Greetings from the Executive and Board of the IATJ.

The IATJ is continuing its organizing efforts towards another successful annual Assembly, this being the 7th Assembly to be held in Madrid, Spain on September 30 and October 1, 2016. Registration particulars can be found on our website, as well as information on suggested accommodations.

The Program Committee, chaired by Judge Wim Wijnen from the Netherlands, is putting the final touches on the program which should be circulated within one month or so. Our program always contains varied and interesting topical discussions on substantive and technical issues which I am sure will be of interest to you and your colleagues.

Please reserve September 30 and October 1, 2016 for the 7th Assembly; register early and book your accommodations. Early registration allows the organizing committee to plan the events more accurately and ensure you have a great experience.

I attach for your information an interesting article by Justice Gaston Jorré of the Tax Court of Canada on "predictive coding". This is a concept now being used in some common law jurisdictions to help deal with the large volume of production documents now required in some litigation.

Thank you for your active and continued support for the IATJ and its efforts.

E.P. Rossiter
President

The 2015-2016 executive for the IATJ is:

Chief Justice Eugene Rossiter (Canada), President; eugene.rossiter@tcc-cci.gc.ca
Judge Philippe Martin (France), 1st Vice-President; philippe.martin@conseil-etat.fr
Judge Michael Beusch (Switzerland, 2nd Vice-President; michael.beusch@bvger.admin.ch
Judge Friederike Grube (Germany), Secretary-General;Friederike.Grube@bfh.bund.de
Judge Willem Wijnen (Netherlands), Treasurer; W.Wijnen@ibfd.org

executive members at large include:

Judge Malcolm Gammie (U.K.) mgammie@oeclaw.co.uk,
Judge Peter Panuthos (U.S.A.) stjpanuthos@ustaxcourt.gov,
Counsellor João Francisco Bianco (Brazil) jfbianco@uol.com.br,
Judge Dagmara Dominik-Ogińska (Poland) dagmara.dominik@wp.pl,
Judge Petri Saukko (Finland) petri.h.saukko@oikeus.fi
Pramod Kumar (India) pramod.itat@gmail.com,
Judge Manuel Garzón (Spain) mv.garzon@ts.mju.es,
President Brahim Zaim, (Morocco) zaim009@hotmail.com,
Judge Manuel Luciano Hallivis Pelayo (Mexico) manuel.hallivis@tfjfa.gob.mx,
Justice Tony Pagone (Australia) justice.pagone@fedcourt.gov.au.
President Massimo Scuffi (Italy) massimo.scuffi@giustizia.it

Note on Predictive Coding and Electronic Documents Discovery in Litigation[1]

Justice Gaston Jorré, Tax Court of Canada

Background

When parties are required to produce all relevant documents, there is sometimes the need to review huge numbers of electronic documents ("e-documents") for relevance at the discovery stage of litigation.[2]

E-documents may be in a variety of locations: in-house or outsourced servers, archival storage media, many copiers[3], or individual electronic devices of one sort or another such as personal computers, tablets, and smart phones.

Providing discovery of all relevant documents requires decisions on:

1. what is relevant,

2. the scope of the search and

3. the process for determining relevance of the documents examined.

Relevance: What is relevant is not always straightforward but the test is the same whether the documents are on paper or in an electronic format.

Scope: For paper documents, this includes determining where relevant documents are located, both geographically and functionally.[4] Other issues include determining the relevant time period for the documents.

E-documents raise additional scope questions such as: Are the relevant documents all in servers and, if so, where?[5] Are relevant documents to be found on individual devices of employees? Is hidden data not normally visible when a person reads the document relevant? Are there relevant documents still in electronic memory even though they have

[1] References herein to documents include any kind of information stored in any format.

[2] In the Tax Court of Canada the production of all relevant documents is not automatic. In General Procedure appeals the default rule is that a party must produce all documents that it intends to rely on. Parties may agree on additional production. Alternatively a party may apply to the court for production of all relevant documents; if granted, the court may impose terms, including terms restricting the scope of production.

