email 1 and Email 2.

Custodian Two receives [**Email 3** + Email 2 + Email 1] and replies to it again, which is another new longer email thread [**Email 4** + Email 3 + Email 2 + Email 1]. that is received by Custodian One, shown in figure 6 below.

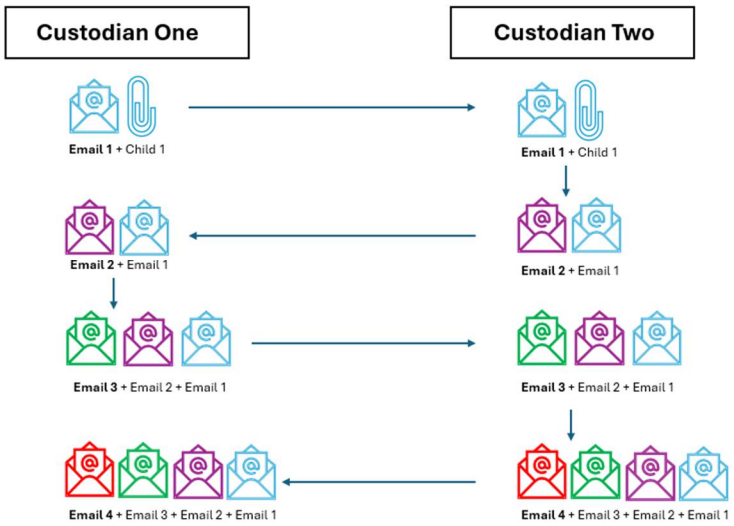

Figure 6 – Email thread between two custodians.

We now have a copy of the individual message [**Email 1**/Child 1] and the email threads [**Email 2** + Email 1], [**Email 3** + Email 2 + Email 1] and [**Email 4** + Email 3 + Email 2 + Email 1] as four individual messages with both custodians. If we collect emails from both custodians for discovery, our first step is to deduplicate both sets of messages. If Custodian One's emails were processed first, then Custodian Two's will be suppressed through standard family-level deduplication and we will be left with Custodian One's copy of message [**Email 1**/Child 1] and the email threads **[Email 2** + Email 1], [**Email 3** + Email 2 + Email 1] and [**Email 4** + Email 3 + Email 2 + Email 1]. Figure 7 below shows that a 50% saving on review effort can be achieved using this simple family-level deduplication method.

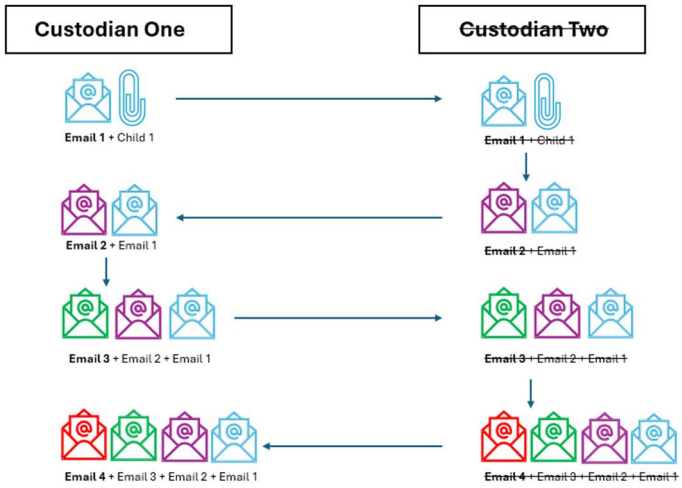

Figure 7– Standard family – level deduplication on an email thread between two custodians.

The challenge is that we have the original [**Email 1**/Child 1] message and attachment; we then have a each of the three email threads which all contain the content of the previous emails, in addition to their new content.

Even after standard family-level deduplication, there is a lot of duplicated text within the email threads. This is due to the fact that each email thread is typically not stored as a single composite thread, but as individual threads, each of which are collected and included in the discovery process.

What further complicates the challenge is that each individual thread will have a date associated with it, which will be the date it was sent or received. If the four emails which are outlined in the example above were sent days or weeks apart, then their dates would be days or weeks apart. If these emails were in a document set with 10,000 other emails, and that document set was sorted for review by date, then the emails would appear quite a distance apart, with many unrelated emails in between. Each email thread could be reviewed by a different member of the review team and a lack of consistency in approach to relevance, privilege or category could result.

E.3.2 Email threads and sorting/deduplication

There is technology readily available which can identify the individual messages which comprise email threads and link them together. This is known as email threading. Email threads will include the initial email and all subsequent replies and forwards linked to that original email. In the above example, this technology would identify the connection between [**Email1**/Child 1], [**Email 2** + Email 1], [**Email 3** + Email 2 + Email 1] and [**Email 4** + Email 3 + Email 2 + Email 1] and present them for review in that order. Sorting by email thread is very useful as each message is shown linked to the next in the thread, which increases the speed and accuracy of the review. However, it still involves the review of the duplicate text contained in [**Email1**/Child 1], [**Email 2** + Email 1], [**Email 3** + Email 2 + Email 1] and [**Email 4** + Email 3 + Email 2 + Email 1] rather than just reading the whole text which is contained in [**Email 4** + Email 3 + Email 2 + Email 1].

The solution to this challenge is referred to as 'email thread deduplication'. Such technology extends on threading identification by also designating each message as 'inclusive' or 'non-inclusive'. In the example above, [**Email1**/Child 1] and [**Email 4** + Email 3 + Email 2 + Email 1] would be designated as inclusive and would be reviewed, while [**Email 2** + Email 1] and [**Email 3** + Email 2 + Email 1] would be designated as non-inclusive and could be suppressed as duplicates from review.

One might assume that [**Email 4** + Email 3 + Email 2 + Email 1] is the only inclusive email as it contains all of the text from all of the individual messages, however because the attachment, Child 1 was dropped after [**Email 1**/Child 1], it is necessary to include [**Email 1**/Child 1] so the attachment is available for review. The email thread would be sorted for review in the order [**Email1**/Child 1] and then [**Email 4** + Email 3 + Email 2 + Email 1], allowing for a more efficient and accurate review.

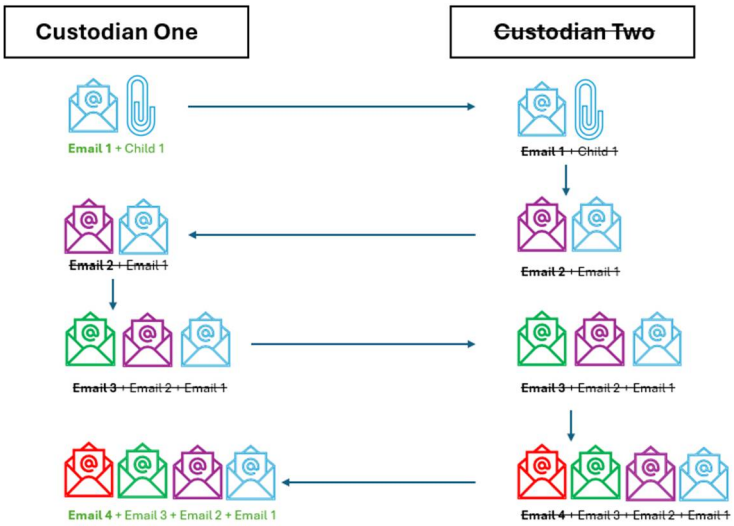

Figure 8– The remaining email thread deduplicated within the thread.

As can be seen from figure 8 above, the number of emails and attachments for review is reduced from 11 to 6, some 45% reduction in review effort (in addition to the 50% already achieved through standard family-level deduplication). This can increase further with longer email threads. Where email thread deduplication has been utilised, there is generally no requirement to reintroduce duplicate/non-inclusive portions of the email thread at a later date. Note: It is not possible to sort by date and by thread as the two concepts are mutually exclusive.

The same principles apply to email chains that are forwarded to a third party and a new email chain involving that third party commences. Emailing threading will identify the most inclusive forwarded email chain in addition to the most inclusive email between the original custodians.

There are some potential downsides to using email thread deduplication, these include the potential for additional redactions where long email threads contain irrelevant, privileged or sensitive personal information. For example, a long email thread in which the first email has no attachment and is both relevant and not privileged but all other messages in the thread are privileged or comprise non-relevant commercially sensitive information. The bottom email will not appear in the review on its own as it is a non-inclusive email so redactions will need to be applied to all of the other messages in the chain. In this example, it may be appropriate to identify the first email as a standalone message and produce it but this may not be possible for email threads where relevant, not privileged messages are dispersed amongst a long otherwise privileged thread of emails.

A further downside of using email thread deduplication is the inability to carry out a focused review of the dataset by reference to a specific date range. This type of focused review can be useful if a particular reviewer or sub-set of reviewers are instructed to review all the documents that could be potentially relevant to a discovery category that has a strict date limitation. Sorting inclusive emails by date will mean they are sorted by reference to the date of the latest email in the thread and not by reference to the dates of the intermediate messages in the thread.

A further example of problems that may arise with e-mail thread deduplication arises where an organisation's server automatically incorporates a warning into external e-mails (e.g. "This e-mail originated from outside of the organization.") E-mails within a conversation which includes autogenerated text of this type may not be identified as part of the same thread group.

These downsides must be balanced on a case by case basis with the overall efficiencies to be gained by email thread deduplication.

This Guide recommends that consideration is given to the utilisation of email threading and deduplication in every discovery as they can improve the accuracy and speed of document review and also reduce the cognitive burden on reviewers, who can focus on the truly unique content rather than sifting through repetitive information.

Appendix F Technology Assisted Review

Technology has been used for many years to assist in the discovery process. In earlier years, document management systems were used only to manage documents which were known to be relevant to the proceedings and were used principally to view and mark such documents. In the past two decades, however, it has become the norm to employ more sophisticated document management systems in order to process, filter and review documents to determine whether they might be relevant to a case. These 'review platforms' and their underlying processing technologies were the first form of Technology Assisted Review (or 'TAR') in that a computer was used to assist in the review process, rather than printing the documents for manual review. Computer Assisted Review (or 'CAR') is another term for TAR.

Technology in the area of eDiscovery is developing at pace. This appendix outlines the current practices with regard to Technology Assisted Review. (Other technologies are being developed and/or are in use, and their absence from this chapter does not imply they are unsuitable for assisting with document review in eDiscovery.)

One such technology is Generative Artificial Intelligence ("GenAI") discussed further in this Appendix. GenAI has mass document capabilities that are likely to be transformative to aspects of discovery workstreams. There is already an emerging market of GenAI products that are specifically tailored for discovery review. Historically, the Irish Courts have been quick to adopt technological solutions which offer assistance when addressing the unique challenges inherent in modern discovery, particularly those associated with the proliferation of data. While GenAI discovery tools are in their infancy, because of their capacity to reduce the hours inherent in conducting large-scale commercial discovery reviews, this technology is likely to play a significant role in the future of commercial discovery in Ireland. Transparency in its use, and control mechanisms to audit, test and verify outputs, will be critical to the acceptance of GenAI by legal practitioners and the Courts in Ireland.

F.1 What is TAR?

There are many ways in which technology may be used to significantly increase the accuracy and speed of a review, thus reducing risk and cost. These can be split into two broad categories; those which arrange the documents in such a way as to make them easier for practitioners to review (referred to as Analytics), and those in which the computer programme is "trained" to identify relevant documents (known as Continuous Active Learning (CAL) or its predecessor, Predictive Coding). Both sets of technologies are often referred to under the umbrella term of TAR. Different technology platforms may refer to these concepts using different terminology.

