Turning Text into Data: An Exploration of the Potential of Natural Language Processing Techniques in Criminal Justice Agencies Using Release Decisions from the Parole Board

Summary

The Lammy Review (2017) highlights significant difficulties in accessing statistics with which to explore ethnic disparities in the Criminal Justice (CJ) system in England and Wales, and urges CJ agencies to redress this issue. In some instances, the solution simply involves publishing anonymised versions of existing datasets. In many other cases, the required data do not exist. Adopting new data collection processes is often unrealistic since it requires additional investments from agencies that have seen their budgets depleted over the last decade. In response to these constraints, this interdisciplinary project will explore an alternative cost-effective route to generate new data capable of highlighting potential disparities in the CJ system – by applying natural language processing techniques to the large volumes of free-text data stored in existing administrative CJ records.

All CJ agencies generate large numbers of records documenting the characteristics of cases processed and decisions adopted. To illustrate, the Crown Prosecution Service produce pre-sentence reports used in court by judges and magistrates, who themselves also document their decisions in sentence transcripts. There is a long tradition of manually processing such records using content analysis to code relevant text information into statistical datasets, thus enabling quantitative analyses (e.g. Hood, 1992; Myers and Talarico, 1987). The key problem with these techniques lies in their scalability. The cost/time of processing records manually is directly proportional to the number of cases to be processed, which renders samples either too small or too expensive (Pina-Sánchez et al., In Press). Following advances in the field of Data Science and Artificial Intelligence, we propose the use of text-mining techniques to undertake such coding process automatically.

Recently, Pina-Sánchez et al. (2019) demonstrated the potential of such techniques for the processing of sentence records. Yet this proof of concept is still affected by important limitations regarding document validity and the sophistication of algorithms presented. Building on this previous work, and partnering with The Parole Board - a key national CJ agency - the project will push the methodological frontier in this crucial area of research. The Parole Board carry out risk assessments to decide whether prisoners can be safely released into the community. In their latest annual exercise The Parole Board (2018a) processed 16,436 'paper-hearings' for which short structured textual summaries (roughly two-pages long) were routinely recorded. These 'hearing summaries' capture the main characteristics of the case, together with demographic factors of the prisoner including their ethnicity.

Analysing a significant sample of these 'hearing summaries' the project objectives are: i) to develop text-mining algorithms capable of processing 'hearing summaries'; ii) to assess the reliability of the data these algorithms generate; and iii) to analyse the data they produce to explore potential racial disparities in Parole Board decisions.

The PhD student will be supervised by two academics with expertise in criminal justice and computational methods and Glenn Gathercole, Head of Business Development at The Parole Board (who will participate in bi-monthly supervision meetings). In year two the student will undertake a four-month internship at The Parole Board, providing invaluable insight into the realities of practice, and offering key opportunities to refine initial methods developed to reflect the environments in which may be applied.

In conclusion, the proposed interdisciplinary project and the strategic partnership it leverages will provide the PhD student with a unique opportunity to explore how data science methods can be applied for public good in the context of criminal justice outcomes. By disseminating the statistical data and methods generated, the project will also enable innovative new lines of CJ research, and, it is hoped, facilitate the adoption of new cutting edge methods by CJ analysts across a range of government agencies (e.g. processing of pre-sentence reports, sentence transcripts, and virtually any other CJ records). Doing so the project has

a very realistic potential of providing a direct response to David Lammy's request for datasets shedding new light on any potential disparities in the CJ system.

Description

Background

The Lammy Review (2017, p.13) established that *"[…] the CJ system should aim to improve the quality and quantity of datasets made available to the public"* to enhance external scrutiny - from academics, journalists, campaigners and the general public at large – and in so doing help detect and redress any potential differentials in the treatment of minorities in the CJ system. The review specifically requires the disclosure of existing datasets, currently operated centrally by various CJ agencies, and the collection of new datasets by those CJ agencies system where no information on individuals' ethnicity is available. This new emphasis on transparency through data collection/publication is welcome, however, we believe the recommendations are not specific enough to ensure the materialisation of such good intentions.

