Social licence and digital trust in data-driven applications and AI: a problem statement and possible solutions

Peter Leonard[1]

‘Social licence’, ‘social capital’ and ‘social contract’ variously describe:

- acting fairly and responsibly, as well as in compliance with laws – in business, sometimes styled as acting with corporate social responsibility,
- acting “ethically”,
- maintaining trust of relevant stakeholders, be they customers, consumers or citizens.[2]

The term ‘social licence’ was first intended to be used in juxtaposition to a ‘legal licence’. To build a mine, or log a forest, you need a legal permit or licence as issued by an authorising authority. But you also should seek broad community support for what you want to do, for which you should engage in community consultations and consensus building. Hence the term social licence, and a clear distinction between what may be legally permissible, and what you should do.[3]

[1] Peter Leonard is a data, content and technology business consultant and lawyer and principal of Data Synergies. He is also a Professor of Practice at UNSW Sydney Business School. Peter was a founding partner of Gilbert + Tobin and following his retirement from that law partnership he continues to assist Gilbert + Tobin as a consultant. Peter chairs the IoTAA’s Data workstream, the Law Society of New South Wales’ Privacy and Data Committee and the Australian Computer Society’s Artificial Intelligence and Ethics Technical Committee.

[2] A good introduction to the history and philosophy of social licence is Morrison, J. (2014) The Social License: How to Keep Your Organization Legitimate. For a functional definition of social licence, see for example https://socialicense.com/definition.html (accessed 2 October 2018). For analysis of the concept of social capital, see Claridge, T. (2014) “Social Capital and Natural Resource Management: An important role for social capital?”, Unpublished Thesis, University of Queensland, Brisbane, 2004, Australia. Available online https://www.socialcapitalresearch.com/literature/definition/ (accessed 2 October 2018). A recent exploration of social licence for data technologies was undertaken in New Zealand by The Data Futures Partnership, which was tasked by the NZ Government to develop guidelines that public and private organisations can use to develop ‘social licence’ for data use. The Partnership characterised social licence as follows: “When people trust that their data will be used as they have agreed, and accept that enough value will be created, they are likely to be more comfortable with its use. This acceptance is referred to as a social licence”. The Partnership summarised its key conclusions in (2017) “A Path to Social Licence: Guidelines for Trusted Data Use”, available at https://trusteddata.co.nz/ (accessed 2 October 2018). The Partnership’s research and a background document with practical advice and examples are also available that that site. The Guidelines focus on eight key questions that organisations can answer to explain how they collect and use data, to better build trust with clients and the wider community, as discussed later in this paper. See further Australian Computer Society (2018) Trust Preserving Data Sharing Frameworks: Data Sharing Taskforce Whitepaper, forthcoming publication.

[3] A proposal to amend the ASX Corporate Governance Principles to include a requirement for ‘social licence to operate’ encountered stringent criticism for crossing the line between legal requirements and corporate beneficence. See consultation paper entitled Review of the ASX Corporate Governance Council's Principles and Recommendations and submissions as to that consultation paper as available at https://www.asx.com.au/regulation/corporate-governance-council/review-and-submissions.htm; also Sally Patten “Social licence: why some companies still don't get it, and how it will cost them”, Australian Financial Review 1 October 2018 available at https://www.afr.com/brand/boss/social-licence-why-some-companies-just-cant-pass-the-test-20180912-h15akg; ‘Chanticleer’, “Hayne royal commission's harsh reminder of the importance of a social licence” Australian Financial Review 21 May 2018, available at https://www.afr.com/brand/chanticleer/royal-commissions-harsh-reminder-of-the-importance-of-a-social-licence-20180521-h10cwv; Patrick Durkin, “Board outrage over push to have a social licence”, Australian Financial Review 1 August 2018 available at https://www.afr.com/leadership/board-outrage-over-push-to-have-a-social-licence-20180731-h13doa and Yolanda Redrup, “How digital inequality and automation led to the rise of Trump, Dutton”, Australian Financial Review 1 October 2018 available at https://www.afr.com/technology/how-digital-inequality-and-automation-led-to-the-rise-of-trump-dutton-20180928-h15zo3 (each accessed 2 October 2018). The debate will probably escalate in anticipation of and following final findings of the ongoing Australian Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry (the Commission’s website is https://financialservices.royalcommission.gov.au/Pages/default.aspx).

Social licence may follow broad social consensus as to what is ‘right’ or ‘wrong’. But in a fast changing world, civil society cannot keep up. Compare the rapid development of genetic engineering technology (CRISPR) with the slower rate of the understanding of CRISPR that is required to underpin engagement as to acceptable applications and outcomes.[4] Compliance with existing laws is usually necessary but often insufficient. Technology and social mores move faster than legislatures and their statutes, so frequently the law reflects yesterday’s problems and standards.