[3] Not only do many copiers have memories but devices that accept voice instructions may have memories not only of the instructions but also of anything heard by the microphone.

[4] For example: Are they in particular locations, such as the head office? All locations? Certain functional divisions, such as the finance department? All divisions? The company archives?

[5] Given that the servers may have documents from across the organization it may be relevant to ask are the relevant documents those of particular individuals or sent to or from particular individuals?

been deleted?[6] Are there relevant documents that only exist in back-up storage? Should a party be obliged to search for deleted documents and to search in back-up storage?[7]

E-documents create new possibilities for technology assisted review that are not available with paper documents. Potentially these possibilities may reduce the amount of work done by counsel. Considerations of proportionality make it important to consider whether new methods may be appropriate ways of limiting litigation costs.

Technology Assisted Review

Software that may be used to aid the examination of the e-documents includes:

- Software that reviews e-documents by file type, quantifies the number and type of documents and, possibly, converts the documents to a more usable format,
- Software that can eliminate documents based on certain parameters, such as being outside a date range,
- Software that can recognize duplicates or near duplicates,
- Software organizing chains of e-mails and
- Software to assist in determining relevance, or privilege, by means such as keyword searching, objective coding (i.e. coding by dates, names …) or predictive coding.

Technology assisted review may frequently involve using more than one kind of software.

What is predictive coding? [8]

---

[6] Deleting an electronic document removes it from the index but does not necessarily delete the document itself.

[7] Some of these questions may be resolved by establishing conventions that tend to limit the scope of discovery, thereby saving costs. For example, Principle 6 of *The Sedona Canada Principles Addressing Electronic Discovery, Second Edition* (November, 2015) ("*The Sedona Canada Principles"*) suggests that absent good reason a party should not have to search for documents that may still be in electronic memory when they have "… been deleted in the ordinary course of business or within the framework of a reasonable information governance structure."
Rule 29.1.03(4) of the Ontario *Rules of Civil Procedure*, for example, requires parties to consider *The Sedona Canada Principles* in preparing their discovery plan.
Another example of a possible useful convention would be that metadata need not normally be produced. This convention has been adopted by at least one American Court.

[8] An extended discussion of Predictive Coding and issues relating to e-documents is in chapter 5 of *Where the Money Goes, Understanding Litigant Expenditures for Producing Electronic Discovery*, Nicholas M. Pace and Laura Zakaras (RAND Corporation, Institute for Civil Justice, 2012) available free in PDF at: http://www.rand.org/pubs/monographs/MG1208.html. The

Predictive coding software assists in determining which documents in a set of documents are relevant in a particular case. It can also search for privileged documents. The software is designed to learn on an iterative basis about a case on the basis of information provided by counsel.[9] It learns to recognize documents likely to be relevant. [10]
Broadly, the training process is along the following lines:

1. A sample of the entire set of documents is selected. The initial selection process may include documents chosen in several ways, including random samples and key-word searches.

2. Counsel categorize each document as relevant or irrelevant.

3. The categorized documents are provided to the software for study.

4. The software then examines a different sample and categorizes the documents.

5. Counsel then review the software's categorization and correct the results.

6. The corrected results are then given to the software for study.

7. Steps 4 to 6 are repeated.

8. At each iteration, the software learns by studying the correct results and the results of the software's categorization are evaluated.

9. Steps 4 to 6 are repeated *until* counsel are satisfied that:

    a. the software is coding as relevant a high enough proportion of the documents that are, in fact, relevant and

    b. the software is coding as relevant a sufficiently low proportion of documents that are, in fact, irrelevant.

10. Finally, the software is applied to the entire data set.

Chapter also discusses issues of accuracy. Chapter 6 discusses barriers to the use of the software. Other chapters discuss the costs of manual discovery and the accuracy of manual discovery.

[9] The use of predictive coding type software in litigation is relatively new.