Recent developments in the area of Generative AI (GenAI) have also had an impact in the manner in which technology can be utilised to assist in identifying relevant documents. While the use of GenAI in document review is still in its early stages, we have included a section at the end of this chapter on the current developments and emerging use case of GenAI in this area.

F.2 Analytics

Analytics technologies arrange documents in ways which make it easier to carry out a traditional review. They do this by automatically extracting relationships and patterns from documents without human intervention. These technologies focus on identifying concepts within document sets and are traditionally viewed as being low risk but leading to moderate reward. In eDiscovery, risk refers to the risk of omitting relevant documents, inconsistent review decisions, budget overruns, etc. Rewards refer to the reduction of time and cost, as well as the reduction of risk. It can be very useful to employ analytics techniques during the processing phase in order to identify groups of relevant or irrelevant documents and manage them accordingly.

F.2.1 Concept Clustering

This technology groups documents together that have been identified as conceptually similar (e.g. they contain shared keywords, concepts, or semantic meaning), allowing, for an overview of the entire data set and the identification of themes or patterns within the documents.

F.2.2 Near Textual Deduplication

This technology groups documents together that are identified as almost identical in content allowing, for example, multiple drafts of the same document to be reviewed together. Unlike exact deduplication, which identifies documents which, on a technical basis, are 100% identical, near textual deduplication focuses on documents that are textually very similar and identifies those documents as "near duplicates".

F.2.3 Email threading

A detailed description of email threads and their sorting and deduplication is set out in Appendix E. While email thread sorting and deduplication is a form of analytics, they have become standard practice in most cases.

F.3 Continuous Active Learning (CAL)

The analytics technologies outlined above focus on making the human review process more efficient and do not involve the computer making decisions as to whether a document is relevant.

Predictive coding (the predecessor to Continuous Active Learning (CALs)) is a process whereby the technology uses machine learning algorithms to identify relevant documents in a document review. The more modern, dynamic form of predictive coding is known as CAL where the system continuously learns from reviewer decisions in real time and re-ranks the remaining unreviewed documents in a review set based on predicted relevance.

This technology allows for real-time updates and prioritisation of documents continuously improving the accuracy and efficiency of the identification of relevant documents. Similar methods may also be employed to determine if a document is privileged.

As the reviewers review and code the relevancy of documents, CAL learns to more accurately predict the likelihood of the remaining documents in the set being relevant. Documents with a likelihood of being relevant above a certain probability can then be manually reviewed and their relevance verified. CAL provides two ways to approach this:

- Prioritised Review which finds the documents most likely to be relevant.
- Coverage Review which aims to quickly separate documents into two categories - relevant and not relevant.

CAL and its predecessor, predictive coding, have been accepted for use by the Courts in various jurisdictions, including the US, UK, and Ireland. It has been shown through numerous studies to take significantly less time and cost than a traditional keyword-driven manual review. It has also been shown to typically provide a far greater level of recall and precision than the traditional linear approach. Although CAL will not usually capture 100% of all relevant documents, no proportionate review method will achieve this statistical perfection. Moreover, CAL technologies typically identify more relevant documents than the traditional keyword and linear review approach. Although there is no specific rule in Order 31 authorising its use, the Irish courts have taken the view that, provided the process used is sufficiently transparent, TAR using CAL discharges a party's discovery obligations under Order 31, rule 12 RSC. The process must, however, contain appropriate checks and balances which render each stage capable of independent verification. A balance must be struck between: (i) the right of the party making discovery to determine the manner in which discovery is provided; and (ii) participation by the requesting party in ensuring that the methodology chosen is sufficiently transparent and reliable.

It has been widely accepted that it is not necessary for users of CAL technology to be experts in the underlying maths and statistics in order to effectively and safely use such systems. It would be akin to suggesting that one should be an expert in internal combustion engines to be a proficient driver.

CAL will not normally be employed in the review process until all data has been collected, processed into a searchable format, and family-level and email thread deduplication have been completed. At that time, early case assessment will allow an analysis of the data to be completed and a determination as to the suitability of CAL for the matter.

The document set will be split into those sub-sets which will be suitable for CAL and those which are not (such as, for example, handwritten notes which have been scanned, documents which do not have sufficient amounts of text such as pictures, photos, graphics, spreadsheets, or handwritten documents and documents that have an excessively large amount of text).

CAL comprises of four main steps:

- **INITIAL REVIEW** – A subject matter expert (a senior lawyer with detailed knowledge of the case) is presented with a small, representative sample of documents - often randomly selected but also may be selected using an agreed selection process - from the entire dataset (referred to as the 'Initial Assessment Set' or 'Seed Set'). The system requires a minimum of 5 relevant and 5 not relevant documents in order to build the active learning model. The subject matter expert codes these documents as relevant or not relevant to the issue(s).

  Note: Documents should also be marked as Technical Issue during the Initial Review step in the event that they cannot be opened or reviewed by the expert reviewer due to a technical issue. The producing party should then work with its technology provider to resolve such technical issues (if possible) and allow the documents to be reviewed and marked as relevant or not relevant.

- **MODEL INITIALISATION** – With the initial review carried out, the system uses the coding from that review to train the model which then predicts the relevance of uncoded documents in the dataset based on their closeness to coded examples. This is known as ranking the documents; the system will rank the documents with a score of between 0 and 100 indicating the likelihood of their relevance.

- **REVIEW AND CONTINUOUS LEARNING –** Once the model has been built and the initial ranking has been assigned, the review team will commence the review. The system will present the highest ranked documents along with a statistical sample of low and medium ranked documents. At this stage of the review, there will be a number of 'false positives' presented to the reviewers. False positives are documents the system has assigned with a high ranking but which the reviewers identify and code as not relevant. As each document is reviewed and coded, the system learns from the coding and re-ranks the remaining uncoded documents. This process continues in real time throughout the review process, allowing the model to adapt and improve continuously. As the system learns from the reviewer, the number of false positives being presented for review will reduce as the relevant documents are prioritised for review.

  Note: As CAL determines the ranking of the document based on the content of the document itself (and not on outside factors such as family members, date range, custodian information etc.), at all stages of the document review, the reviewers should follow the "four corners" rule. This means that a document should only be coded based on text within the body of the document and not based on any surrounding factors.

  Assuming CAL is being used in place of a traditional manual/linear first pass review, this allows the review team to determine the cut-off point based on risk and cost (see below for other use cases leveraging the output of continuous active coding). The cut-off point is the rank below which the technology has determined the documents are unlikely to be relevant. All documents above the cut-off point will be manually reviewed for relevance and those below the cut-off point will be subject to statistical sampling, but otherwise not reviewed. Once the review has reached a stage where the reviewers are, for a sustained period of time,

no longer being presented with relevant documents by the system, the existence of relevant documents remaining in the uncoded data set is statistically very low and this can be determined to be the "cut off" point.

- **PROJECT VALIDATION** – Once the cut-off rank has been reached (assuming CAL is being used in place of a traditional manual/linear first pass review), then it is necessary to carry out a statistical quality test to estimate if any relevant documents may have been missed by the system. The aim of project validation is to estimate the accuracy and completeness of the relevant document set assuming the review was to stop immediately and any unreviewed documents below the cut-off point were not produced. The project validation involves a review by the subject matter expert of a random sample of documents that the system has predicted to be not relevant. The number of documents to be included in the sample will depend on the desired level of confidence and margin of error. The number of relevant documents identified as a result of the Project Validation is known as the "elusion rate".

  Once the review is complete, if few or no relevant documents are found ie the elusion rate is low, this would indicate that the project was successful and the likelihood of any relevant documents remaining in the uncoded document set is statistically very low. If, as a result of the review, many relevant documents are identified and the elusion rate is high, this suggests that the review needs to continue and the model re-trained.

This typical application for CAL is to replace a traditional manual/linear first-pass review for relevancy, but CAL can also be used to complete a second-pass manual review to bring in family members, confirm relevancy, and assign categories, assess privilege and complete redactions, if required. The difference is that this second-pass review is likely to cover all relevant documents, with few, if any, false positives.

There are also a number of other applications for CAL once a probability of relevance has been applied to a set of documents. These include:

1. **Prioritising documents for review** – Documents can be prioritised by those most likely to be of relevance to the matter. This can allow the most relevant documents to be reviewed first by a core review team and documents likely to be irrelevant to a secondary review team.
2. **Remove irrelevant documents** – As part of Early Case Assessment (or '**ECA**'), CAL can be used to identify and separate documents which are clearly irrelevant to the matter.
3. **Verification of keywords** – CAL can be used to perform a discrepancy analysis between those documents identified as potentially relevant through keyword and other filters, and those which the active learning system identifies as potentially relevant. This can be a useful tool at the processing phase when keywords and other filters are being devised.
4. **Quality control during review** – As the review progresses and/or upon its completion, CAL may be used to analyse discrepancies between the human review decisions and the computer. This can be particularly useful to analyse discrepancies in respect of privilege coding at the end of a document review.

As with all technology, CAL is better suited to some matters and not others. CAL is not well suited to matters where:

- There is a large amount of non-text-based data, such as pictures or spreadsheets. (Some predictive coding systems work well with numeric data, whereas others do not.)
- Handwritten documents which have been scanned to electronic format or typed documents where handwritten notes in the margins would be a factor in determining relevance.
- There is a large volume of hard copy documents.

In all matters where CAL is considered, expert advice from a technology service provider (internal or external) should be sought in order to determine if the use of CAL is helpful given the circumstances of the matter and the data types involved. As outlined above, it is usually not possible to understand the data fully until it has been collected and processed into a searchable format and deduplicated at a family level and by email thread. As such, it may not be possible to determine whether CAL will be helpful in advance of that stage in the process.

As with analytics technologies, different CAL technology vendors use different terminology when describing their processes and some of them employ slightly different processes.

F.4 Additional cooperation

As with standard analytics technologies, CAL technologies are commonly accepted and used and it is not typically necessary to notify the requesting party as to their use. The use of CAL will be set out in the Discovery Affidavit.

The time required for the review will depend on data volumes and the richness of the data, that is the preponderance of likely relevant data within the dataset, and it will likely be necessary to defer providing an estimate for the review and production until after the completion of the CAL exercise.

Key to the success of a CAL project is having a system which is transparent, allowing the user (and possible the requesting party and the Court) to see and understand why the system made a decision on any particular document. Equally important is having a statistically valid process and validating the results. It is important to understand that the CAL process, just like the traditional keyword search and review approach, is not perfect. As such, it should be measured against this alternative and not against unattainable perfection.

The discovery plan at Appendix G contains sample text which may be used to provide information to the various parties involved in a matter regarding analytics and/or CAL.

F.5 Generative AI (GenAI)

As mentioned above, document review can now be streamlined through the use of GenAI-powered tools that can analyse vast volumes of data with speed and precision. eDiscovery document review platforms have recently introduced GenAI capabilities that allow legal teams to conduct first-pass review, identify privileged content and generate summaries of documents with greater speed and accuracy. By automating first-pass review and reducing the need for manual training and review the use of this technology can significantly reduce time and costs associated with eDiscovery. The use of GenAI in eDiscovery can transform what would typically take weeks of manual review into just a few days.

