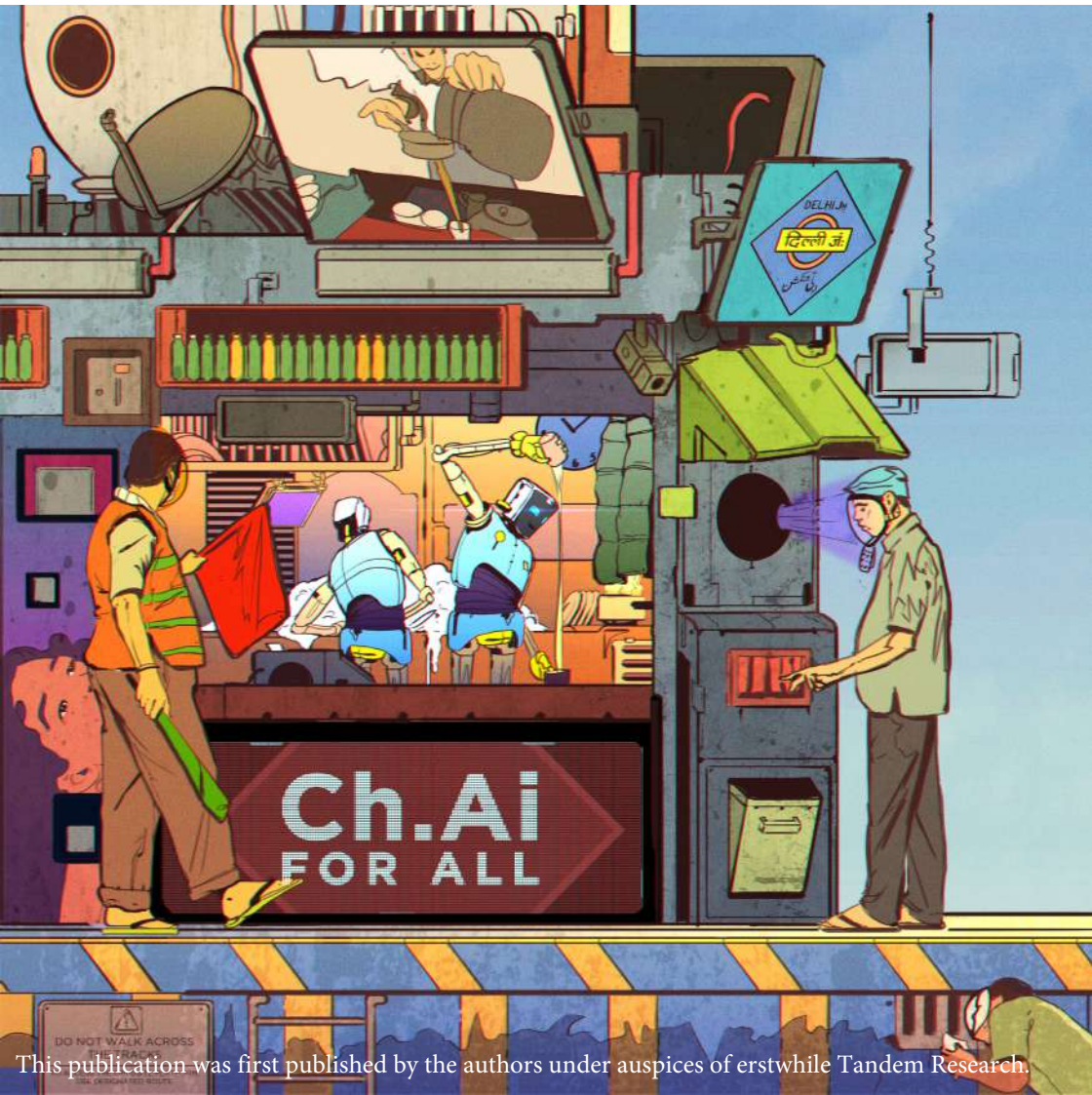
Working Paper

AI for All : 10 SOCIAL CONUNDRUMS FOR INDIA

Tandem Research 2018



This publication was first published by the authors under auspices of erstwhile Tandem Research.

This paper was made possible by the insights and contributions shared at at the first meeting of Tandem Research's AI & Society Technology Foresight Group, July 2018. Participant bios included in Annex.

A big thank you to all!

Tandem Research's Technology Foresight Group (TFG) brings together multiple stakeholders to collectively and iteratively diagnose issues and challenges pertinent to technology and society futures in India.

The brief is based on discussions of the TFG, but should not be seen as a consensus document — participant views differ and this document need not reflect the views of all participants.

Written by: Urvashi Aneja, Vikrom Mathur & Abishek Reddy
Cover Illustration: Yuvraj Singh

Please cite as: Tandem Research, *AI for All: 10 Social Conundrums for India,* Tandem Research - FES, 2018.