[10] *The Grossman-Cormack Glossary of Technology Assisted Review*, Volume 7, Issue 1, 2013 *Federal Courts Law Review*, page 1 provides the following definition of Predictive Coding:

> An industry-specific term generally used to describe a Technology-Assisted Review process involving the use of a Machine Learning Algorithm to distinguish Relevant from Non-Relevant Documents, based on Subject Matter Expert(s)' Coding of a Training Set of Documents.

The term "accuracy" in the field of predictive coding is potentially misleading; see the definition in the *Glossary*.

11. The software will analyse all the documents and produce a predicted relevance score for each document.[11]

12. After looking at the end result counsel decide what is the appropriate cut off point; for example, documents with a score of less than X will not be produced.

There are a number of suppliers of different predictive coding software.

Predictive Coding may be less expensive than manual review of a large set of documents but it is clear that the process of training the software is itself sufficiently time consuming that it will not produce savings if the document set is below a certain size.

Predictive coding will not always make the right decision about a document, but the same is true for human document review.[12] Predictive coding is being used in some civil litigation in Canada and the United States but it is hard to know how common its use is.[13] It is worth bearing in mind:

1. that there are significant costs to training the software and, if it becomes necessary to train it a second time, there will be significant additional costs,

2. that under the rules of our court, and many other courts, a party producing all relevant documents is required to file an Affidavit of Documents that must " … contain a statement that the party has never had possession, control or power of any document relevant to any matter in issue in the proceeding other than those included in the list"[14] and

3. that the quality of the results is dependent on the quality of the particular predictive coding software **and** the quality of the training given by counsel to the software **in the particular case**.

This last consideration, that the software must be good and the training must be good makes predictive coding quite different from the way we used to think of software –

[11] Some software simply classifies documents as relevant or not.
[12] See *Where the Money Goes, Understanding Litigant Expenditures for Producing Electronic Discovery*, footnote 8 above, where there is a discussion of the accuracy of manual discovery at page 55 and following.
[13] Some have expressed reservations about the use of predictive coding. See, for example *Solving the High Cost of the "Review" Stage of Electronic Discovery*, Ken Chasse (posted: May 19, 2014 ; Last revised: June 30, 2015) at http://www.ssrn.com/en/ (SSRN – Social Science Research Network). Other papers by Ken Chasse relating to e-documents at that site may be found by searching for his papers.
[14] *Tax Court of Canada Rules (General Procedure)*, Subsection 82(5).

namely, as being relatively fixed, at least until the next update.[15] Here we are dealing with software that is fluid; any given version from a particular supplier is the same but, after the training specific to a case is completed, it is, in effect, unique software different from the original software.[16]

Some Implications of Predictive Coding for Courts Hearing Tax Litigation

The Tax Court of Canada has not had the occasion to issue a decision dealing with predictive coding.[17] It could arise if a party wanted to use predictive coding in which case there are potentially many issues that could be the subject of dispute.

The range of potential issues with respect to production of electronically stored documents and the desire to minimize costs arising therefrom has resulted in a growing emphasis on the parties working cooperatively in an effort to resolve discovery issues or at least limit them.[18] However, parties will not resolve everything and there will be issues. Let us consider some of those issues in the context of predictive coding.

The quality of the potential results from the use of predictive coding software could be called into question. There could be a debate about the appropriateness of the software or of the training. Because the package of software plus training is unique to every case, in deciding any such dispute one cannot presume that the effectiveness of using such software in any previous cases will be replicated in the case before the court.

Possible issues before the court include: how does the non-producing party evaluate whether the overall result is reasonable without seeing the actual choices made in training the software and without the possibility of access to the whole document set in order to evaluate the process of selection for sample sets? In addition, how can the non-producing party evaluate the reasonableness of the cut off score without actually looking at, or at least sampling, some documents below the cut-off score chosen?

Dealing with issues like these will need to be worked out over time.