GenAI leverages large language models (LLMs) to automatically classify documents as relevant, irrelevant or privileged. Unlike TAR, which relies on extensive manual coding of documents to train the model, GenAI can understand context and nuance across a number of formats – including email, chat, PDFs, spreadsheets, images and video - without extensive initial training from legal teams.

One of GenAI's features is its ability to generate concise summaries of documents. This allows reviewers to quickly understand the content and relevance of documents without reading the full text. Some document review platforms also use GenAI to create explanations for decisions (e.g. why a document is relevant which can be used by reviewers to analyse the results and review the accuracy of the decision making by the GenAI tool.

As GenAI continues to evolve, its integration into eDiscovery workflows marks a significant shift in how legal teams manage and analyse data for document review. While its adoption is still evolving, early results demonstrate significant benefits from its use. in particular in respect of efficiency, accuracy and strategic insight. As legal teams become more familiar with its capabilities and limitations. GenAI is poised to become an integral part of modern eDiscovery document review process.

This guide represents best practice at the time of publication. However, given the pace at which GenAI, is evolving, it is recognised that there may be subsequent developments in GenAI that are not covered by this Guide and their absence from this chapter does not imply they are unsuitable for assisting with document review in eDiscovery.)

Appendix G Sample discovery plan

[Matter reference]

Discovery Plan

[Version 0.1]

[Date]

G.1 Background

This discovery plan (the 'Plan) sets out the steps that [Producing party] has taken to date and will take in future for the purposes of making inter party discovery in [Matter reference]. This is prepared for internal purposes, it is essentially an audit trail and will be an important record if discovery is challenged or questions are raised.

[[Producing party] has engaged the services of [service provider] to assist in this matter. [Service provider] intends to utilise [technology platform(s)] in carrying out the services for the purposes of [producing party] making discovery.]

Notwithstanding this plan it remains the obligation of [producing party] and its solicitors to identify relevant documents and make discovery of such documents in accordance with Order 31 Rule 12 of the Rules of the Superior Courts, 1986 (as amended).

G.2 Summary

The issues in this matter relate to [insert summary]. These issues relate to activities undertaken between [date] and [date].

G.3 Scope

This plan governs all document sources, both electronic and hardcopy, which will be included in the discovery process and provides for how documents will be managed throughout the process, from initial identification through to final presentation in Court.

G.4 Approach and progress to date[18]

A phased approach has been taken to this process, consisting of the following eight phases:

[18] *This sample plan assumes that the producing party is close to the end of Phase Four. i.e. data is collated and filters/searches run, but review not yet commenced. Therefore it provides what has been done to date up to and including Phase Four and what is proposed to be done in the remaining phases. This may need to be amended.*

- Phase One – Identification – the identification of potential custodians and document sources which may contain documents of relevance to the matter.
- Phase Two – Preservation – notifying the custodians and other parties (by way of a litigation hold) of their duty to preserve documents and taking steps to help ensure that documents may not be lost in advance of collection.
- Phase Three – Collection – working with the custodians, IT teams, forensic collection agents and/or other parties to obtain a copy of potentially relevant documents from the document sources identified.
- Phase Four – Processing – converting the document sources collected into a format to facilitate their efficient searching and review, including the use of email threading. Documents were also filtered, using filters such as date range and keywords. [Phases One, Two, Three and Four have been completed [OR are almost complete]. Details of the steps carried out in the four phases can be seen below.]
- Phase Five – Review – documents responsive to the filtering criteria and any documents that do not require filtering to be brought forward for review by [using Continuous Active Learning (CAL)] to narrow the set of documents to those likely relevant and then] manual review to determine the relevance and privilege status of each responsive document.
- Phase Six – Analysis – quality checking and technical analysis as required, for example to determine the provenance of a document.
- Phase Seven – Production – at the conclusion of the review and after quality checks have been completed, generation of discovery schedules and export of documents for disclosure.
- Phase Eight – Presentation – preparation of documents for presentation in court in a manner which facilitates their efficient presentation and the running of the matter.

G.5 Phase One – Identification

The objective of the identification phase is to identify potential custodians and sources of potentially relevant documents.

The first step in this phase was to identify a list of custodians who may hold or have held relevant documents. Custodians included individuals and external organisations who may hold documents on [producing party]'s behalf. The list of custodians collated is at Attachment One below.

We worked with the custodians to identify potential document sources, which included:

- Understanding the likely document types and date ranges through discussions and interviews, using the CLAI custodian questionnaire.
- Interviewing custodians to understand how they typically utilised technology and managed documents.
- Interviewing the IT team [and outsourced IT providers] to understand how data is managed from a technical perspective, using the CLAI IT questionnaire.

We identified the following custodian document sources [note: the below list is not intended to be comprehensive and these sources may not apply for every discovery]:

- [Live email, and instant messaging from each custodian's email account.]
- [Email and instant messaging from backup/archive systems.]
- [Documents from each custodian's private network folder.]
- [Documents from each custodian's cloud storage
- [Documents from each custodian's mobile device(s) and tablet(s)]
- [Documents from each custodian's computer.]
- [Documents from backup/archive systems.]
- [Structured records from accounting/HR systems.]
- [Documents from social media / social networking accounts]
- [Hardcopy documents from personal filing systems.]
- [Hardcopy notebooks]
- [Text messages and WhatsApp messages from custodians' personal mobile devices]

We also identified the following non-custodian data sources:

- [Documents from shared network/project folders/cloud storage.]
- [Structured records from a manufacturing/quality control system.]

- [Hardcopy documents from centralised filing systems.]

We prepared this discovery plan in addition to identifying and addressing data privacy and security concerns with [producing party]. A plan to preserve and acquire copies of the potentially relevant documents from the document sources identified was then formulated.

G.6 Phase Two - Preservation

The objective of the preservation phase is to take steps to preserve documents where they exist, so that they may not be altered or destroyed in advance of collection. Preservation took two primary forms; a legal hold notice and technical measures.

[Producing party] issued a legal hold notice to all potential custodians within [organisation/department name] and [outside service providers who may hold data] on [date]. This notice instructed custodians to retain and not alter or destroy any potentially relevant documents, including electronic data. This notice was also sent to [producing party]'s IT team so that [producing party]'s data retention and deletion policy could be suspended where appropriate. [Producing party] complied fully with the CLAI guidance and obtained explicit acknowledgement from all custodians. A reminder notice was issued [weekly/monthly/quarterly] until all document sources had been collected.

In addition to the legal hold process described above, [Producing party]'s IT team implemented the following technical preservation steps on [date]:

- [The data retention policy and automatic deletion of [email server/file server/application server/cloud storage] was suspended in respect of the identified custodian and non-custodian data sources]
- [Technical controls were implemented which prevented custodians from altering or deleting SMS or WhatsApp messages[19].]
- [Technical controls were implemented which prevented custodians from altering or deleting historical email.]
- [Technical controls were implemented which prevented custodians from altering or deleting historical files.]
- [Access to hardcopy documents were restricted to [litigation team].

All document sources identified during Phase One above were included in the preservation process.[20]

[19] *This includes custodians turning off the disappearing messages function in messaging apps.*
[20] *If all sources were not preserved, then an explanation as to what was excluded and why should be included here.*

G.7 Phase Three - Collection

The objective of the collection phase is to copy potentially relevant documents from the sources identified so that they can be processed and searched for relevant documents.

The custodian data sources identified in Phase One above represents all documents associated with each custodian and as such, will contain a large volume of documents which have no relevance to this matter. It may also contain personal data and, if so, regard must be had to compliance with data protection law, in particular the data protection principles set out in Article 5 of the GDPR, including the principles of data minimisation. accuracy and purpose limitation. Having regard to this and the general principles of necessity and proportionality in discovery, we have not collected entire custodian data sources but [Producing party], has applied the filters outlined in Attachment Three to the custodian data in order to identify only those documents which relate to this matter. [Producing party] then arranged for the resulting documents to be copied and/or extracted. This was completed for all document sources which existed at the time of the collection exercise, details of which can be seen in Attachment One below.

The copying/extraction was completed using different approaches, depending on the document source involved:

- [Hardcopy documents were scanned into electronic format and made searchable through an OCR process. Metadata associated with the documents was compiled into an electronic format as well through a manual coding process. Note: Unlike electronic document sources, where the full document source was acquired, a focused collection of only potentially relevant documents was completed for hardcopy documents.]
- [Electronic document sources were searched and extracted using industry standard tools, such as [name of tools], which help ensure that the original documents and metadata was preserved throughout the extraction process.]
- [The SMS and WhatsApp messages from mobile devices relating to this matter were identified and exported using [name of tool].

[A chain of custody has been maintained for all document sources acquired.]

G.8 Phase Four - Processing

The objective of the processing phase is to remove clearly irrelevant data types and to convert the remaining documents into a format which will facilitate their efficient searching and review. Documents may then be filtered, using filters such as date range, email addresses, keywords, and other analytics, before being brought forward for review.[21] The eight steps outlined below were undertaken during the processing phase.

G.8.1 Remove irrelevant document types

The dataset collected was loaded into an eDiscovery processing system, [name of system]. In the case of each custodian's email data and cloud storage a large volume of irrelevant document types and non-user created data existed, these were removed during the processing phase by the [eDiscovery system] platform.

G.8.2 Convert into searchable format/load into database

The dataset, less non-user and irrelevant file types was found to comprise [number] documents (emails, their attachments and other loose files). This document set represents all of the documents acquired for each custodian, as well as the non-custodian document sources and the hardcopy document sources.

G.8.3 Deduplicate

A family-level deduplication process was run against all documents, suppressing any duplicate families of documents while publishing one copy of each unique family of documents. The list of custodians who held a duplicate family which was suppressed has been recorded and included in the remainder of the process. This allows only one copy of the family to be considered, while also allowing reviewers to quickly understand who held duplicates of each family. The resulting document set consisted of [number] documents from unique families of documents.

G.8.4 OCR

[number] non-searchable documents were identified, including [document types such as PDF and TIFF]. An OCR process was run against these documents in order to convert them to a searchable format.

G.8.5 Apply date range(s)

The date range(s) outlined in Attachment Three were applied to all documents to identify in order to reduce the dataset to documents within the time periods [agreed in the categories of discovery] OR [ordered by the Court].

[21] *Such review may or may not include the use of Technology Assisted Review.*

G.8.6 Thread deduplication

Email thread deduplication[22] was run against all emails and their attachments to identify the inclusive portions of each email thread, along with the non-inclusive (or duplicative) portions of the email thread. The non-inclusive portions of the email threads were then suppressed from further processing. Email threads which were unable to be subjected to the email threading process have not been suppressed and have been included in further processing.

G.8.7 Manage problem documents

[number] [password protected/encrypted] documents have been identified in the remaining document set, the contents of which will not be accessible for keyword searching.[23] We will attempt to access these documents through decryption techniques, where their name, location, and/or associated document (such as parent email) are responsive to filtering criteria.

We will attempt to access any documents that are not subject to filtering [such as those in shared electronic project folders].

[Describe any other class of problem documents and how they will be managed.[24]]

G.8.8Apply keyword filters and perform Early Case Assessment

Hardcopy documents [and other document sources, such as cloud or other shared electronic project folders] have not been subject to filtering. Shared project folders are document repositories for this discovery exercise are likely to contain relevant documents.[25]

The resulting document set, numbering [number] documents, represents all documents collected from each custodian using broad keyword filters after date filters were applied and as such, the will contain a large volume of documents which have no relevance to this matter. It would not be proportionate or practical to manually review each document for relevance.