TANDEM
RESEARCH

AI for All

10 Social Conundrums for India

Working Paper

December 2018

With support from

FRIEDRICH EBERT STIFTUNG

CONTENTS

I. AI for All

Artificial Intelligence (AI) has been a subject of much speculation and debate since the 1950s, when Alan Turing asked if machines could think.[1] However, the development and deployment of AI has proliferated only recently, enabled by access to vast amounts of data, a massive increase in computational power, and better algorithms. From automating warfare to composing art, AI has the potential to radically transform society. The nearly boundless promise of efficiency and productivity gains, along with new forms of value creation, has focused the attention of technology companies and policy makers on AI development.

Yet, AI is not just a new frontier for innovation and technology; its social dimensions and implications are even more complex. The development and deployment of AI is likely to be a thoroughly social affair, shaped not only by technological possibilities but also an interplay of power, interests, values, and user behaviours. AI, thus, needs to be conceived of as a 'socio-technical system', as a system that does not function autonomously but is the outcome of the activities of human actors, and which encompasses the production, diffusion, and use of technology.[2] We need to think both technically and socially; technical feasibility and social viability are deeply interlinked.

---

1 Turing, Alan M. (1950). Computing machinery and intelligence. _Mind_ 59 (October): 433-60.

2 Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems. Research Policy, 33(6-7), 897–920.

The term 'artificial intelligence' does not lend itself to a simple, straightforward definition, at least partly because of the growing hype around it and the resultant tendency to describe various data-driven applications or algorithmic decision-making processes as AI. AI generally refers to the the use of digital technologies to create systems that are capable of performing tasks thought to require intelligence. Machine learning is a technique or sub-system of AI, whereby digital systems can improve their performance on a given task over time through experience.[3] 'General artificial intelligence' still remains a thing of the distant future. 'Narrow AI', involving sophisticated pattern recognition across multiple data points to generate probabilistic models and correlations, is already ubiquitous across multiple spheres of life— from algorithms that filter out spam to those that increase the accuracy of detecting cancerous growths.

The NITI Aayog's discussion paper on India's national AI strategy, titled 'AI for All', seeks to position India as a global AI leader by promoting AI solutions for healthcare, education, agriculture, mobility and smart cities.[4] While the paper hits many of the right notes in suggesting the importance of societal objectives in steering technological trajectories, it arguably overplays the potential of AI and underplays the challenges and risks entailed. This is particularly concerning

---

[3] Miles Brundage et al. ' The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation', Future of Humanity Institute, University of Oxford, and Partners. February 2018.

[4] Niti Aayog (2018) National Strategy for Artificial Intelligence - AI for All. Discussion Paper

in light of the suggestions for creating a data marketplace and positioning India as a 'garage' for testing AI solutions applicable to the developing world.

AI strategy and policy ultimately presents a 'wicked problem' for public policy. Wicked problems are those that have multiple interacting systems— social, ecological, and economic — a number of social and institutional uncertainties, and imperfect knowledge. Possible solutions to existing problems create a new set of additional challenges and the choice between available alternatives are often largely about competing values. BW Head describes wicked problems as representing a confluence of three factors: complexity of subsystems and interdependencies; uncertainty regarding risks and consequences of interventions; and a divergence or fragmentation in values, viewpoints and strategic intentions.[5]

From development and deployment to policy and strategy, conversations around AI reflect these complexities. The AI universe represents a complex set of interlocking sub-systems and issue areas, from the ownership and governance of data to the materiality of AI systems. At the heart of many debates in many, or each, of these subsystems are asymmetries in power and information; the coalition of particular interests and their ability to rally action and opinion; and value judgements about what constitutes a good and fair society and respective entitlements within it. Risks are emerging and consequences

[5] Head, B. W. (2008). "Wicked problems in public policy." Public Policy 3(2): 101. See also: Camillus, J. C. (2008). "Strategy as a wicked problem." Harvard business review 86(5): 98.

are unknown or not fully understood. Yet, the dangers of technological and policy lock-in are real, as are the risks of exploitation and misuse.

Addressing wicked problems requires engaging multiple stakeholders in iterative and adaptive strategies; enabling collaborative sense-making, experimentation, and learning; and building capacities for reflexiveness and foresight. This paper take the first step in the direction of developing 'a capacity for reflexiveness' by engaging a range of actors through a Technology Foresight Group (TFG) in an in-depth diagnosis of the social conundrums pertaining to AI trajectories in India. This brief presents 10 social conundrums for AI trajectories in India, arising from the inherent wickedness of AI futures. The large-scale deployment of AI technologies is still at an early stage in India, and impact is hard to identify and assess. Yet, the range of potential social conundrums need to be identified early and contextualised to the Indian context to be able to generate anticipatory knowledge about plausible and preferable future policy pathways.

II. 10 Social Conundrums for India

1. Reconciling multiple, competing social narratives

Multiple and often competing narratives on AI are emerging as new knowledge is refracted through the mindsets and social frames of various social groups. In science fiction and pop- culture, public discourse on AI has long oscillated between a narrative of progress and of moral panic, a utopian vision of AI saving the world, or a dystopian vision in which it outsmarts and takes over human civilisation.

