Business & Careers

Streamlining document review with technology

Luigi Benetton
Hi-Tech

Today's discovery teams must pick relevant information from mountains of similar-looking documents. To speed up the process, they can now have computers find relevant documents instead.

The smorgasbord of technologies used to do this is called both technology assisted review (TAR) and computer-assisted review (CAR). These technologies are designed to reduce the time and cost required to find relevant documents.

There are other benefits. People on a review "have been trained on

> If you want to lower your risk, you increase your cost and it takes longer. If you want to shorten the timeframe, you heighten your risk and increase cost.

William Platt
PwC

what to look for but they all process it differently," says William Platt, a partner in e-discovery litigation support with PwC Canada. When lawyers train one computer system to perform the initial review, "You get a more consistent approach."

TAR's evolution began with

keyword searching (or pattern matching) similar to the "find" feature in a word processor. It progressed to concept searching, where a computer determines whether concepts in different documents match.

Predictive coding starts with the review team feeding "seed sets" of documents tagged as relevant, non-relevant, privileged and so forth into the system. Subject matter experts review samples of the results and reclassify records as required. The system uses reclassifications to refine its rules, which are then reapplied to the data set in the next iteration. Once the sample results are correct, the computer's rules are assumed to be valid.

From Susan Wortzman's perspective, predictive coding is a static exercise. "You run through batches of records," says the founder of e-discovery and information governance firm Wortz-

mans, "until the machine says 'Stop! I now have enough information to predict what is responsive or not in the rest of the collection.' Then you run through the whole collection of records."

Machine learning starts analyzing records from the first document fed back to it. "As the team codes, the machine continues to learn," Wortzman says. Learning "is an ongoing process. It isn't static." She claims machine learning predicts what records are most likely to be responsive earlier in the process.

Wortzman had "good success" with this continuous learning method. On one project, from a seed set fed into a system, the team quickly zeroed in on between 65,000 and 85,000 responsive documents in a collection of 1.4 million records. "We were confident we were getting the right information because we were continuously training the

machine as the reviewers were going through."

Bill Dimm argues in a 2015 blog post that "continuous active learning (CAL) has important implications for making review efficient, making predictive coding practical for smaller document sets and putting eyes on relevant documents as early as possible, perhaps leading to settlement before too much is spent on document review."

Dimm, founder of Pennsylvania-based Hot Neuron LLC, which launched the document clustering product Clustify in 2008, added: "It also means that meticulously constructing seed sets and arguing about them with opposing counsel is probably a waste of time if CAL is used."

As technologies go, predictive coding still resides in a vague space devoid of cost certainty where few lawyers tread. "Trust is

---

Ministry of the Attorney General

Requires

Case Management Master

"OPEN"

This is an Order in Council appointment pursuant to s. 86.1 of the *Courts of Justice Act*. You will require legal, adjudicative, mediation and administrative skills. Case Management Masters perform judicial duties in civil actions in the Superior Court of Justice, exercising the jurisdiction assigned to Masters and Case Management Masters by the Rules. At the direction of the Regional Senior Judge, you may be assigned to hear motions, carry out case management functions, conduct pre-trial or case conferences and exercise certain trial management functions. You may also be assigned to conduct references under the *Construction Lien Act* and to serve as registrar in bankruptcy under the *Bankruptcy and Insolvency Act*.

LOCATION: 393 University Avenue, Toronto, Ontario. Could also involve travel to other regions as assigned.

QUALIFICATIONS

Membership in good standing of the Law Society of Upper Canada for at least ten years; good judgment, creativity, strong interpersonal skills; effective organization and computer skills. The position requires strong knowledge of civil litigation, Alternative Dispute Resolution and case management. Knowledge of proceedings under the *Bankruptcy and Insolvency Act* and *Construction Lien Act* would be an asset. A G driver's licence and the ability to travel to surrounding jurisdictions, including Newmarket, Brampton, Milton and Hamilton, is a requirement.

POSTING DATE: November 4, 2016

CLOSING DATE: November 25, 2016

Applications must be received by closing date.