We can point at two specific problems. First, existing data is not detailed enough. Capturing outcomes at a particular stage of the CJ - say sentences imposed at the Crown Court - together with nothing more than the ethnic background of the individuals 'processed', constraints the analysis to descriptive statistics, which limits severely the validity of the estimation of unwarranted disparities, and are incapable of delving into the mechanisms causing such hypothetical disparities. Second, the CJ system does not have the resources to collect new data. Asking agencies that have seen their budgets systematically depleted over the last decade to invest in non-essential data collection processes is not very realistic. This is clearly evidenced by the Sentencing Council for England and Wales decommissioning of the Crown Court Sentencing Survey in 2015 - arguably the most detailed sentencing datasets worldwide (Pina-Sánchez and Grech, 2017; Roberts and Hough, 2015).

A recent assessment from the Justice Committee (2019) points at some areas of progress in the data collection front following the Lammy Review, e.g. new initiatives considering gypsy and traveller populations have been highlighted. However, the main issue remains, there are remarkably few CJ datasets publically available capturing information on individuals' ethnicity at a sufficient level of detail. A proof of the unavailability of individual level data – and to some extent, the failure in achieving the goals set out by the Lammy Review - is that, as far as we are aware, no multivariate study exploring ethnic disparities in the CJ system using official data has been published since the Review. Descriptive studies can be useful, but their internal validity is questionable. On the question of discrimination, and to borrow a metaphor, such analyses only allow us to scratch the surface of the problem. Since no relevant legal characteristics are controlled for, 'evidence' of differential treatment obtained from descriptive analyses should be challenged, as it cannot disentangle legitimate forms of differential treatment (e.g. release decisions rejected as a result of lack of crime committed while imprisoned) from discriminatory practices (Hofer et al., 1999; Pina-Sánchez and Harris, In Press; Pina-Sánchez and Linacre, 2016). Descriptive studies pointing at no evidence of differential treatment should be similarly challenged for the same reasons. For example, it is possible that racial disparities are determined not as differences in outcomes, but in how different factors are applied (see Lightowlers 2018, who found that Crown Court judges use the aggravating factor of 'alcohol and drugs intoxications' more punitively against female offenders).

Here we suggest an alternative route to generate the necessary data with which undertake the necessary analyses. We propose tapping onto the vast wealth of documents recorded throughout the CJ system. Text-mining techniques can be employed to code individual CJ records into statistical datasets capturing the main characteristics of individual cases, with which to be able to undertake the required multivariate analyses on ethnic disparities. Text-mining techniques are widely applied in areas such as business analytics or medical research (Jensen et al., 2012; Lim et al., 2013), but they are only in their infancy in their application to CJ research (e.g. Oksanen et al., 2017; Pina-Sánchez et al., 2019; In Press), which is largely due to the difficulty

of accessing CJ records (Baldwin, 2008; Dhami and Souza, 2010; Ulmer, 2012). This project overcomes such hurdle by partnering with The Parole Board.

The Parole Board is a key CJ agency operating at the national level. It is in charge of carrying out risk assessments to decide whether prisoners can be safely released into the community. In their latest annual exercise The Parole Board (2018a) processed 16,436 'paper-hearings' for which short summaries (roughly two-pages long) were routinely recorded. These 'hearing summaries' capture the main characteristics of the case, together with a few demographic factors of the prisoner, amongst them their ethnicity. Equally important, these records are written following a highly structured format, lending themselves to be processed effectively through text-mining techniques.

In summary, the policy focus on measuring the extent and origins of ethnic disparities (established in the Lammy Review), the lack of official statistics with which to undertake such study, and the existing wealth of text data in the CJ system, clearly sets out the potential of text-mining techniques in this area of research. If we add the role of The Parole Board as an active external partner to this project, providing access to their records, offering supervisory guidance and in-kind support, we can see how this is an uniquely promising project. Last but definitely not least, it is worth highlighting the scalability of the methodology suggested here. Once the project is completed, The Parole Board will be able to generate detailed, individual-level, statistical data at virtually no added cost. This is a key practical – but often overlooked – condition for a data collection process to become viable in the current economic context. If successful, we have no doubt that similar processes will be replicated by other CJ agencies, which will provide the necessary building blocks to address many of the challenges set out in the Lammy Review, ensuring that England and Wales *"[…] are once again at the cutting edge of research and analysis in this area"* (Lammy Review, 2017, p.14).

Analytical Strategy

The project is divided in two main stages, each of them tackling a specific research question. The first being mainly methodological the second is of a more substantive nature.

RQ 1: How reliable can data generated through the text mining of 'hearing summaries' be?