Much of the contemporary discourse focuses on the impact of AI on jobs and the future of work.[6] While policy conversations are focussed on identifying relevant coping strategies, particularly through re-skilling, in other corners, a narrative of freedom and liberation is being articulated - while jobs will be lost, this will create more time for creativity and leisure - humans may finally have the freedom they desire![7] Yet, this remains a distant dream for many millions across the world, particularly in India, for whom work is necessary to survive; while proposals for a universal basic income and other redistributive mechanisms have been proposed, these are deeply contested, both in terms of principle as well as the practice and its likely impact.

---

[6] See for example: World Bank (2019) The Changing Nature of Work. World Development Report

[7] See for example: Peter Frase (2016), *Four Futures: Life after Capitalism*, Verso Books.

The potential for AI in warfare is setting in motion an 'AI arms race' among global powers. A 'Winner Takes All' frame is increasingly visible in the national strategies of a number of states. China and the US are in the lead, having invested billions of dollars towards AI research. Vladimir Putin recently claimed, for example, that 'Artificial intelligence is the future… Whoever becomes the leader in this sphere will become the leader of the world.'[8] With large technology companies at the forefront of much AI innovation, creating an enabling environment for the private sector is increasingly seen as a key strategy for winning this race. For the business community, appropriating AI solutions, even if only in name, is increasingly seen as critical to maintaining market competitiveness.

In the past few years, the framing of 'AI for Good' has captured the imaginations of policy makers and technology companies alike. AI is being framed as a silver bullet that can address persistent socio-economic challenges, for the benefit of society at large.[9] This imaginary is already propelling significant investments in health, education, agriculture and urban city management systems. A sense of AI solutionism seems to be

---

[8] Putin says the nation that leads in AI 'will be the ruler of the world ...." 4 Sep. 2017, https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world. Accessed 30 Oct. 2018.

[9] "Could AI Solve the World's Biggest Problems? - MIT Technology Review." 12 Jan. 2016, https://www.technologyreview.com/s/545416/could-ai-solve-the-worlds-biggest-problems/. Accessed 30 Oct. 2018.

driving much research and innovation — with an attempt at match-making between AI-based interventions and social challenges.

Lawyers, academics and activists alike, continue to raise concerns about the threats to civil liberties and the scope for discrimination, misuse, and new unknown risks. They point out that narratives are rarely neutral— they obfuscate certain beliefs and interests, while promoting others. For governments and industry, the narrative of 'AI for Good' helps invisibilize many of the dangerous effects of AI— from surveillance and warfare to new and unknown risks. Opposing or questioning AI is seen as stifling innovation, and now, with an 'AI for Good' framing, as standing in the way of development. Many are concerned about the sense of inevitability that characterises contemporary AI discourses— AI is akin to a discourse of modernity, of which we are all part, whether we like it or not.

2. Concentration & Collusion of Power

The materiality of the AI universe is often overlooked— whether in terms of the menial labor[10] required to train AI systems or the natural resources required to build specific technological devices.[11] Yet, these are a clear and stark illustration of the vast asymmetries in power and wealth that

---

[10] Lilly Irani (2013)," The Cultural Work of Microwork", *New Media & Society*, 0:(0 1-21

[11] see for example: Anders SG Andrae (2017), Total Consumer Power Consumption Forecast, Presentation at Nordic Digital Business Summit, Helsinki, Finland, October 5, 2017

underlie the AI universe. A recent study estimates that it will take 700,000 years for a child working in a cobalt mine in Bolivia[12] to earn what Jeff Bezos, founder and CEO of Amazon, earned in a single day in the first quarter of 2018.[13] A few 'superstar' global technology companies have access to a majority of global digital data, creating new data oligopolies that are being used to reshape behaviours and preferences, disrupting the workings of governments, markets, and communities, to benefit only a few.[14] Whether it's Facebook targeting depressed teens or Cambridge Analytica manipulating elections, these examples show how the interests of those deploying advanced data systems can overshadow public interest, acting in ways that are contrary to individual autonomy and collective welfare, often working in ways that are invisible and unquantifiable.[15] Further, as AI converges with the Internet of Things and the Internet of Living Things, the power of companies is likely to increase at an unprecedented rate, often without informed consent and adequate data security mechanisms.

---

[12] cobalt critical

[13] Kate Crawford and Vladan Joler, "Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources," AI Now Institute and Share Lab, (September 7, 2018) https://anatomyof.ai

[14] Stucke, Maurice E., "Here Are All the Reasons It's a Bad Idea to Let a Few Tech Companies ...." 27 Mar. 2018, https://hbr.org/2018/03/here-are-all-the-reasons-its-a-bad-idea-to-let-a-few-tech-companies-monopolize-our-data. Accessed 30 Oct. 2018.

[15] Foer, Franklin. World Without Mind: *The Existential Threat of Big Tech.* , 2018. Print.

Recent policy announcements in India suggest some attempt to reign in global superstar technology companies through provisions for data localization; a common argument of government officials — both in rhetoric and in law — is that localisation will help Indian law enforcement access data.[16] Yet, this could create new domestic oligopolies— crowding out many smaller players in the Indian market and privileging powerful incumbent players.[17] Many in civil society also note the growing nexus between state power and corporate interests, particularly technology companies. They warn that as government bodies are still grappling with understanding AI technologies and the way they work, this has given greater space for technology companies to influence public policy decision-making. They worry that, in this paradigm, citizens will have nothing to offer but their data— citizens will be reduced to a form of digital labor.