[The Producing Party] has therefore applied the keyword filters outlined in Attachment Three to the document set in order to identify potentially relevant documents, resulting in [number] of responsive documents for manual review.

[Producing party] has worked with their legal advisors to test the filtering criteria for precision and recall and to highlight for review documents likely to be of relevance to the matter, while seeking to reduce the volume of irrelevant documents requiring manual review. [This testing involved initial early case assessment

---

[22] See Appendix F of the CLAI Good Practice Discovery Guide v2.0 for a detailed explanation of email thread deduplication
[23] their location and names, including metadata would however likely be searchable.
[24] Documents that have become corrupted and are not accessible but are known to be relevant should be listed in the Second Schedule to the Affidavit as to Documents.
[25] If the parties have agreed an 'end date' for discovery to facilitate collection of data the shared project folders may fall outside the discovery to be made, and this may not apply.

(or '**ECA**') using analytics tools such as clustering, categorisation, and themes, in addition to sampling the results of each filter.]

G.8.9 Publish for Review

[Responsive documents and those not susceptible to filtering will be published for review. Their families will also be uploaded (i.e. where an attachment is responsive, its parent email will also be uploaded). Where duplicates of responsive documents exist within another unique family of documents, the other unique family of documents will be published for review (i.e. where the same attachment is attached to two different emails, both emails and two copies of the attachment will be published). This allows decisions regarding how duplicates and families of documents are managed to be made throughout the review phase.]

Or

[It is proposed that CAL, be used at the initial review phase. While all unique families of responsive documents will be published for review, only one copy of each document will be included in the CAL review. Where appropriate the parents of all copies of documents identified as relevant will be reviewed separately either during the CAL review itself or manually during the second pass review.]

G.9 Phase Five - Review

[[If CAL is in use] The objective of the review phase is to utilise CAL to highlight documents of potential relevance and then perform a manual review of those documents. This will be completed by [review team]. Each document will have a system generated determination made as to its likely relevance to the issues in the matter and those identified as potentially relevant will be manually reviewed for relevance to categories, privilege and redactions, where appropriate.]

OR

The objective of the review phase is to identify all relevant documents using manual review and mark documents as privileged, relevant to specific categories and requiring redaction, where appropriate. This will be completed on a document by document basis by [review team].

[Choose Model 1, 2, or 3 below and delete other content. See Appendix I of CLAI Good Practice Discovery Guide for guidance on different approaches to review.]

Note: This plan outlines the current approach to Technology Assisted Review (TAR). While emerging technologies - such as Generative Artificial Intelligence (GenAI) - are being developed at pace and show promise in handling large volumes of documents, their application in document review is still in its early stages. As such, GenAI has not been included in this plan. However, its exclusion should not be interpreted as a judgment on the suitability of it or other available technologies.

G.9.1 Two-pass review with CAL [Model 1]

A CAL review will be undertaken consisting of:

- **First Pass –** One copy of all documents plus their parent emails (where the document is an attachment) which remain after date and keyword filtering and family-level and email thread deduplication is applied will be included in this review pass.] The CAL review process comprises four main steps:
    - **INITIAL REVIEW** – The expert reviewer will be presented with a small, representative sample of documents which will be randomly selected [OR will be selected using an agreed selection process] - from the entire dataset (referred to as the 'Initial Assessment Set' or 'Seed Set'). The expert reviewer will mark these documents as relevant or not relevant to the discovery categories as a whole and will highlight if a document is to be withheld due to privilege or other withholding requirement[26].
    - **MODEL INITIALISATION** – The system will use the coding from the initial review to train the model which then predicts the relevance of uncoded documents in the dataset based on their closeness to coded examples. The system will rank the documents indicating the likelihood of their relevance.

      REVIEW AND CONTINUOUS LEARNING – Once the model has been built and the initial ranking has been assigned, the review team will then commence the review. The system will present the reviewers with the highest ranked documents along with a statistical sample of low and medium ranked documents. As each document is reviewed and coded, the system will learn from the coding and re-rank the remaining uncoded documents. This process will continue throughout the review until the cut-off point is reached. The cut-off point is the rank below which the technology has determined the documents are unlikely to be relevant.

      PROJECT VALIDATION – Once the cutoff rank has been reached then a statistical quality test to estimate if any relevant documents may have been missed by the system will be carried out. The aim of project validation will be to estimate the accuracy and completeness of the relevant document set assuming the review has concluded and any unreviewed documents below the cut-off point were not produced. This will involve a review by the expert reviewer of a random sample of documents that the system has predicted to be not relevant. [Number] of documents will be included in the sample based on [percentage] level of confidence and [percentage] margin of error. If few or no relevant documents are identified as a result of the Project Validation the review will be complete. If many relevant documents are identified the review will continue and the model will continue to be

[26] *It is essential that documents are marked for privilege and/or commercial sensitivity at this stage to avoid duplication. Note that CAL filters only for relevance, not for privilege or commercial sensitivity. Documents will only be withheld on the basis of commercial sensitivity if they are not relevant to the discovery categories.*

re-trained until the elusion rate (the number of relevant documents identified as a result of the Project Validation) is sufficiently low.

- **Second pass –** It is not possible to determine in advance what the cut-off point might be. Once the cut-off point has been decided, all documents with a relevance probability above the cut-off point, their families (or related documents) and other unique families which contain duplicates, will be manually reviewed. Documents will be considered in the context of their families and will also be considered for privilege and categorisation. Documents requiring redaction will be identified at this stage.

  To ensure consistency, daily review meetings will take place at which reviewers will flag and discuss query documents or those considered borderline for relevance or privilege with a supervising solicitor.

- **Redaction pass** – Documents requiring redactions will have redactions applied during this review pass.

G.9.1 Two-pass review without CAL [Model 2]

As many responsive documents form part of wider families of documents, a two-pass review will be required consisting of:

- **First pass –** Assessing whether a document is relevant and suppressing clearly irrelevant documents from further review. This review pass will consider documents in isolation and only one unique copy of each document responsive to the filtering criteria will be reviewed.
- **Second pass –** Documents identified as relevant through the first pass review, their families (or related documents) and other unique families which contain duplicates, will be included in this review pass and considered for privilege and categorisation. Documents requiring redaction will be identified at this stage.

  To ensure consistency, daily review meetings will take place at which reviewers will flag and discuss query documents or those considered borderline for relevance or privilege with a supervising solicitor.
- **Redaction pass –** Documents requiring redactions will have redactions applied during this review pass.

At each pass of the review, at least [percentage]% of review decisions by each reviewer will undergo quality checks by [quality check team].

G.9.1 Single-pass review without CAL [Model 3]

A single-pass review will be conducted consisting of:

- **First pass –** consideration of relevance and privilege status of all responsive documents, their families (or related documents) and other unique families which contain duplicates. Documents requiring redaction will also be identified at this stage.

  To ensure consistency, daily review meetings will take place at which reviewers will flag and discuss query documents or those considered borderline for relevance or privilege with a supervising solicitor.

- **Redaction pass –** Documents requiring redactions will have redactions applied during this review pass.

At each pass of the review, at least [percentage]% of review decisions by each reviewer will undergo quality checks by [quality check team].

At each pass of the review, at least [percentage]% of review decisions of each reviewer will undergo quality checks by [quality check team].

G.10 Phase Six - Analysis

The objective of the analysis phase is to take a deeper look at specific documents, for example, to determine their provenance. If necessary, [producing party] may perform a detailed analysis of a document or groups of documents.

G.11 Phase Seven - Production

At the completion of the review, relevant documents will be produced by generating an electronic schedule of the documents in the form of [a spreadsheet/electronic load file/other]. The information to be included in the schedule [and loadfile] can be seen in Attachment Four below. Documents identified as fully privileged will not be produced but will be scheduled separately in accordance with the Rules of the Superior Courts.

Documents will be produced in their native format by default. Exceptions to this include:

- [Hardcopy documents which have been scanned will be produced in PDF format.]
- Corrupt, password protected or encrypted documents may be converted to a different format (such as PDF) which enables their use.
- Redacted documents will be produced in [format, such as redacted PDF or TIFF] and will also be identified in the schedule as having being redacted.
- On rare occasions modified versions of native documents may be produced where it is not practical, possible, or proportionate to image them before redaction. Examples might include very large spreadsheets. Any such documents will be separately identified in the schedule.
- [Documents which require redaction but cannot reasonably be redacted such as very large spreadsheets or databases may be provided for inspection only.]

Document families, such as emails and their attachments, will be produced as follows:

- If a parent email is relevant but its children are not, then only the parent will be produced.
- If a child attachment is relevant, then its parent email will also be produced for context. Orphan child attachments will not be produced in isolation.
- If there are multiple child attachments, where only one is relevant, then only the relevant attachment and its parent email will be produced. Irrelevant attachments will not be produced.

[A slip sheet for irrelevant family members which have not been produced, will be provided.]

[Producing Party]'s production will comprise of the electronic files, [the electronic schedule/loadfile] [and extracted text]. They will be transferred through a secure online transfer system. Both the files and the schedule will be encrypted and the decryption password or key will be provided to [requesting party] separately.

G.12 Phase Eight - Presentation

Once a venue for the hearing of the matter has been finalised, [producing party] proposes that the following electronic system be utilised to facilitate the efficient management of documents throughout the hearing and to allow the documents to be shared with all parties during the hearing.

[Include detailed information as to what system is proposed, any external providers required, and any costs involved. Also include which documents will be presented in hardcopy (such as the core books), and those that will be presented electronically (such as everything else).]

G.13 Attachment One – Custodians and document sources

Custodian-based document sources[27]

No.	Custodian name	Live email	Archived/ backup email	Laptop documents	Private network folder	Cloud storage	Mobile Phone / Personal Devices
1	[Custodian One]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]
2	[Custodian Two]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]
3	[Custodian Three]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]	[Yes/No]

[27] *This should be amended to include all custodian document sources*

Non-custodian-based document sources

No.	Source name	Description
4	[Source name one]	[Documents from cloud storage/shared network/project folder.]
5	[Source name two]	[Hardcopy documents from centralised filing system.]

G.14 Attachment Two – Document type filters

The table immediately below details the document types which have been included:

No.	File type	File extension
1	[Microsoft Office]	[doc, docx, xls, xlsx, ppt, pptx, etc.]
2	[Portable Document Format]	[pdf]

G.15 Attachment Three – Other filters

The following filters have been applied to the document set.

G.15.1 Search terms used for the collection of documents

No.	Search term
1	[Search term one]
2	[Search term two]
3	[Search term three]

G.15.2 Date range

Emails and documents with a date sent, date created, modified, or last accessed, between [date/time] and [date/time] will be included.

G.15.3 Keywords used to filter for review

No.	Search term	Deduplicated hits
1	[Search term one]	[number]
2	[Search term two]	[number]
3	[Search term three]	[number]
n/a	All above combined*	[number]
n/a	Combined including families**	[number]

* This is the combined number of documents which are responsive to one or more of the search terms. i.e. if a document is responsive to two or more of the search terms, it is only necessary to review it once.

** This is the combined number above, plus their families, plus any other unique families which also contain the responsive document.

G.15.4 Other filters

[Any additional filters applied should be detailed here.]