Social licence/capital/contract are problematic terms. Too often particular organisations assume that because of who they are, and a heartfelt and expressed commitment to social beneficence, they are entitled to social licence until they are demonstrated to not be good actors. Social licence relates to particular activities of an organisation: to what an organisation *does*, as opposed to what *it is*. Social licence cannot be directly managed or self-awarded: rather, it is the combined outcome of a number of factors. It is these factors that organizations can manage, not the social licence itself. And it is much easier to notice the absence of the social licence than its presence.[5]

Legislated authority sometimes confers social licence, particularly where the authority is conferred with demonstrable controls and safeguards. But limited social licence arises from a broadly conferred power that can be exercised without external engagement as to reasonableness, proportionality and justified necessity of applications of that power. External oversight of broadly conferred powers may confer legitimacy, but seldom social licence. In our world of rapid change and erosion of trust in institutions, social licence through prior beneficence and lofty purpose is often no longer good enough, even for relatively high trust organisations such as some government agencies (for example, the Australian Bureau of Statistics and the Australian Institute of Health and Welfare). Social media has given voice to a cacophony of competing views often articulated in expectation of inclusion within a debate, even where the contributions are not fully informed and moderately expressed. ‘Society’ (as referred to in the ‘social’ licence) was never a cohesive voice, but increasingly views of different constituencies need to be sought out, empowered to be expressed and taken into account. These constituencies will often need informing as to what it is for which social licence is sought: engagement without empowerment to enable meaningful expression of views and dialogue will not lead to social licence. The relevant constituencies, whether they be citizens, consumers, customers or users, need to be enabled to develop views for articulation in a way that reasonably counterbalances better resourced, more informed and more self-interested stakeholders.[6]

Can ‘consumer advocates’ within organisations provide sufficient proxy for constituencies? It is difficult for any individual to anticipate and synthesise views of diverse voices. Special needs of particular groups, such as addressing diverse accessibility requirements, may be

[4] See further Dronov, R. & Howard, W. (Office of the Australian Chief Scientist), (2017) “Gene Editing and CRISPR”, available at https://www.chiefscientist.gov.au/wp-content/uploads/OCS-Gene-Editing-and-CRISPR.pdf; Australian Council of Learned Academies (2018) The Future of Precision Medicine in Australia, Final Report and input papers as available at https://acola.org.au/wp/pmed/; Brokowski C. and Adli, M (2018) “CRISPR Ethics: Moral Considerations for Applications of a Powerful Tool” J Mol Biol. 2018 Jun 7 (pii: S0022-2836(18)30586-2. doi: 10.1016/j.jmb.2018.05.044); also National Human Genome Research Institute (2017) “What are the ethical concerns about genome editing?” and the publications referenced there, as available at https://www.genome.gov/27569225/what-are-the-ethical-concerns-about-genome-editing/ (each accessed 2 October 2018).

[5] This discussion largely follows Morrison, J. (2014) The Social License: How to Keep Your Organization Legitimate.

[6] Contrast, for example, the detailed contribution to industry dialogues and industry self-regulatory initiatives where sector-specific consumer advocacy bodies are industry-funded to provide consumer input for the sector - such as the Australian Communications Consumer Action Network (ACCAN) (http://accan.org.au/about) and Energy Consumers Australia (http://energyconsumersaustralia.com.au/about-us/), with limited specialist consumer input dedicated to the Australian financial services sectors.

overlooked or not properly evaluated.[7] A well-meaning and empowered consumer advocate may simply be unable to represent the diverse constituencies and societal concerns that characterise vibrant and joyfully argumentative multiracial, multicultural and economically stratified societies such as Australia. And this is also why privacy impact assessments as conducted by professional professionals cannot simply expand to include evaluation as to social licence. Privacy impact can usually be objectively assessed using now relatively mature methodologies, structured broad enquiry and evaluation, but evaluation of social licence for data driven applications requires subjective balancing of competing perspectives that are assessed and weighed using immature frameworks and methodologies.

The problem becomes yet more complicated when we venture into evaluation of uses and applications of data, AI and machine learning applications.[8] It is difficult to empower