3. Work, Mobility and Digital Labor

The deployment of machine learning technologies will reduce the need for low-medium skill labor, and increase demand for higher-order skills. This could pose multiple and contradictory challenges for India. On one hand, high skill labor constitutes a very small portion of the population— most workers are low to medium-skilled and thus are at higher risk of technological unemployment or displacement. Many of these low-medium

---

[16] Chapter VIII. The Personal Data Protection Bill, 2018. MeitY, 2018. http://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf. Accessed 30 Oct. 2018

[17] Mihir Sharma, 'How Data Localisation limits possibilities for India's startups, consumer', *Business Standard*, 14 September 2018

skilled jobs, like in call centres or retail outlets, are what many youth working in informal employment within the unorganised sector aspire towards. An important opportunity for upward social-economic mobility may thus shrink as businesses adjust to new technological possibilities.[18]

Yet, a new class of low-skill jobs are also being created to fuel an AI world— in particular, data annotators that can label and sort data sets needed to train AI systems. This 'invisible work' or 'artificial artificial intelligence'[19] that is the foundation of high-tech AI applications is often underpaid, characterised by poor employment conditions, and entirely disconnected or removed from the artefacts it creates.[20] India could go from being the back-end for global business processing to being the back-end data annotator of the world. Already, Indian workers are one of the largest contributors to online micro-work platforms, such as Amazon Mechanical Turk.[21]

AI is also drastically changing ways of work, and hiring and firing practices. A recent study showed for example, that AI technologies are likely to be used by a majority of American firms in the next decade, as a way of monitoring and

---

[18] Tandem Research. 2018. Emerging Technologies & The Future of Work in India Goa: Tandem Research

[19] Mary L. Grat & Siddharth Suri, "The Humans working behind the AI curtain", *Harvard Business Review*, 9 January 2017

[20] Hope Reese and Nick Heath, "Inside Amazon's clickworker platform: How half a million people are being pennies to train AI'. *Tech Republic*

[21] Neha Gupta, David Martin, Benjamin V. Hanrahan, and Jackie O'Neil, "Turk-Life in India", *Group'14*, November 9-12, 2014.

improving worker productivity.[22] Workplace surveillance could be a serious concern in India, particularly where job competition is high; labor rights are poorly understood; and conversations about data privacy are at a nascent stage. Further, the use of AI for hiring new candidates and measuring their performance could improve firm-level productivity, but constrain opportunities for upward socio-economic mobility and challenge civil liberties.

4. Social / Digital Identity

The deployment of AI technologies risk entrenching old inequities, while creating new ones. In the Indian context, inequities exist across multiple dimensions beyond economic wealth— identity and caste are equally, if not more important markers. A digital identity risks limiting people's capacity to move across these identity markers— for marginalised and oppressed social groups, identity fluidity can be critical for both physical safety and social mobility.[23] Further, individuals have more than one identity, and intersecting identities imply that an individual can simultaneously belong to more than one group.

AI is being prescribed as a tool to enable the delivery of more efficient and scalable government welfare services, replacing existing human intermediaries with automated systems. This requires the creation of an authenticable digital identity that

---

[22] Romy Ribitzky, "Active Monitoring of Employees Rises to 78%", ABC News, 18 April 2018

[23] Judith A. Howard, "Social Psychology of Identities", *Annual Review of Sociology*, 26:2000

can be recognised across multiple interconnected networks; but this also presents a reductionist approach to identity, reducing subjectives to measurable and distinct categories. Further, removing human intermediaries can result in the loss of localized awareness and sensitivity to problems and issues, dismantling existing systems of kinship and patronage. Admittedly, these systems are often exploitative, and create their own sets of winners and losers; yet, this is precisely what renders AI as a wicked problem, whereby a new set of challenges are created in the process of addressing older ones. Further, there are instances in which digital identity is taking precedence over a physical identity, as seen in cases of the denial of welfare services to individuals without an authenticated Aadhar number.[24]

5. Biased Robots

Non-representative or biased data can further entrench existing inequities, as AI systems reproduce the representation gaps and biases of the data sets on which they are trained.[25] Data can be seen through multiple frames: the frame of the uncounted (those who don't exist because they are not included in any sort of database), unaccounted (the portrayals of people with less inclusion into the digital world and therefore not entirely represented, maybe due to economic reasons) and discounted (they exist and are in the system but are not of interest to the people who would serve them such

---

[24] Nikhil Dey and Aruna Roy, "Excluded by Aadhaar," The Indian Express, June 5, 2017

[25] EUBANKS, VIRGINIA. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* S.l.: PICADOR, 2019. Print.

as governments or companies because they do not have enough money). Representation within AI datasets is likely to mirror digital divides. Access is intersectional, especially in rural areas where digital technologies like smartphones are shared by multiple users. Policies hinging on such data would have a distorted picture of social reality and often blindside women and young children.