RQ 2: Is there evidence of unwarranted racial disparities in The Parole Board release decisions?

The first part of the project involves the coding of the main characteristics of cases seen by The Parole Board, as recorded in their 'hearing summaries', into a statistical dataset. This will be done following two parallel avenues: i) through the design of an automatized text-mining algorithm using the full sample of 16,436 'paper-hearings' available from the 2017 exercise; and ii) through a more traditional coding process of a smaller sample of 1,000 cases using content analysis. The latter will be used as a gold standard with which to assess and enhance the reliability of the former. The second part of the project involves using the statistical dataset generated in phase one to explore the presence of potential ethnic disparities in The Parole Board release decisions. This involves the specification of regression models capable of controlling for key case characteristics such as the offence type, age of the offender, and number of previous convictions.

The two tasks are clearly specified. In addition, the data to be used is already available in an adequate format, and ready to be shared by The Parole Board with the PhD student from the onset of the project. These two conditions been met maximise the feasibility of the project and minimise any potential risks. Even if the PhD student to be in charge of this project does not have a background in programming and/or data analysis, the necessary training of those skills, together with the required time to undertake the two phases of the project, can be realistically expected to be undertaken within less than two years. Once those two key goals are achieved there are multiple opportunities for the PhD student to take different areas of the project forward independently. For example, there is the potential of taking phase one forward by applying more sophisticated natural language processing techniques, with which to code subtler forms of information potentially traceable in the 'hearing summaries'. Similarly, depending on the interest of the

student, phase two can be expanded by the adoption of more sophisticated regression models, such as multilevel models, capable of analysing how consistently different case characteristics are applied across practitioners involved in hearings (Pina-Sánchez and Linacre, 2013). Or perhaps the PhD student might want to take advantage of the opportunity brought by The Parole Board to observe how hearings are undertaken in real life, and use the qualitative insights gained to improve the accuracy of the text-mining of hearing summaries or the multivariate analysis of the data to be generated.

Involvement of the Parole Board

The Parole Board collaboration as an external partner – right from the inception of this proposal – represents a major strength of the project. The relevance of their work cannot be overstated, both in terms of the sheer volume of work undertaken (24,573 paper and oral hearings conducted in 2017), but also in terms of the paramount significance of the outcomes of each of those hearings, which are nothing less than 'life-changing' for prisoners, and to the community where they will be released too.

It is also worth highlighting The Parole Board's commitment to research, well evidenced by the publication of their recent 'Research Governance Framework' (The Parole Board, 2018b), where a set of procedures are laid out to facilitate and encourage external research on their work. The Parole Board sense of obligation towards cutting edge research can also be demonstrated from their response to the Lammy Review (2017). The review requested The Parole Board to increase the transparency of their decisions by providing a breakdown of release figures by ethnicity. This request has been carried out and published in their latest annual report (The Parole Board, 2018a), showing no evidence of ethnic disparities in the treatment of BAME prisoners. However, as discussed during our preliminary meetings, this apparent 'no-effect' could be due to the confounding effect of relevant factors not controlled for, or perhaps it could be even masking a discriminatory treatment in how case characteristics are applied differently across demographic groups. Two possibilities that have not been ruled out by The Parole Board.

By taking into account the characteristics of each case, the analyses to be undertaken in this project will provide evidence of higher validity than any previous research on this topic. If evidence of disparities is found, this project will also provide The Parole Board with a series of insights exploring the origin of such disparities. For example, it will be possible to explore whether disparities originate as a result of key case characteristics being applied inconsistently as a result of the ethnic background not of the prisoner, but of The Parole Board practitioner conducting the hearing. This type of causal evidence will become instrumental in the design of policies to redress such potential disparities.

Regardless of the outcome of the statistical analysis, this project will also deliver a unique dataset of release decisions that could be further exploited by The Parole Board and other external researchers to investigate other relevant questions, not necessarily related to release disparities. For example, are more experienced practitioners more lenient? Or, is The Parole Board subject to the same 'penal populism' phenomenon that has been observed at the sentencing stage (Roberts and Hough, 2005)? Lastly, besides the generation of evidence of higher validity on release disparities, and the new dataset coding release decisions, this project will also provide The Parole Board with the text-mining algorithm to be designed by the PhD student. In order to ensure the long-term impact of the project, a tutorial on how to expand this algorithm to take into account changes of formatting or to capture new variables will also be created. A workshop will be conducted to ensure that The Parole Board's analysts are capable of replicating the techniques covered in the tutorial independently.