G.16 Attachment Four – Production format

The following metadata fields will be included in the production schedule:

- Unique document identifier
- Unique family/parent identifier
- Whether the document is a parent or child in the context of any family relationship
- The date and time the document was last modified, or email sent in the format DD/MM/YYYY HH:MM:SS
- The document type
- The document author, or email sender
- The document recipient, or email Cc and Bcc
- The document name, or email subject
- The category the document falls into (where categories have been assigned)
- Whether the document is partially privileged
- The type of privilege that applies
- If the document was redacted and why

[The information provided in the schedule will be sorted by descending parent date/time. i.e. parents will be listed with their children next, and then the next family.]

[The following metadata fields will be included in the production loadfile where applicable:

- Begin Bates (unique document identifier)
- End Bates
- Begin Attachment
- End Attachment
- Author
- Email From
- Email To
- Email CC
- Email BCC
- Email Subject
- Record Type

- Sort Date/Time
- Sent Date/Time
- Last Modified Date/Time
- Created Date/Time
- File Name
- File Extension
- Category (where categories have been assigned)
- Privilege
- Privilege Type
- Redaction Reason
- Hash Value

Note: For electronic documents, the metadata will be derived automatically from the metadata contained within the electronic document. No manual process will be employed to verify or correct such metadata. For hardcopy documents which have been scanned, the metadata fields will be manually populated through the coding process.

Note: Each document will have a unique document identifier and a unique family identifier. These will be included in the relevant schedules and the underlying documents will be named by their unique identifier (the contents of the documents themselves will not be altered to include the identifier). Documents produced will not employ any form of Bates-stamping; rather the unique document identifiers will be used to uniquely identify each document in the production set [OR as the documents are being produced in [image/PDF] format[28] the documents produced will have the unique document identifier bates stamped onto the face of the document.]

[28] *An image might be produced for example in image or pdf format where producing in native or near native format would result in including documents in the production which are privileged or not relevant as they are embedded into the native email.*

Appendix H -Sample Review Plans

This appendix outlines detailed sample approaches to the review phase of the discovery process, presenting two main alternative methodologies (Models 1 and 2) based on common current practices. Whilst these represent the most prevalent review methodologies based on document review technology commonly in use at the time of publication of this Guide, this is not exhaustive or definitive and there may be further alternative approaches available and/or which may emerge as document review technology continues to develop and evolve and in particular as GenAI options are developed.

Selecting the Appropriate Review Model

The first consideration when choosing between the two approaches is whether Technology Assisted Review (TAR) is suitable for the dataset, in which case Model 1 will be the most appropriate starting point. The majority of discovery reviews are now conducted using some form of TAR, one of the most common being Continuous Active Learning ("**CAL**"). However, in certain instances, TAR may not be appropriate—for example, where the dataset is very small, where resources are limited or where the dataset is not appropriate for TAR—and in those scenarios[29], Model 2 (multi-pass manual review without the use of CAL) may be appropriate.

Relevant review statistics can only be determined at the end of the processing phase once filters have been applied and tested, and it is for this reason that the approach to review should not be decided until such information has been gathered and assessed in order to determine the most efficient approach to review.

Note – it may be acceptable to carry out a single-pass manual review of documents without the use of CAL. This would normally only arise where there are a very small number of documents to be reviewed or where documents relate to a discrete issue or category.

[29] *This is not an exhaustive list.*

[Matter reference]

Review Plan

[Version 0.1]

[Date]

H.1Background

This review plan outlines some detail of the possible approach to the review phase of the discovery process.[30]

For **Model 1**, the objective is to carry out a CAL review to identify potentially relevant documents and then manually review those documents in the second pass review, with each document having a determination made as to its relevance to the issues in the matter, and documents being coded for privilege, categorisation and redacted as necessary and appropriate. At the time of publication of this Guide, the majority of discovery reviews are carried out based on the Model 1 approach.

For **Model 2**, the objective is to perform a manual review of documents highlighted as potentially relevant through the filtering applied at the processing phase, completed on a document by document basis, with each document having a determination made as to its relevance to the issues in the matter, and documents being coded for privilege, categorisation and redacted as necessary and appropriate.

H.2 Review Team Structure and Responsibilities

A [Review manager or managers] will be appointed as the review manager(s) for this project. S/he/they will be responsible for all aspects of the review, including:

- Structuring the review, resourcing the review team, and planning the review, including documenting the review protocol
- Training the review team on which tagging/annotations to use and why, including how families of documents will be reviewed and tagged
- Forming and managing the Quality Check ('**QC**') team to perform quality checks on the review decisions made by the review team, and performing QC on the final production (the QC team is usually formed with one or more senior reviewers)
- Managing the assignment of documents for review and liaising with other service providers.

[30] *While the discovery protocol is intended as a means for information sharing and agreement between the parties, this review protocol is typically not shared between the parties, but rather between the various stakeholders (producing party, internal and external legal advisors, review team, and any external service providers) responsible for conducting the review.*

Additionally, an [eDiscovery service provider (which may be internal or external as appropriate)[31]] will be engaged and will be responsible for:

- Providing access to the review team from the law firm/producing party to the review platform
- Providing training on how to use the review platform and providing support on the use of the review platform and dealing with technical issues
- The final production of documents from the review platform
- [Any other roles which will be performed by the eDiscovery service provider should be specified]

H.3 Approach to Review

Model 1: Multi-Pass Review with CAL

The first step is deduplication and identification of documents not suitable to CAL (which are to be included in second-pass review – see below). Following this step, a CAL review is undertaken.

Step 1: CAL Review Process - The CAL review comprises four main steps:

1. Initial Review

The expert reviewer is presented with a small, representative sample of documents which are randomly selected (or selected using an agreed selection process) from the entire dataset (referred to as the 'Initial Assessment Set' or 'Seed Set'). The expert reviewer marks these documents as Relevant or Not Relevant to the discovery categories as a whole and highlights if a document is to be withheld due to privilege or other withholding requirement.

2. Model Initialisation

The system uses the coding from the initial review to train the model which then predicts the relevance of uncoded documents in the dataset based on their closeness to coded examples. The system ranks the documents with a score of between 0 and 100 indicating the likelihood of their relevance, with a minimum of 5 documents coded 'Relevant' and 5 documents coded 'Not Relevant' required to train the model.

3. Review and Continuous Learning (First Pass Review)

Once the model has been built and the initial ranking has been assigned, the review team commences the review, with the system presenting the reviewers with the highest ranked documents along with a statistical sample of low and medium ranked documents. As each document is reviewed and coded, the system learns from the coding and re-ranks the remaining uncoded documents. This process continues throughout the review until the cut-off

[31] *An eDiscovery service provider may be internal where the relevant law firm has an internal specialist eDiscovery technical team.*

point is reached—the rank below which the technology has determined the documents are unlikely to be relevant. All documents above the cut-off point are manually reviewed for relevance and those below the cut-off point are subject to statistical sampling, but otherwise not reviewed.

Once the review has reached a stage where the reviewers are, for a sustained period of time, no longer being presented with relevant documents by the system, the review is stopped as the likely existence of relevant documents remaining in the uncoded dataset will be statistically very low, indicating the cut-off rank has been reached.

4. Project Validation

Once the cut-off rank has been reached, a statistical quality test is carried out to estimate if any relevant documents may have been missed by the system. The aim of project validation is to estimate the accuracy and completeness of the relevant document set assuming the review has concluded and any unreviewed documents below the cut-off point will not be produced. This involves a review by the expert reviewer of a random sample of documents that the system has predicted to be not relevant, with the number of documents in the sample based on a specified level of confidence and margin of error.

If few or no relevant documents are identified as a result of the Project Validation, the review is complete; if many relevant documents are identified, the review continues and the model continues to be re-trained until the elusion rate (the number of relevant documents identified as a result of the Project Validation) is sufficiently low.

Step 2: Second-Pass (Manual Review)

It is not possible to determine in advance what the cut-off point might be, but once the cut-off point has been decided, all documents which have a relevancy rank above the cut-off point, their parent emails are included in this manual review pass. Documents are considered in the context of their families and are also considered for privilege, categorisation and redaction as appropriate and necessary. Documents which were not suitable for CAL are also included in the second-pass review.

The second-pass review should commence shortly after the CAL review has started and should run simultaneously to the CAL review so that any errors in coding throughout the CAL review are identified and corrected.

Document Assignment Methods - Documents for review are either:

(1) **Batch Reviewed:** documents are split into batches which are checked out by reviewers. Reviewers will review the batch before checking it back in as complete. The batches will contain all relevant documents, email threads and attachments, **or**

(2) **Queue Reviewed**: documents are placed in a second pass review queue which serves families of documents to the reviewers ensuring that the same documents are not reviewed by different reviewers at the same time.

Coding Layout

The reviewer is presented with a coding layout which will likely contain, at a minimum, the following choices (or equivalent):

Tag	Choices	Mandatory
Relevance (Single choice)	Relevant Not Relevant Query Technical Issue	Yes
Category (Multiple choice)	Category 1 Category 2 Category 3	Yes, but only if Relevant
Privilege, Data Protection and Commercial Sensitivity (Multiple choice)	Privileged – Legal Advice (withhold) Privileged – Litigation (withhold) Part-privileged – Legal advice(for redaction) Part-privileged – Litigation (for redaction) Data protection (for redaction)* Commercial sensitivity (for redaction)+	No
Comments (Free text field)	n/a	No
Discuss at Review Meeting (Single choice)	Yes Resolved	No
Quality Control (Multiple choice)	QC Complete QC Changes made	Only available to QC team

*This refers to documents which contain personal data which is not relevant, the disclosure of which may interfere with data subject rights.

+This refers to documents which contain commercially sensitive information which is not relevant to the dispute.

Step 3: Quality Control

Quality control can be implemented using either a **batch-based** or **review queue-based** approach:

Batch-Based QC Approach

When a reviewer completes a batch, the review manager assigns it to the QC team who will:

1. Review all 'Query' documents and re-classify them as Relevant, Not Relevant, or Technical Issue
2. Randomly sample a specified percentage of the batch to verify tagging accuracy
3. Tag documents as 'QC Complete' (adding 'QC Changes made' if corrections were required)
4. Return the entire batch for re-review if errors exceed the [percentage]% threshold, providing additional guidance to the reviewer.

Review Queue-Based QC Approach

As documents are coded, they automatically flow into QC queues:

1. 'Query' documents → QC query queue for re-classification by the QC team

2. Random sample of 'Relevant' documents → QC relevant queue for verification
3. Random sample of 'Not Relevant' documents → QC not relevant queue for verification
4. Documents exceeding the [percentage]% error threshold are returned to the review queue with additional guidance provided to the reviewer.

Note - the key difference between batch-based QC and review queue-based QC is that with batch-based QC, the QC team reviews entire batches after completion, whilst with review queue-based QC, the QC team reviews documents continuously as they're coded, with documents automatically sorted into different QC queues as the second-pass review progresses. With the increasing use of technology including the use of GenAI, the review queue method is likely to become more common.

Technical Issue Resolution

Throughout the review process, the eDiscovery service provider regularly assesses documents coded as 'Technical Issue'. Once resolved, these documents are tagged as 'Query' and either:

- Directed to the review manager for final tagging as Relevant or Not Relevant, or
- Placed in the QC Query queue for the QC team to finalise

Step 4: Redaction-Pass Review

This pass includes all documents requiring redaction: Part-privileged (Legal advice or Litigation or both), Data Protection, or Commercial Sensitivity. Again, there are two different types of approaches:

Batch-Based Approach

Documents are organised into batches by type:

- **Emails And Documents** – emails and standard documents
- **Spreadsheets** – spreadsheets (typically requiring native redaction)
- **Other** – other formats (typically requiring native redaction)

Batches are sorted by email thread (for emails) or near duplicate (for other documents) to ensure consistency.