Data is expensive and hard to come by at scale; AI training relies on available data sets, rather than complete data sets. This type of data can easily privilege socio-economically advantaged populations, those with greater access to connected devices and online services. In India, less than 30 percent of India's internet users are women and only 14 percent of women in rural Indian own a mobile phone.[26] Existing data sets in India, whether for labor markets or health records, are often fragmented, outdated, or unrepresentative.[27]

The impact of such data bias can be seriously damaging in India, particularly at a time of growing social fragmentation. It can contribute to the entrenchment of social bias and discriminatory practices, while rendering them invisible and pervasive through the AI systems. For example, historically certain communities were forced towards thievery due to caste discrimination and were labeled as born criminals.

---

[26] LIRNEasia. (2018). AfterAccess: ICT access and use in Asia and the Global South (Version 1). Colombo: LIRNEasia

[27] Samarth Bansal, "From missing data to unreliable numbers, India's statistical ecosystem needs an overhaul', *Hindustan Times*, 21 September 2017.

Alienation and stereotyping of these communities due the historical association still continues today by the police and media.[28] The AI trained on this historic data is likely to view people of these communities as thieves even if none of them currently continue to thieve. According to a 2014 report, Muslims, Dalits, and tribals make up 53 percent of all prisoners in India; National Crime Record Bureau data from 2016 shows in some states, the percentage of Muslims in the incarcerated population was almost thrice the percentage of Muslims in the overall population.[29] If AI applications for law and order and the delivery of social justice and welfare systems are built on this data, it is not unlikely that they will be prejudiced against these groups.

6. Concentration of Knowledge (and Power)

The increasing ubiquity of AI systems controlled by a few can also subjugate certain forms of knowledge, while creating new dependencies on technological applications. Companies and governments deploy AI systems as superior sources of credible information and insights, advocating unquestioned adherence to AI recommendations. This can end up displacing existing and hybrid knowledge systems, who are now confronted with AI outputs as objective truths. This risks creating a situation in

---

[28] "How Denotified Tribes In India Face Discrimination ... - Youth Ki Awaaz." 2 Aug. 2016,https://www.youthkiawaaz.com/2016/08/denotified-tribes-discrimination-and-violence/. Accessed 30 Oct. 2018.

[29] "Crime In India 2016, NCRB - National Crime Records Bureau." 10 Oct. 2017,http://ncrb.gov.in/StatPublications/CII/CII2016/pdfs/NEWPDFs/Crime%20in%20India%20-%202016%20Complete%20PDF%20291117.pdf. Accessed 30 Oct. 2018.

which entire sectors and underlying knowledge systems are reconfigured in their entirety. For instance, the merger of Bayer and Monsanto can be seen as an attempt to fuse complementary data sets for soils and seeds— they can then flood the market with subsidised seeds and recommend best practices to farmers through AI analysis of the soil and seed data, arguably pushing a dependency model with the company's best interest at heart. It is believed this alliance can corner as much as 61% of the global seed and pesticide markets, having serious implications on the local autonomy of the farmer.[30]

7. Privacy frameworks: Unfit for purpose?

From recent controversies around the use and misuse of Aadhar data, to targeted social media messaging to influence electoral outcomes, data privacy is already an urgent concern in India. AI systems depend on gathering the maximum amount of available data and drawing correlations across a disparate and often unknowable set of data points. In this sense, AI technologies are fundamentally at odds with current privacy frameworks based on the idea of consent and data minimisation.[31] The Sri Krishna report, which articulates a draft data protection framework for India, hinges on the idea

---

[30] IT for Change, ' Data Frameworks for a Right to Development' in UNRISD, *From Disruption to Transformation? Linking Technology and Human Rights for Sustainable Development*, 2018.

[31] "How Companies Learn Your Secrets - The New York Times." 16 Feb. 2012, https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html. Accessed 30 Oct. 2018.

of consent.[32] Meaningful consent is hard enough to ascertain in India, where low levels of education and awareness hinder capacities to exercise informed choice. Moreover, with AI systems, data is used, shared, and made sense of in ways that cannot even be imagined. How can consent be given in such a context? In the case of misuse or manipulation, can consent be taken back? Is this even technically possible— when individual data is being fed into complex and deeply layered algorithms, where the ways in which data is correlated remains unknown and unpredictable?

Further, even the anonymization of personal data may not be adequate— by triangulating between multiple data points, re-identification of individuals can be possible. With new IoT home appliances that record energy usage, or the use of AI for urban planning through the deployment of IoT devices, patterns about an individual lifestyle or movements can be ascertained.[33] Individuals may also choose not to share certain data about themselves, but data shared by a larger group or collective can still affect the particular individual's agency. In other words, even digitally excluded or disconnected will be impacted by the deployment of AI systems in spheres of social life.

Paradoxically, India is the one of the most connected countries in the world in terms of the number of people online, but

---

[32] Chapter III. The Personal Data Protection Bill, 2018. MeitY, 2018. http://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf. Accessed 30 Oct. 2018.

[33] "Smart Cities May Be The Death of Privacy As We Know It – Futurism." 7 Nov. 2017, https://futurism.com/privacy-smart-cities. Accessed 30 Oct. 2018.

perhaps one of the least connected in terms of percentages. The poorest of poor constitute a large part of the unconnected population among which women are often the least connected— 88% of rural Indian women are not digitally-connected.[34] A large proportion of India's population can thus be thought of as digitally-excluded but data-included. AI systems that are deployed, especially by the government, impact the entire population immaterial of whether they are a part of the decision-making process or even if they are digitally-included.