[7] Many organisations are considering proxies for broader consideration of data trust and trust in the form of statutorily appointed Privacy Commissioners, internal or external Ethical Review Committees, Risk and Audit Committees or external Governance groups. For discussion of such proxies, see Calo, R., Chizeck, H.J., Joh, E. & Hannaford, B., "Panel 2: Accountability for the Actions of Robots", (2018) 41 Seattle U. L. Rev. 1101-21 (2018); Calo, R. "Consumer Subject Review Boards: A Thought Experiment" 66 Stanford Law Review Online 97-102 (2013) and Polensky, J., Tene, O. & Jerome, J. (2015) "Beyond the Common Rule: Ethical Structures for Data Research in Non-Academic Settings", Colo. Tech. L.J. 13, 333. Some commentators, including Ryan Calo and Jules Polensky, advocate adaptation and extension to AI development of the Institutional Review Board/Human Ethics Review Board model as adopted in many countries. This review board model follows the Belmont ethical principles for research involving humans. The Belmont ethical principles were summarized in the Belmont Report and first published in the U.S. Federal Register in 1979, which were adopted by the U.S. National Commission for the Protection of Human Subjects of Biomedical & Behavioral Research. See the National Research Act (Pub. L. 93-348) and the Belmont Report "Ethical Principles and Guidelines for the Protection of Human Subjects of Research, Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research", 44 Fed. Reg. 76, 23192 (Apr. 18, 1979). In the United Kingdom, the Caldicott Committee's December 1997 "Report on the Review of Patient-Identifiable Information", usually referred to as the Caldicott Report (after its author Dame Fiona Caldicott), identified weaknesses in the way parts of NHS handled confidential patient data. The Caldicott Report highlighted key principles which have come to be known as 'the Caldicott principles' and which are now embodied in the NHS confidentiality code of practice. Follow up reports by Dame Fiona in 2012 and 2016 added a further principle and included many further recommendations. The 1997 Caldicott Report recommended appointment of "Caldicott guardians", members of staff of each NHS organisation with formally designated responsibility to ensure patient data is kept secure. It is now a requirement for every NHS organisation to have a Caldicott guardian. The Guardian is responsible for ensuring that their organisation adheres to the Caldicott principles see "UK Caldicott Guardian Council', available at https://www.gov.uk/government/groups/uk-caldicott-guardian-council, and the materials there referenced. The Information Security Management: NHS Code of Practice (at paragraph 31 and available at https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/codes-of-practice-for-handling-information-in-health-and-care/information-security-management-nhs-code-of-practice) also requires each organisation to designate a senior medical officer to oversee all procedures affecting access to personally identifiable health data. This role and that of the "guardian" may be combined, providing there is no conflict of interest. In 2017, the Australian Productivity Commission in 2017 considered criticisms of the role and functioning of Human Research Ethics Committees, in the context of considering how to encourage more effective sharing of health data and possible extension of ethics review to review of other data sets. See the Productivity Commission "Inquiry Report on Data Availability and Use" (No. 82, March 2017, available at https://www.pc.gov.au/inquiries/completed/data-access#report), in particular at section 3.3, (pp140-145), section 6.5 (pp272-278 and Appendix E: Case Study: Health Data (pp509-540). See also Lipworth, W & Kerridge, I. (2017) "The Future of Precision Medicine in Australia: Social and Ethical Implications of Precision Medicine" in the Horizon Scanning Series published by the Australian Council of Learned Academies' (ACOLAs) precision medicine project (https://acola.org.au/wp/pmed/) and available at https://acola.org.au/wp/wp-content/uploads/IP-4-Ethics.pdf. The NZ Data Futures Partnership commissioned a useful background paper on uses through sharing of health data: Moore, D. & Niemi, M. (2016) "The Sharing of Personal Health Data – A Review of the Literature", available at http://datafutures.co.nz/assets/Uploads/The-Sharing-of-Personal-Health-Data-Sapere-FINAL.pdf. See also Public Health Agency of Canada (March 2017) "Framework for Ethical Deliberation and Decision-making in Public Health", available at https://www.canada.ca/en/public-health/corporate/transparency/corporate-management-reporting/internal-audits/reports/framework-ethical-deliberation-decision-making.html; and Mittlestadt, B., "Designing the Health-Related Internet of Things: Ethical Principles and Guidelines" Information 2017, 8, 77

[8] Two leading reports by regulators are Commission Nationale Informatique & Libertes (CNIL, December 2017) "How can humans keep the upper hand? The ethical matters raised by algorithms and artificial intelligence", an English language version of which is available at https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf (and summarised at https://www.cnil.fr/en/how-can-humans-keep-upper-hand-report-ethical-matters-raised-algorithms-and-artificial-intelligence), and UK Information Commissioner's Office, "Big data, artificial intelligence, machine learning and data protection" as available at https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. For an interesting industry specific study, see (Germany) Federal Ministry of Transport and Digital Infrastructure's Ethics Commission (2017) "Report on automated and connected driving", available at www.bmvi.de/report-ethicscommission. See also the extensive work of the Information Accountability Foundation, including forthcoming "Fair and Just Analytics and AI, The Hong Kong Ethical Data Stewardship and Assessment Framework, Ethics Truly by Design" (as announced at https://informationaccountability.org/fair-and-just-analytics-and-ai-the-hong-kong-ethical-data-stewardship-and-assessment-framework-ethics-truly-by-design/); Information Accountability Foundation "Report for the Comprehensive Assessment Oversight Dialog: Canadian Ethical Data Review Boards Project" (March 2018) (available at http://informationaccountability.org/wp-content/uploads/Report-for-the-Comprehensive-Assessment-Oversight-Dialog-

dialogues as to unfamiliar and sometimes unknowable outputs and outcomes. At the stage of data analysis and search for correlations and inferences, evaluation as to social licence is speculative and must be qualified. Any impact assessment conduced at commencement of a research and development phase will usually requires iterative re-evaluation – and sometimes entirely new evaluation - once outputs from data discovery, research and development become clear. Once you know the possible outputs and their limitations, it is possible to engage stakeholders with a properly grounded and informed assessment of possible outcomes through deployment.