Review Queue-Based Approach:

Two separate queues are used:

- **Queue 1**: Emails and standard documents
- **Queue 2**: Spreadsheets and other document types (typically requiring native redaction)

Both queues are sorted by email thread or near duplicate to maintain consistency across similar materials.

Redaction Coding Layout

The reviewer will be presented with a coding layout which will contain the following choices:

Tag	Choices	Mandatory
Redaction (Single choice)	Complete No longer required Technical Issue	Yes
Comments (Free text field)	n/a	No
Discuss at Review Meeting (Single choice)	Yes Resolved	No
Quality Control (Multiple choice)	QC Complete QC Changes made	Only available to QC team

Redaction QC Process

Batch-Based Approach:

Upon completion, the reviewer notifies the review manager, who assigns the batch to the QC team. The QC team:

1. Randomly samples a specified percentage of documents
2. Reviews redactions and tags documents as 'QC Complete' (and 'QC Changes made' if amendments were made)
3. Returns batches exceeding the [percentage]% error threshold for re-review with additional guidance.

Review Queue-Based Approach:

Redacted documents automatically enter a QC review queue where:

1. A random sample is reviewed by the QC team
2. Documents are tagged as 'QC Complete' (and 'QC Changes made' if necessary)
3. Documents exceeding the [percentage]% error threshold are re-entered into the redaction queue with additional guidance provided.

Technical Issue Resolution

Throughout and at the conclusion of the redaction pass, the eDiscovery service provider addresses 'Technical Issue' documents, directing the review manager on how to access and redact them once resolved.

Model 2: Multi-Pass Review Without CAL

As many of the documents which are responsive to the filtering criteria form part of wider families of documents, a multi-pass review is required.

Step1: First-Pass Review

The first-pass identifies whether a document is relevant to the discovery categories and suppresses clearly irrelevant documents from further review. This review pass considers documents in isolation and only one copy of each document responsive to the filtering criteria is reviewed.

Document Assignment Methods:

Documents for review are either:

(1) **Batch Reviewed:** documents are split into batches which are checked out by reviewers. Reviewers will review the batch before checking it back in as complete, **or**

(2) **Queue Reviewed**: documents are placed in a first-pass review queue which one copy of each documents to the reviewers ensuring that the same documents are not reviewed by different reviewers at the same time.

First-Pass Coding Layout

The reviewer is presented with a coding layout which contains the following choices:

Tag	Choices	Mandatory
Relevance (Single choice)	Relevant Not Relevant Query Technical Issue	Yes
Comments (Free text field)	n/a	No
Discuss at Review Meeting (Single choice)	Yes Resolved	No
Quality Control (Multiple choice)	QC Complete QC Changes made	Only available to QC team

Step 2: Quality Control (First-Pass)

First-Pass QC Process

Quality control can be implemented using either a **batch-based** or **review queue-based** approach:

Batch-Based QC Approach

When a reviewer completes a batch, the review manager assigns it to the QC team who:

1. Review all 'Query' documents and re-classify them as Relevant, Not Relevant, or Technical Issue
2. Randomly sample a specified percentage of the batch to verify tagging accuracy
3. Tag documents as 'QC Complete' (adding 'QC Changes made' if corrections were required)
4. Return the entire batch for re-review if errors exceed the [percentage]% threshold, providing additional guidance to the reviewer.

Review Queue-Based QC Approach

As documents are coded, they automatically flow into QC queues:

1. 'Query' documents → QC query queue for re-classification by the QC team
2. Random sample of 'Relevant' documents → QC relevant queue for verification
3. Random sample of 'Not Relevant' documents → QC not relevant queue for verification
4. Documents exceeding the [percentage]% error threshold are returned to the review queue with additional guidance provided to the reviewer.

Technical Issue Resolution

Throughout the review, the eDiscovery service provider addresses 'Technical Issue' documents. Once resolved, these are marked as 'Query' and either:

- Directed to the review manager for final tagging as Relevant or Not Relevant, or
- Placed in the QC Query queue for the QC team to finalise

Step 3: Second-Pass Review

This pass reviews documents tagged 'Relevant' in the first pass, along with their families. Documents are assessed in the context of their families for privilege, categorisation, and redaction requirements.

Document Assignment Method

Again, documents for review are either:

(1) **Batch Reviewed:** documents are split into batches (with families being kept intact) which are checked out by reviewers. Reviewers will review the batch before checking it back in as complete, **or**

(2) **Queue Reviewed**: documents are placed in a second-pass review queue that serves complete document families to reviewers, preventing duplicate review of the same documents.

Coding Layout

The reviewer is presented with a coding layout which will likely contain, at a minimum, the following choices (or equivalent):

Tag	Choices	Mandatory
Relevance (Single choice)	Relevant Not Relevant Query Technical Issue	Yes
Category (Multiple choice)	Category 1 Category 2 Category 3	Yes, but only if Relevant
Privilege, Data Protection and Commercial Sensitivity (Multiple choice)	Privileged – Legal Advice (withhold) Privileged – Litigation (withhold) Part-privileged – Legal advice (for redaction) Part-privileged – Litigation (for redaction) Data protection (for redaction)* Commercially sensitive (for redaction)+	No
Comments (Free text field)	n/a	No
Discuss at Review Meeting (Single choice)	Yes Resolved	No
Quality Control (Multiple choice)	QC Complete QC Changes made	Only available to QC team

*This refers to documents which contain personal data which is not relevant to the matter, the disclosure of which may interfere with data subject rights.

+This refers to documents which contain commercially sensitive information which is not relevant to the dispute.

Step 4: Quality Control (Second-Pass)

Quality control can be implemented using either a **batch-based** or **review queue-based** approach:

Batch-Based QC Approach

When a reviewer completes a batch, the review manager assigns it to the QC team who:

1. Review all 'Query' tagged documents and re-classify them as Relevant, Not Relevant, or Technical Issue
2. Randomly sample a specified percentage of the batch to verify tagging accuracy
3. Mark documents as 'QC Complete' (adding 'QC Changes made' if corrections were required)
4. Return the entire batch for re-review if errors exceed the [percentage]% threshold, providing additional guidance to the reviewer

Review Queue-Based QC Approach

Documents automatically flow into separate QC review queues as they are coded:

1. 'Query' documents → QC query queue for re-classification by the QC team
2. Random sample of 'Relevant' documents → QC relevant queue for verification
3. Random sample of 'Not Relevant' documents → QC not relevant queue for verification
4. Documents exceeding the [percentage]% error threshold are returned to the review queue with additional guidance provided to the reviewer.

Technical Issue Resolution

Please refer to page [ ]. It will be the same process as set out in Model 1.

Redaction-Pass Review

Please refer to page [ ]. It will be the same process as set out in Model 1.

Important Notes

- Separate Redaction Pass: Redactions should be performed in a separate pass due to the technical imaging process and comprehensive QC requirements. Only expert users should 'image and redact on-the-fly'. Consistent redactions across document copies are essential to avoid inadvertent waiver of privilege.
- Skipping First-Pass: Documents not subject to filtering criteria (such as scanned hardcopy documents or dedicated project folders) can proceed directly to second-pass review, bypassing the first pass.

H.4 Production Criteria

At the conclusion of the redaction-pass, documents coded 'Relevant' and not fully privileged will be produced.

As outlined in the discovery protocol, documents will be produced in their native format by default. Exceptions to this include:

- [Hardcopy documents which have been scanned will be produced in PDF format.]
- Corrupt, password protected or encrypted documents may be converted to a different format (such as PDF) which enables their use.
- Redacted documents will be produced in [format, such as redacted PDF or TIFF] and will also be identified in the schedule as having being redacted.

- In rare occasions, modified versions of native documents may be produced. This may be the case where it is neither practical, possible, or proportionate to image before redaction. Examples might include very large spreadsheets. Any such documents will be separately identified in the schedule.
- [Documents which require redactions, but cannot reasonably be redacted, such as very large spreadsheets or databases, may be provided for by inspection only.]

Document families, such as emails and their attachments, will be produced as follows:

- If a parent email is relevant, but its children are not, then only the parent will be produced.
- If a child attachment is relevant, then its parent email will be produced for context. Orphan child attachments will not be produced in isolation.
- If there are multiple child attachments but only one is relevant, then only the relevant attachment and parent email will be produced. Irrelevant attachments will not be produced.

Note: It is important that families of documents are coded according to the agreed criteria during the review. This will help ensure that effort is not expended late in the process correcting the consistency of family markings.

The following QC steps will be completed prior to production by the eDiscovery service provider:

- Verify that documents containing redacted data have been coded for production in image format, if required.
- Verify that families of documents have been tagged appropriately from a relevancy and privilege perspective.
- Verify that: (i) documents coded for redaction are redacted; (ii) no documents have been redacted which were not coded for redaction; and (iii) documents have been checked for redaction consistency.
- Verify that categories and other tagging have been applied to documents where required.
- Verify that there are no containers within the production set which could result in an embedded item being inadvertently disclosed.
-

H.5 Review Phase Timelines

The current deadline for production is [date/time]. In order to complete QC and production and make the necessary copies of the production (documents and schedules) available, the review phase would need to be complete by [date/time].

The following provisional timelines have been agreed for the review phase:

- First-pass to commence on [date/time] and finish by [date/time].
- Second-pass to commence on [date/time] and finish by [date/time].
- Redaction-pass to commence on [date/time] and finish by [date/time].

It is very difficult to estimate in advance the number of documents which will require second-pass review and/or redactions, therefore the times outlines are indicative only.

Appendix I Sample request for voluntary discovery

[From solicitor for requesting party]

[To solicitor for responding party]

[Date]

[Matter reference]

Dear Sirs,

We refer to the above matter and to previous correspondence in relation to these proceedings. This letter constitutes our formal request pursuant to the Rules of the Superior Courts seeking voluntary discovery from the [Plaintiff/Defendant].

TAKE NOTICE that, in accordance with the terms of Order 31, Rule 12 of the Rules of the Superior Courts (as amended), the [Plaintiffs/Defendants] hereby requests the [Plaintiff/Defendant] to make voluntary discovery of all documents which are or have been within its possession, power, or procurement, and which are within the categories of documents listed below. References in this letter to a "document" or "documents" are to all documents in the possession, power or procurement of the [Plaintiff/ Defendant], their servants, agents or any other person from whom they have a legally enforceable right to procure said documents.

The term "documents" comprises all media on which information of any kind is stored and includes electronically stored information, metadata and hard copy documents including but not limited to e-mails, text messages, WhatsApp and other messaging formats, Word or similar files, spreadsheets, databases, computer generated files or communications, notes, agreements, memoranda, correspondence, financial records, reports, minutes, calculations, transcripts, sound files or tape recordings, or communications and records of any kind containing relevant information however maintained, whether electronically, in hard copy format or in any other manner.

Please note that the "Agent" of a particular party includes all representatives including all expert advisers and all other persons acting on behalf of that party.

Category 1

[Describe in detail the category of document being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Reasons

[Describe in detail the reasons for the category being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Category 2

[Describe in detail the category of document being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Reasons

[Describe in detail the reasons for the category being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Category 3

[Describe in detail the category of document being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Reasons

[Describe in detail the reasons for the category being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Category 4

[Describe in detail the category of document being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Reasons

[Describe in detail the reasons for the category being requested by reference to specific pleas to demonstrate relevance to the issues in the pleadings.]