Framing data privacy policies hinges on the understanding of how personal data is viewed, as a right or property. If treated as a right, regulation needs to specify what data can be collected and traded and what data can not. While treating it as property enables commodification of data, the collecting and trading of which is then under the control of the individual. The former holds the government responsible for protecting privacy while the latter pushes the onus on the individual, but, the individual can profit from their data. The private sector obviously favours the latter. Another option is licensing or conditional ownership, where the individual claims stake in the product or purpose for which the individual's data is used.

[34] Osama Manzar, "Rural India: Living Under Digital Exclusion", NDTV, 6 January 2017

8. Reporting Citizens / Surveillance State

A number of states in Asia, including India, are investing in mass surveillance systems— from facial recognition technologies to social media analysis cells. Government agencies are already using automated tools to allocate resources and monitor people. This raises significant concerns about civil rights and liberties.[35] Contemporary AI systems intensify practices of surveillance systems, which require the collection of massive amounts of data. Marginalized communities and populations already subjected to disproportionate government scrutiny will bear the brunt of these new surveillance technologies.[36]

Some voices within civil society suggest that India risks resembling characteristics of a surveillance state; the government is a major stakeholder in the current process of data collection, and individuals are expected to be reporting citizens with little option but to comply. The state now also has an interest in data that would not otherwise be relevant, but which will now be collected since AI can process this information. With the Collection of Statistics Act, 2008,[37] refusing to part with information or providing inaccurate data to the state a punishable crime. The draft Data Protection Bill,

---

[35] "The Big Eye: The tech is all ready for mass surveillance in India ...." 13 Aug. 2018, https://factordaily.com/face-recognition-mass-surveillance-in-india/. Accessed 30 Oct. 2018.

[36] Eubanks, Virginia. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. S.l.: PICADOR, 2019. Print.

[37] The Collection of Statistics Act, 2008. https://indiacode.nic.in/bitstream/123456789/2081/1/A2009-07.pdf. Accessed 30 Oct. 2018.

submitted by Srikrishna Committee, also gives the state the power to access and process personal information for reasons of national security. The bills further proposes data localization which requires companies to store all personal data in India, which the government can then access citing national security as a reason, furthering concerns among civil society about the surveillance capacities of the state.[38]

9. What / How to Govern

As AI systems are deployed across a number of socio-political domains, the transparency and accountability of these systems is becoming an urgent concern. One one hand, the outcomes of deep learning are unpredictable and unknowable at the outset. This renders ineffective many existing frameworks for accountability. How can the the unpredictable and invisible be governed?

Data used for AI-based public service delivery has been asked to be auditable, to check for biases, and provide iteratively improved solutions. Some suggest that peer review processes need to be put in place, whereby AI algorithms are reviewed by neutral peers.[39] Another essential check suggested is

---

[38] Chapter VIII. The Personal Data Protection Bill, 2018. MeitY, 2018. http://meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf. Accessed 30 Oct. 2018.

[39] "Cognitive Bias in Machine Learning – The Data Lab – Medium." 17 Aug. 2018, https://medium.com/ibm-watson-data-lab/cognitive-bias-in-machine-learning-d287838eeb4b; see also: Mittelstadt, Brent. "Automation, Algorithms, and Politics| Auditing for Transparency in Content Personalization Systems." International Journal of Communication [Online], 10 (2016): 12. Web. 30 Oct. 2018

output analysis to ensure the output and working of the AI systems is fair, unbiased, inclusive and non-exploitative. Independent state regulatory bodies might be instated to verify, test and approve of AI algorithms before being deployed in the market— similar to the FDA's role in the pharmaceutical industry. Requiring an explanation or interpretation of AI and machine learning systems might also serve as an accountability measure.[40]

There is the critical question of who should be held accountable— the developer, the designer, the deployer, or governments? Among technology companies at the forefront of AI development, the conversation about accountability tends to be framed in the language of ethics. Yet, ethics as a code is not legally enforceable and is a fuzzy point of discussion that can be agreed or disagreed with; it is a convenient way for technology companies to claim they are self-regulating. Ethics, however, cannot be substituted for legal responses to harm.

While AI development is still at a nascent phase, so are understandings of the risks and unknowns around AI. We need to ask the purpose of development and deploying certain technologies, else we risk embarking on technological trajectories that will soon outpace society's capacity for control. This question of social and technological choice must

---

[40] Knight, Will. "There's A Big Problem With AI: Even Its Creators Can'T Explain How It Works". MIT Technology Review, 2018, https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/. Accessed 12 Nov 2018.

figure in conversations around accountability.[41] Accordingly, new interdisciplinary knowledge clusters comprised of political scientists, technologists, sociologists, anthropologists, and lawyers, among others, are a need of the hour. Community-based visioning and public engagement around AI should be a normative process and there is a need to push the state towards providing platforms for constructive public engagement. The Telecom Regulatory Authority of India's (TRAI) platform for citizens to voice opinions over net neutrality is a good antecedent example. Expert activism through evidence gathering and demonstration of alternative trajectories is required, along with the traditional policy consultation methods, particularly to counter the current grip of technocrats on the decision-making process. Public engagement initiatives and public art can also help increase awareness by unveiling the cloak of invisibility that usually envelops issues around AI.