It should follow that social licence for data aggregation and linkage and search for useful correlations and inferences within a controlled and safeguarded data laboratory environment is likely to be much more accommodating than social licence for letting outputs out of the lab and into the wild of applications affecting humans. Some consultations with Australians appear to confirm a readiness to distinguish between what goes on in a lab under controlled conditions and what is allowed out of the lab. However, engagements as to possible 'data sharing' often conflate and therefore confuse controls and safeguards for the lab, and controls and safeguards as to what may be allowed out of the lab. Where statutes permit 'data sharing', or constituencies are canvassed to consider 'data sharing', many citizens rightly fear carte blanche aggregation of all identifiable data that bodies holds about relevant individuals, and carte blanche for outputs from the lab and outcomes that might affect individuals (at least, so long as those outputs and outcomes don't include data capable of association with other data for re-identification). What constituencies should be asked to consider is laboratory grade controlled data linkage of particular data sets using reliably pseudonymised data, with no data leaving the lab and with separate and reliable evaluation of insights, models and algorithms before they are allowed out of the lab and then whether they go back to the contributing bodies, or into the wild, and subject to what conditions. Where there is bundling or conflation of permissions as to what goes on in the lab and what leaves the lab, the discussion quickly becomes complex and confusing. Statements as to possible outputs and outcomes must inevitably be more vague, over-inclusive and speculative. At least some constituencies are likely to fear the worst possible outputs and outcomes.

In summary, engagements as to social licence:

- must be multi-faceted and cannot sit with a single source of truth,
- often best conducted in and for distinct phases (i.e. for data linkage enabling R&D, for outputs used only for an outcome being development of policy, and for outputs with applications that have a possible outcome of significant effects upon how individuals are evaluated or dealt with),
- are highly contextual, and accordingly best broken down into specific contexts of data application with articulation as to appropriate and demonstrably reliable and verifiable controls and safeguards which ensure that the context of evaluation is sufficiently understood by engaged stakeholders and their evaluation may therefore be less speculative,

Canadian-Ethical-Data-Review-Boards-Project.pdf); Abrams, M.J., Cullen, P., Goldstein. L, "Artificial Intelligence, Ethics and Enhanced Data Stewardship" (September 20, 2017) as available at http://informationaccountability.org/wp-content/uploads/Artificial-Intelligence-Ethics-and-Enhanced-Data-Stewardship.pdf. See also Aaron Fluitt (2018), "Report from the Georgetown Law Round Table on the Ethical Reuse of Data in a Machine Learning World", available at https://www.georgetowntech.org/news-fullposts/2018/3/24/march-26-2018-institute-publishes-report-from-roundtable-on-the-ethical-reuse-of-data-in-a-machine-learning-world; Broad, E., Smith, A. and Wells, P. (Open Data Institute, 2017), "Helping organisations navigate ethical concerns in their data practices" and associated materials at http://aims.fao.org/activity/blog/identify-and-manage-data-ethics-new-odi-approach-data-ethics-canvas.

- should never be on the basis of unsubstantiated assurances as to possible re-evaluation in the future, as distinct from commitments as to both controls and safeguards as to the lab and as to nothing coming out of the lab without appropriate separate evaluation of outputs, potential application of those outputs and possible outcomes upon individuals.

So let us now try to further unpack the rights and ethics underpinnings of social licence.

Sometimes compliance with existing laws does not confer any social licence. Civil disobedience may be condoned through social licence, or even required. Were Dr Martin Luther King, Vincent Lingiari[9] and Nelson Mandela each lawbreakers, or operators at the edge of social licence? Generally, acting to promote broadly accepted human rights is considered ethical and nurturing social licence.[10] And some statements of human rights, such as the International Covenant on Civil and Political Rights[11], are broadly accepted, even if not enacted in national laws. However, interpreting broad principles of human rights into the diverse minutiae of everyday social and business interactions can be difficult. And many assertions as to 'new' human rights are contested – to take but a few, a right to know, to be treated fairly, and not to be refused service by a robot.

So assertions as to contravention of 'rights' are often contestable. Do we do any better if we add a focus upon 'ethics' to consideration as to compliance with human rights?[12] To a degree, but only if we test whether an act or practice is ethical from a variety of viewpoints: in no particular order and without being exhaustive, from perspectives as to:

- fairness and justice,

- achievement of the common good,

- virtue,

- utilitarian ethics.[13]

[9] A brief introduction to Vincent Lingiari is available at Avani Dias (ABC News), "Activist captured 'original' 1975 photo of Whitlam pouring sand into Lingiari's hand" http://www.abc.net.au/news/2016-09-03/activist-took-original-gough-whitlam-vincent-lingiari-sand-photo/7805880; see further Ted Egan, Lingiari, Vincent (1919–1988) in Australian Dictionary of Biography, available at http://adb.anu.edu.au/biography/lingiari-vincent-14178.