Ministry of the Attorney General
Office of the Director of Court Operations for Toronto Region
Court Services Division
700 Bay Street, Suite 1601
Toronto, ON M5G 1Z6

Attention: Beverly Leonard

AREA OF SEARCH: Within commuting distance of Toronto

Le ministère du Procureur général

est à la recherche d'un(e)

protonotaire chargé(e) de la gestion des causes

« OUVERT »

Il s'agit d'une nomination par décret, aux termes de l'article 86.1 de la *Loi sur les tribunaux judiciaires*. Vous devez posséder des aptitudes juridiques, ainsi que des aptitudes à la prise de décisions, à la médiation et à l'administration. Les protonotaires chargés de la gestion des causes assument des fonctions judiciaires dans des actions civiles à la Cour supérieure de justice, dans l'exercice des fonctions que les Règles confèrent aux protonotaires et protonotaires chargés de la gestion des causes. Selon les directives du (de la) juge principal(e) régional(e), vous pourriez être assigné(e) à l'audition de motions, à la gestion des causes, à la conduite de conférences préparatoires au procès ou de conférences de gestion de la cause et à l'exercice de certaines fonctions de gestion du procès. Vous pourriez également être appelé(e) à gérer des renvois en vertu de la *Loi sur le privilège dans l'industrie de la construction* et à remplir les fonctions de registraire de la faillite en vertu de la *Loi sur la faillite et l'insolvabilité*.

LIEU DE TRAVAIL : 393, avenue University, Toronto (Ontario). Des déplacements vers d'autres régions au besoin seront possibles.

EXIGENCES

Être membre en règle du Barreau du Haut-Canada pendant au moins dix ans; bon sens du jugement, créativité, solide entregent; capacités organisationnelles et connaissances informatiques. Le poste exige une solide connaissance des litiges civils, des modes de règlement extrajudiciaire des différends et de la gestion des causes. La connaissance des instances relevant de la *Loi sur la faillite et l'insolvabilité* et de la *Loi sur le privilège dans l'industrie de la construction* serait un atout. Le candidat ou la candidate doit posséder un permis de conduire de catégorie G et être capable de se rendre dans des villes voisines, dont Newmarket, Brampton, Milton et Hamilton.

DATE D'AFFICHAGE : Le 4 novembre 2016

DATE DE CLÔTURE : Le 25 novembre 2016

Les demandes doivent être reçues avant la date de clôture.

Ministère du Procureur général
Bureau du directeur de l'administration des tribunaux pour la région de Toronto
Division des services aux tribunaux
700, rue Bay, bureau 1601
Toronto (Ontario) M5G 1Z6

À l'attention de : Beverly Leonard

SECTEUR DE RECHERCHE : Toronto et environs

Business & Careers

Predictive coding: Eventually it will be part of routine

on the low side these days because this is fairly new technology," Dimm says.

He has been watching this nascent industry evolve, and he's noticed some confusion. "People make up their own terms for things," he offers as an example. His forthcoming book, tentatively titled *Predictive Coding: Theory and Practice* (predictive-codingbook.com), is a response to this confusion.

Platt offers a basic perspective by describing TAR projects as triangles. The total area of a triangle represents the amount of work involved in a TAR project. The three sides of the triangle are cost, timing and risk. The area

stays the same, so shortening any one side of the triangle means lengthening the other two sides.

"If you want to lower your risk, you increase your cost and it takes longer," Platt says. "If you want to shorten the timeframe, you heighten your risk and increase cost."

"Law firms must find people who are interested in technology," to champion TAR, Wortzman says. Once those people come forward, they can run test cases. "Consider what review would cost as a manual review," Wortzman suggests. "Then do a predictive coding exercise while tracking the costs. The cost savings can be so significant."

Brett Burney describes consult-

ants versed in modern TAR as "meta-project managers" who "work with the law firm, their IT support, the client, their IT support, different vendors," he says.

There's also a training component: "I help lawyers get comfortable with the process," adds the principal of Ohio-based e-discovery and litigation support firm Burney Consultants LLC. "Lawyers have heard about TAR but that's often the limit of their understanding."

Even though TAR has been around for years, learning the skills involved still amounts to amassing experience with both the technology and the discovery process while reading literature published on the topic.

PwC's Platt insists TAR consultants must have a track record on large matters both in the traditional and technology sense. "They must understand the process, how to get documents into a state for manual review," he says, as well as how data analytics works in a broader context, in processes like concept searching, keyword searching and clustering (grouping potentially similar relevant documents).

Any consternation this complexity causes may not last. "Spam filters are an analogy I use to help people get a better comfort level with what a predictive coding tool does," Burney says. "It's a computer making a decision. Is it the ultimate decision? It doesn't have to be, but it's more

often correct than not."

"Most lawyers don't get any spam today, much less than five or 10 years ago," he says, even though law firms routinely filter out untold thousands of messages every day.

Lawyers may find some false positives in their spam filters, so they indicate it's not spam, thus training the spam filter. "You don't even think about this technology. You take this for granted," Burney says, and he figures that in 10 to 15 years most lawyers will take predictive coding for granted, too.

We want to hear from you!
Send us your verdict:
comments@lawyersweekly.ca