10. Reconfiguring human agency

The focus in many conversations about AI is on the potential range of solutions and innovations it can deliver. AI has become an industry in itself to promote, rather than one among many tools that can be used to serve socially-identified goals. But, AI technologies also risk transforming what it means to be human, shaping not just human behaviour, but also desires and preferences, and now with the emergence of

---

[41] Wendall Wallach, *A Dangerous Master: How to Technology from slipping beyond our control*, Basic Books, 2015.

biotechnology, the fundamental biological building blocks of humans.

Technological advances in genomics and synthetic biology are increasingly converging with automation, artificial intelligence, and cloud computing. What if humans were no longer required to perform the analysis, writing, and editing of DNA?[42] The merging of information technology with biotechnology will hit at the core of what it means to be human, to have the capacity for free-will and independent decision-making.[43]

This suggests that we need to look beyond the applications of AI to the kind of relationship we have with AI. This may enable putting humans at the centre of the conversation, their desire and needs, rather than just the technological possibilities and limitations of AI. Data, for example, is often treated as a disembodied subject, dissociated from individuals.[44] Yet, data is ultimately about the lives of real people, their needs, preferences, and beliefs. The notion of 'augmentation' is increasingly being enrolled to address concerns about AI taking over human jobs, or humanity more generally; yet again, the focus with 'augmentation' is on the capacities of AI rather than human needs or societal priorities. Manual scavenging, for example, continues as a regular

---

42 Garrett Dunlap and Eleonore Pauwels, "The Intelligent and Connected Bio-Labs of the Future: Promise and Peril in the Fourth Industrial Revolution", Wilson Briefs, September 2017.

43 Yuval Noah Harari, 21 Lessons for the 21st century, Random House 2018

44 see for example: Martin Dodge & Rob Kitchin, " The Ethics of Forgetting in an Age of Pervasive Computing", CASA Working Paper Series, No. 92, 2005.

practice in India, claiming over 200 lives in the past year alone; yet, investments towards automating such dangerous, demeaning, and dirty tasks, are few and far between.

With growing evidence of the harmful impacts of AI, technologists and others have advanced the idea of 'parenting AI' i.e. more time and investments are needed to teach the AI, to eliminate bias and other negative consequences.[45] But, who should take on the burden of this parenting; when has it been trained adequately; and who will suffer in the meantime? Already vulnerable and marginalised communities are likely to bear the burden and dangers of un-parented AI, reducing the incentives for timely and rigorous parenting.

[45] See for example: Lila Tretikov, " We should not raise AIs like parents, not programmers - or they'll turn into terrible toddlers", Quartz, 5 January 2018

III. Building a Capacity of Reflexiveness

Responsible Research and Innovation

The policy discourse on 'responsible research and innovation' has emerged over the last decade in Europe and elsewhere as policy makers and policy analysts grapple with the special challenge of regulating technologies characterized by both technical and social uncertainty, and complexity and ambiguity; these technologies require fundamentally different policy-making processes and approaches.[46] These conversations have their roots in earlier discussions about the ethical and social implications of areas such as nanotechnology and genomics, and call for the need to address the social and ethical dimensions of technology and innovation early on. The framework of anticipatory governance[47] emphasizes the need for deliberation on the social conundrums of technology at an early stage of the policy conversation.

Various forms of public engagement and open opportunities for deliberation can help build a 'capacity for reflexiveness' in science and technology institutions and decision-making

---

[46] R.Owen, P. Macnaghthen and J. Stilgoe, " Responsible Research and Innovation - from Science in Society to Science for Society", *Science and Public Policy* , 39 (2012)

[47] Risto Karinen and David H Guston, "Towards Anticipatory Governance: The Experience with Nanotechnology", in M. Kaiser et al., *Governing Future Technologies: Nanotechnology and the Rise of an Assessment Regime*, Springer 2010

processes.[48] Technology assessment in 'real-time' is imagined in contrast to the linear model of innovation policy; and an awareness and consideration for stakeholder values, aspirations and risk perceptions rather than management of just 'technical' risks.

A framework for reflexivity for AI development in India would need to have five key elements:

1. ***Thinking Both Technically and Socially:*** The dominant narrative around AI in India frames the development and use of AI for social good, or otherwise, as a technical problem— to be addressed through the creation of better innovation and startup ecosystems and investments in technical manpower. A range of complex social concerns— access, equity, privacy and power — are acknowledged, for example in the governments AI strategy, but not addressed and often brushed under the carpet. Policy makers and policy analysts will need to pay attention to both the social and technological issues around AI and support better socio-technical integration.