To facilitate research co-production, to help in the steering of the project, and to provide practitioner-level guidance, Glenn Gathercole, Head of Business Development at The Parole Board, will participate in bi-monthly supervision meetings throughout the duration of the project. Glenn will also design a four-month internship at the beginning of year-two of the project to provide the PhD student with the opportunity of working alongside professionals conducting parole hearings. This internship is a crucial stage of the project. The shadowing of practitioners will offer more refined qualitative insights into the hearing process, which

will enable the student to assess the potential of extending the text mining algorithm to include relevant variables that had been initially missed, but also to assess the reliability of the records upon which such algorithm will be applied. At the end of the project the PhD student will also be scheduled to spend another week working from the offices of The Parole Board to conduct the workshop ensuring that the analysts at The Parole Board are capable of replicating and extending the text-mining algorithm designed.

The benefits to all parties involved, and more generally to the CJ field of research, are therefore obvious. This project will enhance our understanding of release decisions from The Parole Board, will tackle specific questions on potential ethnic disparities set out by the Lammy review, and, potentially, demonstrate and spread the use of text mining techniques across CJ agencies, in the UK and elsewhere. More specifically, The Parole Board will receive a dataset coding their decisions, new evidence on whether ethnicity plays a role in those decisions, and a text-mining algorithm ensuring the long run impact of the project. The supervisory team will benefit by gaining a key CJ agency as a new research partner. A series of scholars from the School of Law at the University of Leeds share a strategic plan to develop a centre of Empirical Criminal Justice Studies within the School. Strengthening links with The Parole Board will contribute importantly towards achieving that goal. While the student will benefit from unique access to CJ records, fieldwork activities that normally take a long time to materialise, and a supervisory team composed of experts on the topic and methods to be employed. These opportunities and the remarkable potential for future academic growth are discussed in the next and final section of this proposal.

Student Training and Academic Growth

The proposed project will offer the student a unique opportunity to develop truly interdisciplinary research expertise in an area with great potential for societal impact. It will also provide the student with first-hand experience of how a national CJ agency operates, an opportunity not always available to academic researchers. The project requires the student develop mastery in both social science and quantitative/computational skills – a highly desirable combination in the modern world of policy, practice and academia – that will place the student at the cutting edge of criminal justice and social science research.

A candidate with a social science background will have the opportunity to gain skills in data science methods, while a student with a more technical background will develop key social science research skills. Irrespective of their background, the student will gain practical experience of decision-making in a key CJ agency and engage in important research in a timely and significant area of enquiry with wide reaching societal implications. While the core structure of the project is proposed above, there is considerable scope for the student to make the project their own and focus their research and development in one or more of several potential areas of interest. To illustrate, a student who wants to emphasise the text mining side of the project could apply more complex natural language processing techniques capable of deciphering subjective traits underlying within documents, such a positive/negative emotional charge, the use of professional or more lay person English, the presence of patronising remarks, etc. Alternatively, the student could focus a greater proportion of their efforts into the statistical analyses of data generated using the text mining methods. For example, using advanced regression methods, such as multilevel or non-parametric models to explore whether characteristics are applied uniformly. Another possibility could be to explore issues related to the reliability of documents using measurement error models. Ultimately, given the methodological breadth and the compatibility of the different members of the supervisory team, there are many interesting avenues the PhD student could pursue.

The student will be supervised by an interdisciplinary team who are both internationally recognised research leaders in their respective fields. The team have strong connections with a range of relevant public sector and academic networks which the student will have direct access to throughout the project including the European Society of Criminology Working Group on Sentencing and Penal Decision Making, The Leeds Institute for Data Analytics, and The Alan Turing Institute (where Birks is a Turing Fellow). In particular The Leeds Institute for Data Analytics, and The Alan Turing Institute will provide the student with access

to advanced natural language processing expertise and training where necessary. Throughout the project, the supervisory team will encourage the student to disseminate the outcomes of their research both through lead-authored academic manuscripts, where appropriate policy/practitioner briefs, and through the attendance of national and international conferences/symposium.

Furthermore, the active involvement of the Parole Board in the supervisory team will ensure the student and their research are uniquely grounded within an actual policy-making context. In particular the structured work placement will provide the student with unique opportunities to test and refine their ideas in a real world setting and develop their own networks with practice/policy in a national CJ agency. Collectively, these opportunities enabled by the strategic partnership between the University of Leeds and the Parole Board will ensure that the student's development is firmly grounded in realities of practice.