[10] The Australian Human Rights Commission's ongoing major project on human rights and technology was launched in July 21068 with publication of an issues paper which includes useful discussion of application of principles of human rights to new technologies: see AHRC, Human Rights and Technology Issues Paper July 2018 available at https://tech.humanrights.gov.au/consultation; also Berkman Klein Center at Harvard University, "Artificial Intelligence & Human Rights: Opportunities & Risks", September 25, 2018, available at https://cyber.harvard.edu/sites/default/files/2018-09/2018-09_AIHumanRightsSmall.pdf? and Tene, O. Polonetsky, J. & Ahmad-Reza Sadeghi, A.-R, "Five Freedoms for the Homo Deus", IEEE Security & Privacy 16, 3 May/June 2018, pp15-17.

[11] Available at https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx. See also The United Nations Human Rights Office, Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework as available at https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf and the European Convention on Human Rights (in particular, Article 10) as available at https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx.

[12] As to data ethics, see the excellent introduction by Luciano Floridi and Mariarosaria Taddeo "What is data ethics?" and other articles in the theme issue entitled "The ethical impact of data science" of Philosophical Transactions of the Royal Society December 2016 vol 374 issue 2083, as available through http://rsta.royalsocietypublishing.org/content/374/2083. See Jake Metcalf and Kate Crawford, 2016 'Where are the Human Subjects in Big Data Research? The Emerging Ethics Divide,' Big Data & Society, special issue on Critical Data Studies, Spring 2016 available at http://bds.sagepub.com/content/3/1/2053951716650211.full.pdf+htm; Ifeoma Ajunwa, Kate Crawford and Joel Ford, 2016 "Health and Big Data: An Ethical Framework for Health Information Collection By Corporate Wellness Programs", Journal of Law, Medicine and Ethics, 44(2016).

[13] As to ethical decision making generally in business, see the excellent materials available at the website of the Markkula Center for Applied Ethics at Santa Clara University (at https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/) which consider basic ideas in applied ethics, such as utilitarianism, rights, justice, virtue, and the common good. The Makkula Framework for Ethical Decision Making (as available at https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/a-

This is complicated enough. But real life is worse. A commitment to an ethical approach is often asserted as a self-regulatory alternative to black letter law and can be no more than ethics-washing: cloaking bad behaviour with ethical self-justification. A provocative essay by Dr Ben Wagner of the Privacy & Sustainable Computing Lab at Vienna University examines the topic of "Ethics as an Escape from Regulation: From ethics washing to ethics-shopping"[14], and concludes that in order for approaches to technology design and development, whether ethical or human-rights based, to be taken seriously, they should at minimum meet the following basic criteria:

- early and regular engagement with all relevant stakeholders,
- a mechanism for external independent (not necessarily public) oversight,
- transparent decision-making procedures on why decisions were taken,
- a stable list of non-arbitrary standards where the selection of certain values, ethics and rights over others can be plausibly justified,
- ethics do not substitute for fundamental rights or human rights,
- a clear statement on the relationship between the commitments made and existing legal or regulatory frameworks, in particular on what happens when the two are in conflict."

When we focus upon specific attributes of the Fourth Industrial Revolution[15] that we are now living through – computationally inferred behaviours of humans using diverse data, algorithmic decision making and other applications of artificial intelligence – we enter unfamiliar ethical territory where accepted social mores provide little guide as to social licence. The basic criteria described above should be met, but more is necessary. Meeting these criteria makes it more likely that we will ask the right questions of the right stakeholders in a structured way, but do not provide us with all that we need to get to the right answers.

In my view, there are two important elements that are missing:

- a frame to assess social licence for a particular activity of a particular type of organisation, and
- a decision making methodology to be applied consistently and reliably in data governance and technology planning within an organisation.

framework-for-ethical-decision-making/) is extensively used and cited. A good example of application of the Framework to emerging technology is provided by a case study of as to application of the principles to government proposals to mandate access to encrypted communications: Raicu, I. "Balancing privacy and the needs of law enforcement" (December 2, 2016) at https://www.scu.edu/ethics/focus-areas/technology-ethics/resources/ethical-questions-about-encryption/.

[14] Wagner, B. (2018) "Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?", available at https://www.privacylab.at/ethics-as-an-escape-from-regulation-from-ethics-washing-to-ethics-shopping/ (accessed 2 October 2018)

[15] As to which, see World Economic Forum (2016) "The Fourth Industrial Revolution: what it means, how to respond" as available at https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/; and the projects of the World Economic Forum Centre for the Fourth Industrial Revolution; Kay Firth-Butterfield (Head, Artificial Intelligence and Machine Learning, World Economic Forum LLC) "To realise the full potential of AI, we must regulate it differently" (17 September 2018), available at https://www.weforum.org/agenda/2018/09/the-potential-and-pitfalls-of-ai-artificial-intelligence/; project reports of the World Economic Forum's Centre for the Fourth Industrial Revolution (as available through https://www.weforum.org/centre-for-the-fourth-industrial-revolution/areas-of-focus), and project reports of The Royal Society's Data and AI workstream (https://royalsociety.org/topics-policy/data-and-ai/).