Other

[Insert any definitions, as required to provide clarity to the categories and reasons.]

And **TAKE NOTICE** that:

1. Voluntary discovery is requested pursuant to Order 31, Rule 12 (4) RSC.
2. Any agreement to make discovery would require it to be made on oath in a manner and form and will have such effect as if directed by order of the Court.
3. Discovery is required to be made with the documents listed in a manner which allows the categories which they respond to be clearly identified.

4. Where documents of which discovery is sought exist in electronic format, production of the same in searchable form is requested. [Plaintiff/Defendant] reserves its position as to whether it will be necessary to seek the provision of inspection and searching facilities using any information and communications technology system owned or operated by the [Plaintiff/Defendant].
5. Objection will be taken to any attempt to adduce in evidence a document which has not been discovered.

In circumstances where the [Plaintiff/Defendant] confirms that they will make voluntary discovery, we require discovery to be made by affidavit sworn by them and furnished to us, together with copies of all documentation, within a period of [14] weeks from the date hereof.

Kindly note that in circumstances where [Plaintiff/Defendant] do not confirm that they will make voluntary discovery of all documentation referred to above, or if such confirmation is not received within a period of [2] weeks hereof, we will have no option but to issue a motion seeking Orders directing the [Plaintiff/ Defendant] to make discovery of the categories of documents identified above without further notice to you. Furthermore the content of this letter will be used to seek to fix your client with the costs of any application necessitated by reason of your clients' failure to make discovery as requested.

Yours faithfully,

[Solicitor for Plaintiff/Defendant]

Appendix J Sample affidavit of discovery

THE HIGH COURT

[Commercial]

Record No. [INSERT YEAR] [INSERT NO.] [P/S]

BETWEEN

:

[INSERT PARTY[ies]

Plaintiff[s]

-and-

[INSERT PARTY [ies]

Defendant[s]

DRAFT / [SUPPLEMENTAL][32] AFFIDAVIT OF DISCOVERY[33]

I, **[INSERT NAME]**, [**INSERT PROFESSION**], of [**INSERT ADDRESS**], aged eighteen years and upwards, **MAKE OATH AND SAY** as follows:

1. I am the [**INSERT DETAILS OF DEPONENT**] of the [**Plaintiff/Defendant**] herein and I Make this Affidavit of Discovery on its behalf and with its authority and consent. I do so from facts within my knowledge save where otherwise appears and whereso otherwise appearing I believe same to be true and accurate.

2. [I make this Affidavit of Discovery pursuant to an agreement made on behalf of the **[Plaintiff/ Defendant]** to make discovery in terms recorded in correspondence dated **[INSERT DATES AND DETAILS OF RELEVANT CORRESPONDENCE]**] **OR [**I make this Affidavit of Discovery pursuant to an Order of this Honourable Court (the Hon. Mr./ Ms. Justice [**INSERT NAME OF JUDGE]** dated the **[INSERT DATE OF ORDER]**].

---

[32] *Where Affidavit is to be an Affidavit Supplemental to an Original Affidavit of Discovery then wording in the following terms should be inserted into the Supplemental Affidavit of Discovery: "This Affidavit is supplemental to my Affidavit of Discovery sworn on [**INSERT DATE**] in these proceedings (the Original Affidavit of Discovery"). If required to provide reasons as to why a Supplemental Affidavit is being sworn then the following sample text may be of assistance: "This Supplemental Affidavit of Discovery is sworn because further documents have been located which ought to have been referred to in the Original Affidavit of Discovery] but were not because **[insert reasons why the documents were not included in the Original Affidavit of Discovery]**"*

[33] *Where a party is ordered to make discovery, Order 31, Rule 13 RSC requires that this is done on affidavit made out in the format required by Form 10, Appendix C RSC. The proper title of this Affidavit, as set out in Form 10, Appendix C is "Affidavit as to Documents", although in practice is almost universally referred to as an "Affidavit of Discovery".*

Categories of Documents to be Discovered

3. I am advised that the categories of documents in respect of which the **[Plaintiff** / **Defendant**] has [agreed to make discovery (the "**Agreed Discovery**")] **OR** [been ordered to make discovery (the "**Ordered Discovery**")] are as follows:

[INSERT FULL DETAILS OF ALL CATEGORIES OF DISCOVERY HERE]

Methodology

4. In order to comply with its discovery obligations, the [**INSERT NAME OF PARTY**], as advised by its solicitors, has conducted a wide-ranging and extensive search for documentation which is within the scope of [the Agreed Discovery **OR** the Ordered Discovery]. In particular …

[This section should outline the steps that were taken to ensure that the party making discovery has complied with its discovery obligations. The Affidavit should address the stages of the discovery process as outlined in this Guide, including Identification of the relevant document universe (see Chapter 5); Preservation (see Chapter 7); Collection (see Chapter 8); Processing (see Chapter 9); Review (see Chapter 13); and Production (see Chapter 14).]

[While the content of the Affidavit will depend on the circumstances of the case and the particular discovery process, topics to be addressed should include:

- Whether a third-party discovery service provider was retained.
- Procedures applied for de-duplicating and approach to Family/ Parent and Child Documents should be discussed.
- Whether Computer Assisted Learning or similar process was applied for reviewing and deciding upon potentially relevant documents;
- Approach to redactions should be discussed, specifying whether redactions have been applied to portions of documents which contain information that is privileged and or commercially sensitive and or confidential and or personal data that is not relevant to the proceedings.]

Categorisation of Documents

5. The [**INSERT NAME OF PARTY**] in making discovery has for ease of reference listed each document being discovered under one of the [**INSERT NUMBER**] categories of documents within the [Agreed Discovery **OR** Ordered Discovery]. It is the case, however, that there is an overlap between various categories in the [Agreed Discovery **OR** Ordered Discovery] in that many of the documents being discovered might reasonably be considered to fall within the terms of a number of categories. The [**INSERT NAME OF PARTY**] has been advised that it is not obliged, and could not reasonably be expected, to identify every category under which a particular document might be listed. Accordingly, while the documentation listed in the First Schedule under the [**INSERT NUMBER**] categories comprises the totality

of documentation which the [**INSERT NAME OF PARTY**] is discovering pursuant to the [Agreed Discovery **OR** Ordered Discovery], the [**INSERT NAME OF PARTY**] does not thereby suggest nor is it the case that all the documents listed under any particular category are the only documents being discovered which are within the scope of that category.

Averments as to Discovery:

6. The [**INSERT NAME OF PARTY**] has in its possession, power or procurement the documents[34] [and electronically stored information][35] relating to the matters in question in this suit and falling within [the Agreed Discovery **OR** the Ordered Discovery][36] as set forth in the First and Second parts of the First Schedule hereto.

7. The [**INSERT NAME OF PARTY**] objects to producing the documents [and electronically stored information] set forth under [**SPECIFY PARAGRAPH NO.**] in the Second Part of the First Schedule hereto on the grounds that they are privileged and that they comprise communications of a confidential nature passing between the [**INSERT NAME OF PARTY**] and its legal advisers for the purposes of obtaining legal advice for or giving legal advice to the [**INSERT NAME OF PARTY**]. The [**INSERT NAME OF PARTY**] objects to producing the documents set forth under [**SPECIFY PARAGRAPH NO.**] in the Second Part of the First Schedule on the grounds that they are privileged in that they comprise documents that came into existence after these proceedings were contemplated or commenced and were created with a view to defending such proceedings either for the purposes of giving or obtaining advice in relation to them or of obtaining and collecting evidence to be used or of obtaining information which might lead to the obtaining of such evidence or for the purposes of defending these proceedings.

8. The [**INSERT NAME OF PARTY**] has had, but does not now have, in its possession, power or procurement the documents [and electronically stored information] relating to the matters in question in this suit that are set forth in the Second Schedule hereto[37].

9. The last mentioned documents [and electronically stored information] were last in my possession, power or procurement on [**INSERT DATE**].

10. That [here state what has become of the last-mentioned documents or information, and in whose possession they now are].

---

[34] Documents of the same or a similar nature, when numerous, must so far as possible, be grouped together and numbered or otherwise sufficiently marked so as to be identifiable.

[35] The Rules Amend Order 31 Rule 12 of the Rules of the Superior Courts and in particular make provision for the discovery of electronically stored information.

[36] Parties providing discovery shall list documents or categories of information, and shall provide documents and information for inspection, in a manner corresponding with the categories in the agreement or order for discovery, or in a sequence corresponding with the manner in which the documents or information have been stored or kept in the usual course of business by the party making discovery.

*[37] Additional text that might be considered here, where the context permits is: "In addition to the documents set out at [SPECIFY PARAGRAPH] in the Second Schedule, it is possible, having regard to the scope of the [Agreed Discovery **OR** Ordered Discovery] and the time period to which it relates, that some documents as described at [SPECIFY PARAGRAPH] in the Second Schedule, within the scope of the [Agreed Discovery **OR** Ordered Discovery] have not been retained or may have been overwritten in the ordinary course of business of the [INSERT PARTY] prior to the commencement of these proceedings."*

11. According to the best of my knowledge, information, and belief, [**INSET NAME OF PARTY**] has not now, and never had in its possession, power or procurement or in the possession, custody or power of its solicitors or agents, or in the possession, custody or power of any other persons, or person on its behalf, any document of any kind or any electronically stored information, or any copy of or extract from any such document or information relating to the matters in question in this suit or any of them, or wherein any entry has been made relative to such matters, or any of them and falling within the relevant categories of documents specified in the [Agreed Discovery **OR** Ordered Discovery] other than and except the documents [and electronically stored information] set forth in the said First and Second Schedules hereto.

SWORN by the said [NAME OF DEPONENT] on the [DATE] day of [DATE] 20XX at [ADDRESS], before me a Commissioner for Oaths / Practising Solicitor and [I know the Deponent] / [the Deponent has been identified to me by [NAME] who is personally known to me] / [prior to the swearing of this affidavit, the identity of the deponent has been established by me by reference to [insert particulars of photographic ID e.g. a passport (passport no. [*Passport number*] issued on [*date of issue*].

_____ _____

[NAME OF DEPONENT] Commissioner for Oaths/Practising Solicitor

This Affidavit was filed by [Law firm name], Solicitors for the [**INSERT PARTY**], [Law firm address], on the [**INSERT DAY**] day of **[INSERT DATE]**

FIRST SCHEDULE

First Part

[INSERT INDEX OF DISCOVERABLE DOCUMENTATION]

Draft Schedule

Production Number	Family ID	Doc ID	Date	Doc Type	Author	Recipient	Cc	Bcc	Name /Subject	File Extension	Category	Redaction

FIRST SCHEDULE

Second Part

1. [Insert Schedule of Privileged Documentation in relation to same]

Production Number	Family ID	Doc ID	Date	Doc Type	Author	Recipient	Cc	Bcc	Name /Subject	File Extension	Category	Privilege[38]

(a) All documents including correspondence, notes and memoranda passing between the [**INSERT PARTY**] and [Legal advisors] seeking advice in relation to the various aspects of the matters the subject of the proceedings herein.

(b) All correspondence with and advices, draft pleadings and opinions of Counsel in relation to the various matters the subject of the proceedings herein.

(c) All correspondence and advices received from expert witnesses retained on behalf of the [**INSERT PARTY**].

(d) All documents including various memoranda, notes of meetings, correspondence, reports and drafts thereof produced by the [**INSERT PARTY**], [Legal advisors], Counsel [and experts] for the purposes of these proceedings.