---

[15] We are all more and more exposed to software that evolves through interacting with us. One example is dictation software which can be trained as we use it to better understand what we are dictating.

[16] An alternative way of thinking about this is that, once the software is trained, one has a unique combination of software plus training different from the combination of the same software plus training in any other case.

[17] The Court has had issues relating to metadata and discovery; see paragraphs 215 to 243 of *Canadian Imperial Bank of Commerce v. The Queen*, 2015 TCC 280 (CanLII), a decision of Chief Justice Rossiter.

[18] See principles 2, 3 and 4 of *The Sedona Canada Principles*, cited in footnote 7 above. Also apt is the following passage from paragraph 35 of the decision of the Ontario Superior Court of Justice in *Descartes v. Trademerit*, 2012 ONSC 5283 (CanLII):

> … The court does not expect that full agreement will always be possible in an adversarial system. What the court should be insisting upon however is that counsel use their very best efforts to collaborate and there should be little tolerance for overly rigid, technical or petty squabbles that threaten to bog down the litigation in motions.

The 24 February 2012 decision in *Monique da Silva Moore v. Publicis Groupe & MSL Groupe* of the United States District Court (Southern District of New York) is instructive in illustrating one approach to these issues.[19] There the Court ordered the parties to follow a predictive coding process that involved the Plaintiffs to a fair degree in the Defendants' process of setting up and carrying out the predictive coding; much of that process was agreed to by the parties. The protocol governing the process adopted by the Court as its order is very detailed and is 22 pages long.[20]

If a process involving both parties is set up there may be issues arising from the inadvertent loss of privilege.[21] The process set up in *Monique da Silva Moore* raised such privilege issues. In that case, the parties had agreements protecting against the inadvertent loss of privilege; these arrangements are referred to in the protocol.[22]

In manual review a conscious decision is made for each document that it is or is not relevant. However, with predictive coding the producing party potentially faces some difficulty in determining when it has fulfilled its obligation to produce all relevant documents where the software provides the results in terms of a predicted relevance score. If the parties agree on a cut-off point that is not an issue. If they do not agree, the court may have to rule on the question.

Where it is appropriate to try non-traditional methods for reviewing documents the following comments made in paragraph 36 of *Descartes v. Trademerit*[23] are apt:

> Secondly, there is a policy reason to encourage creativity in the discovery process and the court should be slow to criticize parties for mistakes that may be made while exploring the relatively uncharted territory of how to manage e-discovery issues efficiently. Indeed it is inherent in the nature of e-discovery itself that there may have to be trial and error and will likely be false starts. Often the results of

---

[19] Available at: https://thesedonaconference.org/node/4337.

[20] The United States Tax Court has dealt with predictive coding in: *Dynamo Holdings Limited Partnership v. Commissioner of Internal Revenue*, 143 T.C. No. 9 (Sept. 17, 2014).
In *Dynamo* the Respondent, the Commissioner of Internal Revenue, sought production of electronically stored information on two back up tapes. The Petitioner, Dynamo, resisted production but, in the alternative said it should be allowed to use predictive coding to respond to the request because it would cost substantially less than the cost of examining all the documents manually in order to identify what was responsive and what was privileged. The Respondent opposed the use of Predictive Coding on the basis that it was unproven and said that the Petitioner could simply provide it with the tapes and it would agree to a claw-back agreement to protect any privileged material. The court allowed the use of predictive coding.

[21] Indeed, a producing party might use predictive coding to try to identify privileged documents.

[22] The court rules also helped deal with this issue.

[23] 2012 ONSC 5283 (CanLII). There is reference in the decision to key word searches. While predictive coding does not appear to have been used, the comments appear to be generally applicable to e-discovery.

> electronic searches are not what is expected and it is often necessary to try different search strategies to key in on the required information.