I suggest that we need to start from a focussed examination of digital trust and then apply a diversity of ethical perspectives and legal analysis through good governance.

We should start by asking whether a particular act or practice (if and when exposed to scrutiny and debate) will maintain and not erode trust of a substantial majority of a relevant constituency - citizens, voters, consumers, customers or users – in how data relating to them is used, disclosed or applied.

If a proposed act or practice will erode trust, we should ask whether we should do it at all, or whether it can be done in a different way that would not be corrosive of trust.

Our consideration of a variety of ethical perspectives should start from a number of propositions:

- In the fields of information governance and governance of social analytics, relatively narrow coverage of data laws has left a lacuna, which has of itself contributed to erosion of digital trust.
- Services and technologies that today challenge information privacy are a relevant factor, but not the key factor, eroding digital trust. Citizens perceive that data privacy is challenged, in particular through pervasive tracking and near ubiquitous surveillance, but regulating tracking and surveillance will not of itself significantly improve digital trust.
- Digital trust is hard won and easily lost.
- Trust in data use is an important part of social licence. When people trust that their data will only be used as they understand and they consider that enough value will be created, they are likely to be more comfortable with its use.
- Value may be social benefit and does not necessarily imply financial reward or preference to the individual to whom the data relates.
- Failure to disclose possible detriment to the individual to whom the data relates is likely to be significantly corrosive of trust.
- Trust and fairness and often bound together, but in different ways in different social groups.
- A new element is becoming clear: fair disclosure as to derivation of value of digital data, and who derives that value, is increasingly relevant to trust.
- The issue as to fair sharing of value is often masked by a largely meaningless debate as to who owns data.[16] The ownership debate is meaningless because it is misframed. The more appropriate questions are: (1) who controls what uses of relevant data, (2) how that control is exercised, and (3) under what circumstances that control ends or ceases to be exercisable.

My basic proposition is that digital trust is not principally an ethical issue: it is also an outcome from properly aligning perceptions that the value derived through a particular data

[16] See further Peter Leonard (2018) "The Good Oil on 'the New Oil", available at https://www.scl.org/articles/10164-the-good-oil-on-the-new-oil; and Peter Leonard (2018) "AI Challenges and the Law: Being smart enough to boss around smart devices and AI", available at https://www.scl.org/articles/10295-ai-challenges-and-the-law-being-smart-enough-to-boss-around-smart-devices-and-ai.

use is mutually understood and reliably fair. Compare the continuing rise in the value of loyalty schemes, which reflects continuing willingness of users to ‘trade’ data within closed network loyalty schemes run by big trusted brands, with growing distrust of social networking services perceived (wrongly or rightly) as appropriating too much (or all) data value through complex or opaque practices.

Of course, distribution of data value is not the whole story. Erosion of digital trust of citizens has many sources. Important contributors are:

- failure by many digital entities to share benefits derived through data use and sharing with users,
- rapid growth of digital giants and data driven platforms and their market wealth and influence, which fuels resentment as to a growing digital divide within the Fourth Industrial Revolution,
- concerns as to pervasive tracking and surveillance,
- fallout from Cambridge Analytica and Facebook, after demonstrated failure by Facebook to proactively address clearly evident misuse of customer data by Cambridge Analytica,[17]
- failure by government agencies to transparently engage with citizens how data about them is curated and used by and between government agencies, coupled with stories around unaccountable algorithms, and (in Australia) realities of poor handling in switching of defaults in My Health Records and data breaches by Federal Government agencies,[18] and
- anxiety about technology continuing to destroy jobs, within limited opportunities for many workers to reskill within the digital economy.

And social licence is not political mandate. Political mandate remains the rule of the majority. Often now it is minorities that are excluded, or that perceive themselves as excluded, that loudly express views as to what is acceptable or not and thereby define social licence and levels of societal trust of digital technologies. Indeed, majority support for particular data practices may increase the sense of diverse minorities of disenfranchisement and therefore alienation. And that disenfranchisement is finding popular voice, in many economies, and eroding the mandate of mainstream political parties.

Sometimes minorities are right in feeling disenfranchised. Algorithmic decision making can work to their disadvantage, either by choice (suppliers excluding or discriminating against less attractive customers) or inadvertence or poor application (e.g. use of bad algorithms that are the output of poor statistical methods, particularly where derived from thin data at the edges of the societal Bell curve).[19]

[17] See further Peter Leonard and Toby Walsh (2018), “#MeToo for AI? Could Cambridge Analytica happen again?”, available at https://www.scl.org/articles/10210-metoo-for-ai-could-cambridge-analytica-happen-again

[18] For example, Matthew Beard (6 March 2017) “When it comes to trust, a good offence is your worst defence” The Mandarin (http://www.themandarin.com.au/76454-high-price-to-pay-to-correct-the-public-record/) and Anna Johnston (30 June 2017) “A litany of privacy disasters: how to ruin public faith in just 12 months” The Mandarin (http://www.themandarin.com.au/80791-litany-privacy-disasters-ruin-public-faith-just-12-months/).