[38] *Depending on the context, it may be appropriate to include separate "Privilege Narrative" column. See discussion at Chapter* ***12,*** *page* ***64.***

SECOND SCHEDULE

Appendix K Overview of legal privilege

Whether a particular document or category of documents might be considered privileged under Irish law requires an examination of its content and/or the context of the document and its creation.[39] Ultimately, only the Courts have jurisdiction to decide whether a claim of privilege is justified.

Solicitor's Duties

Privilege belongs to the client, not the solicitor. That being so, a solicitor cannot waive privilege without client authority. A solicitor must assert privilege in respect of a document that appears to be privileged, unless the client expressly waives privilege.

Categories of Privilege

The main categories of privilege are:

1. Legal professional privilege;
2. Without prejudice privilege;
3. Common interest/Joint Interest privilege;
4. Public interest privilege;
5. Journalistic privilege; and,
6. Privilege against self-incrimination.

These are considered below, in turn.

K.1 Legal professional privilege

Legal professional privilege includes two distinct categories: legal advice privilege and litigation privilege. In both cases, the onus is placed upon the person invoking legal professional privilege to justify it.

[39] *A comprehensive review of legal privilege is beyond the scope of this Guide. For a more detailed treatment of the topic, see, Abrahamson, Dwyer and Fitzpatrick; Discovery and Disclosure, (3rd Ed., 2019), Chapters 39 – 45; Heffernan Legal Professional Privilege (2011).*

K.1(i) Legal advice privilege

a) In order to attract **legal advice privilege**, the material in question must satisfy a number of criteria.

(i) First, the material must constitute or refer to a communication between lawyer and client;

(ii) Secondly, that communication must arise in the course of the professional lawyer–client relationship;

(iii) Thirdly, the communication must be confidential in nature; and,

(iv) Fourthly, it must be for the purpose of giving or receiving legal advice. Legal advice is generally advice dealing with legal rights, obligations and remedies.

b) By contrast, legal advice privilege does not extend to seeking or providing **legal assistance** that is administrative, transactional, or business-related in nature, even when provided by a lawyer.

c) **Continuum of communications:** Where a transaction involves protracted dealings, the continuum as a whole may attract privilege if the communications are part of seeking/giving legal advice (e.g. client providing facts to enable lawyer to advise on legal position).

Examples

The distinction between legal advice and legal assistance can be difficult to apply in practice. The following are illustrative examples only in the context of legal advice privilege (a separate assessment would need to be undertaken as to whether they benefit from litigation privilege, discussed below):

Likely privileged:

- Advice on legal compliance with statutory obligations
- Opinion on legal risks in proposed transactions
- Analysis of contractual rights and remedies

Likely not privileged:

- Purely factual information gathering (absent request for or provision of legal advice);
- Administrative coordination of transactions
- Commercial negotiation strategy (absent legal advice element)
- Routine conveyancing steps without legal advice content
- Emails arranging meetings, passing on documents, or coordinating activities.

Part-privileged documents and redactions

Even within a single document, courts will distinguish privileged legal advice (requiring redaction) from non-privileged legal assistance. When making discovery of documents over which legal advice privilege is asserted, where only a portion of the document contains legal advice, this should be redacted (on agreement between the parties) and the document discovered as part privileged unless the parties expressly agree otherwise.

Families of documents

It is important to note that each individual document in a document family must be assessed for the purposes of being considered whether it is privileged. Where one document in a family is privileged, it is not necessarily the case that all documents in that same family will also be privileged.

A claim of litigation privilege should not be maintained over attachments to an email where those attachments are not of themselves privileged and the relevant attachment is not elsewhere disclosed in the First Schedule of the First Part of the affidavit of discovery (unless disclosing that attachment would undermine the privilege attaching to the email).

In house lawyers

Under Irish law, in-house lawyers and their employers are entitled to the same legal professional privilege as applies to external lawyers. The definition of 'lawyer' for this purpose includes solicitors, barristers, salaried in-house legal advisers, foreign lawyers and the Attorney General. However, practitioners should note:

a) *Evidence:* If challenged, cogent evidence is required as to the professional relationship between the lawyer and client.

b) *Non-practising solicitor / no practising certificate:* Legal Advice Privilege will not extend to advice provided by a solicitor acting outside of a professional relationship and without a practising certificate.

c) *Qualification requirements*: Advice given by an adviser who is not a qualified lawyer (or a trainee supervised by a qualified lawyer) does not attract legal advice privilege.

d) *EU law*: under EU law, legal professional privilege does not extend to communications between in-house lawyers and their clients in the context of competition law investigations by the European Commission.

e) *Dual roles*: Where an in-house lawyer performs both legal and commercial/managerial functions, careful consideration must be given to whether communications were made in

their capacity as legal adviser (potentially privileged) or principal/business executive (not privileged).

Third Party Communications

Unlike litigation privilege (below), legal advice privilege does not protect communications between the client or lawyer and a third party such as a witness or an expert. Exceptionally, legal advice privilege can extend to third parties but only where the third party is an agent for the purpose of communicating with the other party to give or obtain legal advice (e.g. an interpreter), not just an agent in the general sense.

K.1.(ii) Litigation Privilege

a) This is the privilege that exists over confidential documents created because of an apprehension or contemplation of litigation or for the dominant purpose of prosecuting or defending litigation. This may include proceedings before a tribunal and/or regulatory or criminal proceedings.

b) Litigation privilege also exists over documents which come into existence after the commencement of litigation and for the dominant purpose of the litigation.

c) The test applied by the Courts in assessing litigation privilege is known as the "dominant purpose" test. The Court must be satisfied that the primary or dominant purpose of the client in creating the specific document was litigation, either (i) actual or pending, or (ii) reasonably in contemplation or apprehended at the time of creation. "Reasonable contemplation" requires more than a vague possibility; there must be a real prospect of litigation. If there are other equally important purposes and litigation is not the dominant purpose, the document will not attract litigation privilege.

d) Litigation privilege may be claimed over communications between clients/lawyers and third parties if they are created for the dominant purpose of the litigation.

e) The burden of proof rests on the party claiming privilege to establish dominant purpose. –

Part Privilege

It is possible that part of a communication may be privileged notwithstanding that the document itself is not privileged. An example may be board meeting minutes, a section of which refers to litigation strategy. Privilege may be claimed over that part of the document which contains or refers to privileged information. In these circumstances, and on agreement between the parties, the privileged information may be redacted and the document disclosed as part privileged.

K.2 Without prejudice privilege

Communications by parties to a dispute which are written or made for the purpose of settling that dispute and which are either expressed to be or are otherwise proved to have been made on a "without prejudice" basis are privileged. The purpose of the communication must be to try to settle the dispute/proceedings. This is a joint privilege; it cannot be waived unilaterally by one party.

For a claim of without prejudice privilege to succeed the party claiming it must establish that the communication in question was made:

a) In a *bona fide* attempt to settle a dispute between the parties, and

b) With the intention that, if the negotiations failed, the communication could not be disclosed without the consent of the parties.

The use of the words "without prejudice" are not sufficient in themselves to invoke privilege.

Separately, the Mediation Act 2017 provides statutory protection for mediation communications and all records and notes relating to a mediation.

K.3 Common interest privilege / Joint privilege

Common interest privilege preserves privilege in documents that are disclosed to third parties where a person voluntarily discloses a privileged document to a third party who has a common interest in the subject matter of the privileged document or in litigation in connection with which the document was brought into existence (e.g. a co-defendant). The common interest must exist at the time of the disclosure and it applies to both legal advice privilege and litigation privilege. The common interest must be legal in nature (relating to litigation or legal advice), not merely commercial. A shared commercial interest is insufficient.

Examples of relationships which are capable of giving rise to or supporting necessary common interest include:

- an insurer and the insured;
- a reinsurer and reinsured;
- a principal and agent;
- companies within a corporate group;
- joint venture partners; and,

- co-defendants in litigation.

In examining whether a document is the subject of common interest privilege it is important to consider: Whether the document would, in the hands of a single party, have had the benefit of privilege in the first place. If not, then no question of common interest privilege can arise.

If, however, the document passes the first test and has been released by one party to a second party it is necessary to ask whether the release was on foot of a common interest in either the litigation or advice.

a) If so, then the document remains privileged, notwithstanding its release by virtue of the doctrine of common interest privilege.

b) If not, then the release might be taken to be a waiver of any privilege which would otherwise have attached to the document.

K.4 Public interest privilege

Public interest privilege is not confined in its application to the executive functions of the State. It is also available where the balance of the public interest favours non-disclosure.

Where a claim of public interest privilege is made the Court is required to balance public interest in the proper administration of justice against the public interest put forward for non-disclosure in order to decide which interest is the superior public interest in the circumstances of the case.

The Executive cannot prevent the Courts from examining documents relevant to any issue in a civil trial for the purposes of deciding if they should be produced.

The categories of public interest in favour of non-disclosure include:

a) National security;

b) International relations;

c) The proper functioning of the public service; and

d) The prevention and detection of crime.

In order for a claim of public interest privilege to succeed, it is essential to show that the communication was brought into being in circumstances of confidentiality. In addition, the Courts will refuse to allow a claim in favour of non-disclosure of a class of documents. In order for the claim of privilege to succeed, it must be particularised and the damage

identified to the public interest in question which will accrue from disclosure of each individual document.

K.5 Journalistic privilege

A journalist may be entitled to withhold from production documents which tend to reveal their confidential sources on grounds of journalistic privilege but such documents must be discovered by listing in the affidavit as to documents as with other privileged documents.

K.6 Privilege against self-incrimination

The privilege against self-incrimination provides a general immunity against any compulsion to produce information or documents which may incriminate the producing party. Where an order for discovery is made and the producing party wishes to assert this privilege in respect of a document or documents, the documents must be listed in the usual way in the first schedule second part and the fact that the privilege against self-incrimination is asserted identified expressly in the affidavit. It is important to note that the privilege must be asserted by the person claiming the privilege, or rather by the person who would be incriminated if the documents were disclosed.

K.7 Waiver of privilege

In order to be privileged a document must be confidential, but confidentiality in itself does not give rise to privilege. It is possible that highly sensitive client documents will not be privileged and must be disclosed.

Disclosure to a third party or for a limited purpose does not always waive privilege and there is no universal rule that the disclosure of documents produced for the sole purpose of seeking legal advice or litigation to a stranger to that litigation constitutes a waiver of privilege in the document. It is advisable to record in writing any conditions regarding a limited disclosure.

It is open to a client to waive privilege and they may do so at any time in proceedings. However, a client may not re-assert their right to privilege once it has been waived either expressly or by implication.

Where a client destroys the confidentiality of a document by choosing to disclose it to the opposing party or to the public generally, any entitlement to assert privilege will be waived.

K.8 Inadvertent waiver

The Courts have upheld the privilege attaching to documents disclosed in error. A solicitor in receipt of a document which appears to be privileged should immediately contact the solicitor for the party whose document has been disclosed to confirm the status of the document and should not read or deploy the document until its status has been confirmed. Upon confirmation, the solicitor should make all reasonable attempts to return the documents and should not make use of them. In general terms, if a document over which privilege may be asserted is inadvertently disclosed without asserting privilege over it where:

a) not to do so was a clear mistake, and

b) privilege was asserted over another copy of the document within the discovery,

the Court may not allow the opposing party to rely on the document disclosed in error.