Historically, in most cases the scope of discovery in the Tax Court of Canada has been limited thereby keeping the number of documents down to manageable levels under traditional methods. It seems likely that this will continue and the need for methods such as predictive coding to deal with very large numbers of documents will only arise in a modest number of cases.[24] However, it is likely that some document heavy cases will eventually force our court to consider technology assisted review by means such as predictive coding. [25]

[24] One thing that probably helps this is that certain documents are generally organised in a logical and easy-to-retrieve way: for example, accounting records and certain transactional records.

[25] Of course, circumstances where there is full production using technology assisted review in no way relieves a party of the obligation of producing relevant documents that it is aware of. Paragraph 37 of *Descartes v. Trademerit*, 2012 ONSC 5283 (CanLII), put this well:

> Finally, there is nothing in the requirement for a discovery plan nor in the mandate to adopt proportional focused processes of production and discovery which is intended to relieve a party from the responsibility of producing documents (using the expanded definition in the rules) that it knows are relevant to the issues in dispute. In particular a party must produce documents it intends to use at trial or which it should know will be important elements of the proof required by the opposing party.

Madrid, Spain 2016

The IATJ 7th Assembly will be held in Madrid, Spain on September 30 and October 1, 2016 at the Tribunal Supremo de España, Palacio de Justicia, Plaza de la Villa de París, 28071 Madrid, Spain:

The Assembly Agenda is now available; the timetable schedule is as follows:

September 30, 2016

- 8:00 a.m. to 9:00 a.m. – Registration
- 8:45 a.m. to 9:00 a.m. – Welcome by Eugene Rossiter, President and Manuel Garzon, Host 7th Assembly
- 9:00 a.m. to 9:15 a.m. Guest speaker
- 9:15 a.m. to 12:00 p.m. – Substantive Issues
- 12:00 p.m. to 1:30 p.m. – Lunch
- 1:30 p.m. to 4:45 p.m. – Substantive Issues
- 5:45 p.m. to 7:45 p.m. – Cocktail Reception

October 1, 2016

- 8:30 a.m. to 12:00 noon – Substantive Issues
- 12:00 p.m. to 1:30 p.m. – Lunch
- 1:30 p.m. to 3:30 p.m. – Substantive Issues
- 3:30 p.m. to 4:00 p.m. – IATJ Business Meeting (cont'd)
- 6:00 p.m. to 9:00 p.m. - Closing Dinner

The Assembly Registration Fee is $250.00 USD per attendee. Attendees must be Tax Judges or former Tax Judges which (as per the IATJ Statutes) includes Courts, Tribunals or Administrative bodies, judges or retired judges which or who irrespective of their official title, are or were, nevertheless empowered to adjudicate in tax disputes.

[Note: Registration Fee is for the 7th Assembly Conference only and does not include the Closing Dinner on October 1, 2016. The fee for the Closing Dinner on October 1, 2016 is approximately 80 Euros per person. Participants will be responsible for payment of their dinner and that of their guest. Participants must pay for the closing dinner attendance at the morning registration on September 30, 2016.

Please complete the form below and submit **on or before June 30, 2016**. Registration is complete upon confirmation of payment of the registration fee and Closing Dinner fee, if attending.

By cheque/money order or credit card:

Please print, fill out the form and send it together with your payment by cheque or money order to the order of the *International Association of Tax Judges*; or in the alternative, please submit your Visa/Mastercard number, expiry date and three digit security code directly to Mary Doran. Any questions regarding registration or the payment of registration fees should be directed to Ms. Mary Doran, Tax Court of Canada, 200 Kent Street, ON, Canada K1A 0M1 or at iatj.net@gmail.com.

If submitting the Registration form by mail, e-mail or fax, please also include arrangements regarding payment and forward to:

IATJ
c/o Mary Doran
Tax Court of Canada
200 Kent Street
Ottawa, Ontario
Canada K1A 0M1
Fax:613.996.5863
iatj.net@gmail.com

Registration for the conference is a two step procedure:

1. submission of registration form with payment; and

2. the hotel reservation.