[19] There are many analyses of actual and prospective unfair applications of algorithmic decision making, e.g. Cathy O’Neill (2016), Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy; Tene, O. & Polonetsky, J, (2017) “Taming the Golem: Challenges of Ethical Algorithmic Decision Making” N.C.J.L.&Tech 19,125; Crawford, K. & Schultz, J. (2014) “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harm”, 55 BCL. Rev 93;

How should these challenges inform development of public policy and decision-making by Government agencies?

Government agencies are affected by loss of trust in organisations and institutions, just like all other institutions. Trust is no longer about compliance with privacy and other laws. Commitment by Government agencies to demonstrate at least absence of improper purpose, and at best, fairness and impartiality in decision making and commitment to maximise public benefit, is no longer enough, in and of itself.

Social licence (citizen trust) is not demonstrated through process, or even transparency of process. In any event, Government agencies often erode public trust through poor process – impenetrable explanatory memoranda, self-justifying regulatory impact statements, unrealistically short consultation periods, set-piece enquiries dominated by well-resourced stakeholders, and so on.

And organs of Government generally should not seek to rely upon notice and/or consent as a basis for processing of personal information. Many citizen engagements with government agencies do not entail real choice. I am happily an Australian taxpayer and proud of many uses of Australian tax to improve civil society. However, I don't freely and voluntarily elect to pay Australian tax. By way of example only (the Australian Taxation Office (ATO) rightly does not do this), the ATO should not bundle any provision by me of tax information with "consent" to uses and disclosures of my tax information to other agencies in any way that affects the individual treatment by those other agencies of (individual) me, regardless of whether that tax information is passed over in a form which is personally identifying or not. And in circumstances where I do not have a choice as to my dealings with government agencies, I rely upon legislated or regulated control as to what the government agency may or must not do. If (to continue with my hypothetical example of the ATO, again with caveat that to my knowledge the ATO does not do this) wants do something with my tax information that is beyond calculating and collecting my tax and directly related secondary uses, or that affects the individual treatment by other agencies of (individual) me (as distinct from informing development of policy), the empowering statute (or transparent subordinate instrument) should specifically address what the ATO wants to do. The statute or instrument should only specify other uses and disclosures that affects the treatment of me (as distinct from informing development of policy) that are transparent and understandable, reasonable, proportionate and reasonably necessary.

Implementation of these principles should be axiomatic as a matter of good regulatory practice. Unfortunately, there are not implemented reliably and consistently. And when the legislature or Government organs don't comply with these principles, they erode digital trust as to what the Government, the Parliament and the bureaucracy are up to. If a citizen can't work out which body the citizen should hold responsible for eroding her or his trust, the citizen may mistrust all three.

So whether for good reasons of bad, Government agencies should no longer consider that they are immune from the discontent of many citizens with the professional political caste, with banks and insurers, and with other long perceived miscreants such as lawyers, property developers and robber barons (whether of the forest, railway or silicon variety).

Sloan, R.H. & Warner, R., "When Is an Algorithm Transparent? Predictive Analytics, Privacy, and Public Policy", IEEE Security & Privacy (16, 3, pp18-25, May/June 2018); Peter Leonard (2017), "Emerging Concerns for Responsible Data Analytics: Trust, Fairness, Transparency and Discrimination", available at https://www.commsalliance.com.au/__data/assets/pdf_file/0018/58104/Peter-Leonard-Emerging-Concerns-for-Responsible-Data-Analytics_-Trust-Fairness-Transparency-and-Discrimination-Paper-for-the-NSW-Data-Analytics-Centre-Showcase-12-Jul.pdf.

The fightback will be challenging. But for data governance, curation and sharing, the prescription is starting to be reasonably clear. Nurturing data trust requires a **FEAT**[20] of execution:

- **F**airness (including as to sharing of benefits),
- **E**thics,
- **A**ccountability and
- **T**ransparency.[21]

Social licence can be expressed as demonstrating to most affected individuals (not a mere majority) that a proposed data use will be to their benefit and can be trusted to remain to their benefit (that is,. without mission creep).

If a proposed data use cannot be fully transparent and explained so at least those that choose to be interested will understand and then carry the doubters with them, the proposed data use probably shouldn't be done.

Assessment of effect upon digital trust must be made as with transparent understandings as to:

- the purpose of collection of particular data sets,
- the purpose of sharing and linking;
- safeguards and controls protecting that sharing and linking data environment and what is allowed out of that data environment,
- permissible in-scope outputs (including applications through derived algorithms) and permissible outcomes achieved through use of such outputs,
- clearly excluded and never permissible outputs and outcomes.

In general, people tend to be more comfortable when data is used in context, as a necessary part of delivering a product or service.[22]

People expect to be clearly informed about the purpose of collecting data in specific and detailed terms, including a list of the data being collected, a description of any algorithms used, and disclosure of any possible future uses.

As the New Zealand Data Futures Partnership sees data social licence, for people to feel comfortable about a proposed data use, they first need good information on eight key questions that can be grouped under headings of **Value**, **Protection** and **Choice**.

