

Centro Studi

Diritto **A**vanzato

Edizioni

Comitato scientifico:

Simone **ALECCI** (Magistrato) - Elisabetta **BERTACCHINI** (Professore ordinario di diritto commerciale, Preside Facoltà Giurisprudenza) - Mauro **BOVE** (Professore ordinario di diritto processuale civile) - Giuseppe **BUFFONE** (Magistrato addetto alla direzione generale della giustizia civile presso il Ministero della Giustizia) - Tiziana **CARADONIO** (Magistrato) - Costanzo Mario **CEA** (Magistrato, già Presidente di sezione) - Paolo **CENDON** (Professore ordinario di diritto privato) - Gianmarco **CESARI** (Avvocato cassazionista dell'associazione Familiari e Vittime della strada, titolare dello Studio legale Cesari in Roma) - Caterina **CHIARAVALLOTI** (Presidente di Tribunale) - Bona **CIACCIA** (Professore ordinario di diritto processuale civile) - Leonardo **CIRCELLI** (Magistrato, assistente di studio alla Corte Costituzionale) - Vittorio **CORASANITI** (Magistrato, ufficio studi del C.S.M.) - Mirella **DELIA** (Magistrato) - Lorenzo **DELLI PRISCOLI** (Consigliere Suprema Corte di Cassazione) - Paolo **DI MARZIO** (Consigliere Suprema Corte di Cassazione) - Francesco **ELEFANTE** (Magistrato T.A.R.) - Annamaria **FASANO** (Consigliere presso la Suprema Corte di Cassazione) - Cosimo **FERRI** (Magistrato, già Sottosegretario di Stato alla Giustizia) - Francesco **FIMMANO'** (Professore ordinario di diritto commerciale, Preside Facoltà Giurisprudenza) - Eugenio **FORGILLO** (Presidente di Tribunale) - Andrea **GIORDANO** (Avvocato dello Stato) - Mariacarla **GIORGETTI** (Professore ordinario di diritto processuale civile) - Giusi **IANNI** (Magistrato) - Francesco **LUPIA** (Magistrato) - Giuseppe **MARSEGLIA** (Magistrato) - Roberto **MARTINO** (Professore ordinario di diritto processuale civile, Preside Facoltà Giurisprudenza) - Francesca **PROIETTI** (Magistrato) - Serafino **RUSCICA** (Consigliere parlamentare presso il Senato della Repubblica) - Piero **SANDULLI** (Professore ordinario di diritto processuale civile) - Stefano **SCHIRO'** (Presidente del Tribunale Superiore delle Acque pubbliche) - Bruno **SPAGNA MUSSO** (già Consigliere di Cassazione ed assistente di studio alla Corte Costituzionale) - Paolo **SPAZIANI** (Magistrato dell'Ufficio del Massimario della Corte Suprema di Cassazione) - Antonella **STILO** (Magistrato, Presidente di sezione) - Antonio **URICCHIO** (Professore ordinario di diritto tributario, Magnifico Rettore) - Antonio **VALITUTTI** (Presidente di Sezione presso la Suprema Corte di Cassazione) - Alessio **ZACCARIA** (Professore ordinario di diritto privato, già componente laico C.S.M.).

PREDICTIVE JUSTICE

(about Viola, Interpretation of the law through mathematical models, London 21 june 2019)

Simon WHITE

(Principal consultant – Data & Predictive Analytics at DWF Consulting)

Good evening to you all and it is a privilege to speak with you today.

I work for DWF Law LLP as a consultant for data and predictive analytics. I spend a significant proportion of my time building predictive models to forecast outcomes of all types, across many sectors, but I started my career in analytics building models to forecast the outcomes of civil claims, particularly in insurance. I am not a qualified lawyer although I have a reasonable understanding of the Civil Law of England & Wales.

Building predictive models to forecast the outcomes of civil claims does not involve the application of the relevant law – rather, the model looks for case commonalities in historical data, each with varying correlation strengths, and uses the model to forecast the percentage propensity for a given outcome. As my firm specialises in defending cases, the forecast outcome typically relates to the propensity of a successful defence.

The practical use of the models has been as an aid to the decision-making of the lawyer. If a lawyer considers they have strong evidence to defend the case, then they will not, and consider that they do not need, to make use of the model, and that is entirely understandable.

Where the models have proven effective is on cases where the evidence to defend has been finely balanced – which, you might argue, constitutes the majority of litigated civil claims – or worse, where the evidence available to defend is a long way short of even finely balanced. Here, the models have proven useful in the successful defence of higher volumes of finely balanced cases along with some evidentially poor cases, where the model predicted a reasonable chance of a successful defence. I would say this is an example of Intelligence Augmentation (or IA) rather than AI.

Data sources have typically been an issue in terms of completion, accuracy and granularity of available data. All of these issues affect the accuracy of the model's propensity scoring and therefore affect human trust of the model. In the discipline of IA, human trust in the model is essential. As text mining and automation technology improves at an exponential rate, these issues are

eroding accordingly. When dealing with non-numerical data within a predictive model, accurate taxonomy of historical data is of significant importance and still requires human input with domain expertise – our lawyers, in other words.

The single biggest issue we have had to overcome has been to gain human trust in the accuracy of the outputs of the predictive model. Lawyers can be conservative by nature and when dealing with predictive data models, generally gravitate towards the side of caution and in some cases, outright distrust. That distrust exists for a number of very good reasons – can human legal expertise ever be reflected with full accuracy in algorithms? In addition, if so, will the algorithms replace human lawyers in time? How do we know the source data is fully complete, accurate and properly classified?

We have overcome those trust issues in several ways. We have demonstrated our models to our lawyers in detail – showing them why each variable within the model exists, the strength of each variable's correlation to the forecast outcome, and by testing hypothetical cases with them and inviting them to predict the result – finding that their predictions were largely matched by the model. I would say that, on average, around 80% of the model workings are validated by our lawyers' anecdotal experience and expertise. This 80% human trust then sees lawyers generally accept and trust the 20% of the model they did not know to be true.

The concept of predictive justice is of course a far more complex arena, although similar principles of data science are applicable. Where the standard of proof is 'on the balance of probabilities' – does that equate to a model propensity of 50.01%? The interesting question is – where the standard of proof is 'beyond reasonable doubt', what is the mathematical propensity that 'beyond reasonable doubt' equates to? Is it 80%? 90%? Higher? What is 'reasonable' in the discipline of mathematics?

I cannot foresee a time – within my lifetime at least – when human lawyers or judges no longer exist – instead, the role of the human lawyer is evolving, largely at the same rate as data and technology capabilities albeit with a significant time lag between the two. Human lawyers are needed to inform data mining technologies to convert unstructured data from documents into

structured and categorised data. In a world where IA is more prevalent than AI, human lawyers are needed to make decisions, albeit better decisions informed by data.

Moreover, what of judges? When we reach a stage where predictive justice technologies can be theoretically harnessed effectively, will judges begin to operate in an IA world that gradually moves towards true AI? What of claimants, defendants, prosecutors, and the lawmakers themselves? I would suggest that the biggest obstacle to overcome for all stakeholders in the fields of predictive justice is trust in the predictive models and the raw data that informs them. Trust is a human emotion, not easily reproduced by AI.

In conclusion, the potential operationalisation of predictive justice lies in the balance of benefits against risks for all stakeholder groups. Rather like the criminal standard of proof, at what point will the risks of predictive justice be low enough to be 'beyond reasonable doubt?'

Thank you for listening. I am happy to take any questions.