Hotel reservations must be made directly with the hotel in order that you may be guaranteed a room at the best rate available at the time. It is strongly recommended that you register as soon as possible to ensure that you obtain a room as rooms are limited in availability.

A list of recommended hotels is attached.

IATJ 7th Assembly Conference
September 30-October 1, 2016
Hotel Information

MADRID

5* Gran Meliá Fenix *(Mr. Jaime Santiago)*
The hotel provides a link for bookings
(Calle de Hermosilla, 2 - Madrid 28001 - Telf: (+34) 91 4316700)
www.melia.com/Gran-Melia-Fenix reservas.gmfenix@melia.com
http://meetings.melia.com/en/7thASSEMBLYIATJ.html#

Reservations made more than 90 days in advance: 15 % discount on public rate hotel
Advance bookings made between 89 days and 15 days: 10 % discount on public rate hotel
Bookings made less than 15 days in advance: 5% discount on public rate hotel

5* Hospes Puerta de Alcalá
(Plaza de la Independencia, 3 – Madrid 28001 – Telef: (+34) 91 4322 911)
www.hospes.com reservations.madrid@hospes.com
single room (breakfast included-wifi)195 € + 10% IVA
double room (breakfast included-wifi) 210 € + 10% IVA

5* Hotel Wellington *(Mr. Pablo Mora)*
(Calle Velázquez, 8 – Madrid 28001 – Telf: (+34) 91 5754400)
www.hotel-wellington.com reservas@hotel-wellington.com
Standard room (breakfast included-wifi) 200 € (IVA included)
Superior room 220 € (IVA included)
Executive room 315 € (IVA included)

4* Hotel Petit Palace Art Gallery *(Mrs. Alba Aguilera)*
The hotel will add a discount code 10 % at the price of the room
(Calle Jorge Juan, 17 – Madrid 28001 – Telef: (+34) 914355411)
www.petitpalace.com/art-gallery artgallery@petitpalace.com
single room (breakfast included-wifi)160 € (IVA included)
double room (breakfast included-wifi) 189 € (IVA included)
triple room (breakfast included-wifi) 229 € (IVA include)

4* Hotel Petit Palace Embassy Serrano *(Mrs. Alba Aguilera)*
(Calle Serrano, 46 – Madrid 28001 – Telef: (+34) 914 313 060)
www.petitpalace.com/Embassy-Serrano embassy@petitpalace.com
The hotel will add a discount code 10 % at the price of the room

single room (breakfast included-wifi)192 € (IVA included)
double room (breakfast included-wifi) 204 € (IVA included)
triple room (breakfast included-wifi) 233 € (IVA include)

4* Hotel Petit Palace Santa Barbara *(Mrs. Alba Aguilera)*
(Plaza de Santa Bárbara, 10 – Madrid 28004 – Telef: (+34) 91 3914421)
www.petitpalace.com/hotel-santabarbara santabarbara@petitpalace.com

The hotel will add a discount code 10 % at the price of the room
single room (breakfast included-wifi)177 € (IVA included)
double room (breakfast included-wifi) 189 € (IVA included)
triple room (breakfast included-wifi) 218 € (IVA include)

4* Hotel NH Collection Madrid Colón *(Mr. Angel Jerez)*
(Calle Goya, 5 – Madrid 28001 – Telef: (+34) 91 5760800)
www.nh-hoteles.es/NH-Madrid-Colon nhsanvy@nh-hotels.com
single room (breakfast included-wifi)155 € (IVA included)
double room (breakfast included-wifi) 185 € (IVA included)

4* Apart Hotel Serrano Recoletos *(Mr. Carlos J. Portal)*
(Calle Villanueva, 2 – Madrid 28001 – Telef: (+34) 914319100)
www.apart-hotelserranorecoletos.com info@apart-serranorecoletos.com
single room (breakfast included-wifi)90 € (IVA included)
double room (breakfast included-wifi) 97 € (IVA included)
(Third person supplement 20 €)