[20] 'FEAT' is also 'FATE": see Microsoft Research, "FATE: Fairness, Accountability, Transparency, and Ethics in AI" at https://www.microsoft.com/en-us/research/group/fate/.

[21] These principles are analogous to the so-called "core principles of good corporate governance; fairness, accountability, responsibility and transparency", as developed from the U.K. Cadbury Report as released in 1991. See further The Committee on the Financial Aspects of Corporate Governance (chaired by Sir Adrian Cadbury) "Financial Aspects of Corporate Governance" available at http://cadbury.cjbs.archios.info/report.

[22] See the summary "What Do New Zealanders Think?" in Part 2 of New Zealand Data Futures Partnership "A Path to Social Licence, Guidelines for Trusted Data Use" (August 2017) as available at https://trusteddata.co.nz/wp-content/uploads/2017/08/Background-Trusted-Data.pdf.

Value

1. What will my data be used for?

2. What are the benefits and who will benefit?

3. Who will be using my data?

Protection

4. Is my data secure?

5. Will my data be anonymous?

6. Can I see and correct data about me?

Choice

7. Will I be asked for consent?

8. Could my data be sold?[23]

Some of these concepts are embedded in existing Australian data privacy laws, with the important qualification that in general these laws only address collection, uses and disclosures of personal information and not uses of non-personal information (such as through algorithmic guided decision making) that may have significant legal effects upon individuals.

This important gap is leading some commentators, including me, to advocate both:

- creation of new rights, such as a right to know (of transparency), a right to (only) reasonable inferences[24], and a right of human intervention, that would supplement and partially fill in gaps in coverage of existing data privacy rights and the limited human rights-based legal protections against discrimination,

- new requirements of ethical oversight or review, such as expanding the current requirements to conduct (1) privacy impact assessments for perceived high risk uses of personal information, and (2) independent ethics review board assessment of research projects involving humans or animals, and about humans as conducted with public funding, to include (3) ethics and fairness assessment of proposed applications of data in ways that could have significant impacts upon rights, benefits or entitlements of individuals either individually or in segments as compared to other individuals.

Unfortunately there is little reason to be confident that organisations will voluntarily offer full explanations as to the process, justification and accordance of high impact data driven and algorithmic decision making unless legally compelled to do so.

[23] As above.

[24] See the excellent analysis in Wachter, S. and Mittlestadt, B (2018) “A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI", Columbia Business Law Review, forthcoming; also Wachter, S., also Mittelstadt, B.D.M. and Floridi, L. (2017) "Transparent, explainable, and accountable AI for robotics", Science Robotics. 2 (6) eaan6080; Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. and Floridi, L. (2017) "Artificial Intelligence and the 'Good Society': the US, EU, and UK approach", Science and engineering ethics. 24 (2) 505-528; Wachter, S., Mittelstadt, B.D.M. and Russell, C. (2017) "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR", Harvard Journal of Law and Technology; Wachter, S., Mittelstadt, B. and Floridi, L. (2016) "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation", International Data Privacy Law 7(2) 76-99.

Nor is there reason to believe that organisations will be able to unravel high level statements of ethical principles into concrete frameworks, governance forums, processes and methodologies for making fair and ethically based decisions within those organisations.

Organisations need guidance and tools to aid them. Guidance and tools are emerging.[25] However, disproportionate effort is being placed into creating what is now a plethora of statements of ethical principles for algorithmic applications and AI. Not nearly enough effort is going into 'making it real', regardless of whether the 'it' is framed as digital trust, ethics, fairness, social licence, social capital or human rights. We need to refocus upon developing practical tools and methodologies for use in our workplaces today to ensure that we appropriately assess outputs and outcomes from computationally inferred behaviours of humans using diverse data, algorithmic decision making and other applications of artificial intelligence.

Peter G Leonard BEc(Hons) LLM (Syd)
pleonard@datasynergies.com.au

[25] Probably the most comprehensive effort to translating principles into practice is the ongoing work of the various workstreams of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems under the program "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems" (EADv2): see https://ethicsinaction.ieee.org/. For decision tools, see for example the Makkula app at https://www.scu.edu/ethics-app/; the "Ethics & Algorithms Toolkit" available at http://ethicstoolkit.ai/; the "Ethics checklist for data scientists" available at http://deon.drivendata.org/; the Open Data Institute's "Data Ethics Canvas" and associated materials at https://theodi.org/article/data-ethics-canvas/; Winnipeg Regional Health Authority (2015) "Ethical Decision Making Framework: Evidence Informed Practice Tool" and "Ethical Decision-Making Framework Workbook available at http://www.mb-phen.ca/files/EthicsEIPT.pdf and http://www.wrha.mb.ca/extranet/eipt/files/EIPT-037-002.pdf ; and the UK Government's "Data Ethics Framework" and associated guidance notes and workbooks as updated in August 2018 and available at https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework and https://www.gov.uk/government/publications/data-ethics-workbook.