References

Baldwin, J. (2008). Research on criminal courts. In: King, R., Wincup, E. (Eds.*), Doing Research on Crime and Justice.* Oxford University Press, Chippenham, pp. 375-398.

Dhami, M. K., Souza, K., 2010. *Breaking into court*. In: Streiner, D., and Souraya, S. (Eds.), When Research Goes off the Rails. Guildford Press, New York, pp. 81-87.

Hofer, P. J., Blackwell, K. R., & Ruback, R. B. (1999). The effect of the federal sentencing guidelines on inter-judge sentencing disparity. *The Journal of Criminal Law and Criminology*, 90(1), 239-322.

Hood, R. G. (1992). *Race and Sentencing: A Study in the Crown Court. A Report for the Commission for Racial Equality.* Oxford University Press.

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics*, 13(6), 395.

Justice Committee (2019) Oral evidence: Progress in the implementation of the Lammy Review's recommendations. *House of Commons.* Available at: https://www.parliament.uk/business/committees/committees-a-z/commons-select/justice-committee/inquiries/parliament-2017/lammy-review-progress-inquiry-17-19/

Lammy, D. (2017). *The Lammy Review: An Independent Review into the Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/643001/lammy-review-final-report.pdf

Lightowlers, C. (2018). Drunk and Doubly Deviant? The role of gender and intoxication in sentencing assault offences. *The British Journal of Criminology*, 59(3), 693-717.

Lim, E. P., Chen, H., & Chen, G. (2013). Business intelligence and analytics: Research directions. *ACM Transactions on Management Information Systems (TMIS)*, 3(4), 17.

Myers, M. A., and Talarico, S. M. (1987). *The Social Contexts of Criminal Sentencing.* Springer, New York.

Oksanen, A., Tuominen, J., Mäkelä, E., Tamper, M., Hietanen, A., & Hyvönen, E. (2017). Law and Justice as a Linked Open Data Service. Available at: https://www.semanticscholar.org/paper/Law-and-Justice-as-a-Linked-Open-Data-Service-Oksanen-Tuominen/b28d208de7da06b9c8b40ac0bc81a06c32ae1698

Pina-Sánchez, J., & Grech, D. C. (2017). Location and sentencing: To what extent do contextual factors explain between court disparities? *The British Journal of Criminology*, 58(3), 529-549.

Pina-Sánchez, J., Grech, D., Brunton-Smith, I., & Sferopoulos, D. (In Press). Exploring the origin of sentencing disparities in the Crown Court: Using text mining techniques to differentiate between court and judge disparities. *Social Science Research.*

Pina-Sánchez, J., & Harris, L. (In Press). Sentencing gender? Investigating the presence of sentencing gender disparities in the Crown Court. *Criminal Law Review.*

Pina-Sanchez, J., & Linacre, R. (2013). Sentence consistency in England and Wales: Evidence from the Crown Court sentencing survey. *British Journal of Criminology*, 53(6), 1118-1138.

Pina-Sánchez, J., & Linacre, R. (2016). Refining the measurement of consistency in sentencing: A methodological review. *International Journal of Law, Crime and Justice*, 44, 68-87.

Pina-Sánchez, J., Roberts, J. V., & Sferopoulos, D. (2019). Does the Crown Court discriminate against Muslim-named offenders? A novel investigation based on text mining techniques. *The British Journal of Criminology*, 59(3), 718-736.

Roberts, J.V. and Hough, M. (2005). *Understanding Public Attitudes to Criminal Justice*. Maidenhead, Kent: Open University Press.

Roberts, J. V., & Hough, M. (2015). Empirical sentencing research: Options and opportunities. In *Exploring Sentencing Practice in England and Wales* (pp. 1-17). Palgrave Macmillan, London.

The Parole Board (2018a). *Parole Board for England and Wales Annual Report and Accounts 2017/18*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/727619/Parole_Board_Annual_Report___Accounts_2017-18.pdf

The Parole Board (2018b). *Research Governance Framework*. Available at: https://www.gov.uk/government/organisations/parole-board/about/research

Ulmer, J. T. (2012). Recent developments and new directions in sentencing research. *Justice Quarterly*, 29, 1-40.