2. ***Anticipation and Futuring:*** Extant policy approaches are limited in managing emerging and socially-challenging technologies such as AI because of the imperative to demonstrate short term policy impact and success. AI is not likely to be a silver bullet, which

[48] *Guston, D. H., & Sarewitz*, D. (2002). Real-time technology assessment. Technology in Society, 24(1-2), 93-109

could be deployed instrumentally to solve problems: long-term strategies will be needed to navigate the complex social and technical challenges to applying AI to address persistently insoluble challenges like health and education. Reflexivity about social conundrums around AI will be needed to identify the diverse plausible trajectories in order to formulate policies to navigate towards desirable socio-technical futures.

3. ***Knowledge Systems:*** While the temples of technological research in India— the IITs and IIITs— are gearing up to develop AI applications, social science research and knowledge systems around AI are less developed and attract limited funding support from either the public or private sector. More investment and capacity is needed for research on complex social and ethical issues around AI development, to a point where they can be meaningfully assessed before wide-scale deployment.

4. ***Policy Portfolio and experiments*:** Responsible research and innovation on AI will require strategies and policies that cut across traditional decision-making silos in India. AI is being primed for use in agriculture, education and health– but any applications will need to be built upon an understanding of both the needs of the AI 'users' in these sectors and the institutional and policy context of the challenges that the sectors face. The NITI Aayog has proposed 'proof of concept' pilots for AI applications across sectors, but untested technical fixes will not address the policy failures that

plague these sectors.

5. ***Deliberative decision making*:** Globally, many models of technology assessment based on public engagement, participation and deliberation— for example, through consensus conferences— are being tested.[49] There also is growing evidence that in the case of ethically problematic and socially complex technologies, early engagement with the public, on the wider dimensions of the technology, leads to more socially acceptable design and development. Widely-debated norms and general governance principles around development and deployment of emerging technologies are likely to be more widely-accepted. In a democratic political context, like India, deliberative process can lead to more dynamic technology strategies and can help avoid static regulatory and legal approaches to technology control.

The complex social dimensions around the research, development and deployment of AI need to understood, rather than brushed aside. Innovation and regulatory frameworks will need to co-evolve. This should be designed to happen in such a way that there is an opportunity to progressively strengthen mutual expectations and collaboration, and buy-in of all the parties involved: government, industry, academia, civil society organisations and wider publics. A continuous interaction between

[49] Simon Joss and John Durant eds. *Public Participation in Science: The Role of Consensus Conferences in Europe*, Science Museum, 1995.

research, experimental action, regulation and assessment of this kind, within a framework to which public deliberation will make a central contribution, will support AI trajectories that align with societal goals.

ANNEX

Technology Foresight Group

Tandem Research's *Technology and Society* research initiative unpacks the social dimensions of technology trajectories in India, to identify relevant knowledge and policy needed to align technological trajectories with societal goals. We seek to look beyond universalizing narratives on technology to, instead, interrogate specific local interactions and impact.

In partnership with the Friedrich-Ebert-Stiftung, Tandem Research is convening a set of six meetings of the Technology Foresight Group (TFG) to collectively evaluate the political, social, and value-based contestation underlying the framing of problems and solutions around AI futures in India. The TFG brings together multiple stakeholders for collaborative and iterative public policy solutions. Previous meetings of the TFG have focused on the Future of Work in India.

The aim of the first meeting, on 24 & 25 July 2018 in Aldona, Goa, is to identify the key narratives, social frames, actors and institutions shaping AI trajectories in India, and the unique opportunities and challenges facing India. Subsequent labs will adopt a sector-specific approach, looking at the role of AI in health, education, agriculture, environment and government.

Anita Gurumurthy is a founding member and executive director of IT for Change, where she leads research

collaborations and projects in relation to the network society, with a focus on governance, democracy and gender justice. Her work reflects a keen interest in southern frameworks and the political economy of internet governance and data and surveillance. Anita is part of the Feminist Alliance for Rights, a global advocacy platform that emerged during the SDGs process and on the International Steering Committee of the Global Alliance on Media and Gender. She also serves as a member of the Science and Technology Constituency of the Asia Pacific Regional CSO Mechanism.

Anja Kovacs directs the Internet Democracy Project in Delhi, India, which works for an Internet that supports free speech, democracy and social justice in India and beyond. Anja's research and advocacy focuses especially on questions regarding freedom of expression, cybersecurity and the architecture of Internet governance. She has been a member of the of the Investment Committee of the Digital Defenders Partnership and of the Steering Committee of Best Bits, a global network of civil society members. She has also worked as an international consultant on Internet issues, including for the Independent Commission on Multilateralism, the United Nations Development Programme Asia Pacific and the UN Special Rapporteur on Freedom of Expression, Mr. Frank La Rue, as well as having been a Fellow at the Centre for Internet and Society in Bangalore, India.

Arnab Kumar leads NITI Aayog's initiatives in Emerging Technologies (Artificial Intelligence, FinTech, BlockChain etc.). He is also the co-author for NITI Aayog's "National Strategy for Artificial Intelligence" discussion paper and is leading several

proof-of-concept projects in AI and BlockChain viz. advanced farming advisory using AI and satellite data, land records and fertilizer subsidy using BlockChain and Imaging BioBank for Cancer. Arnab is also the Founding Manager for Atal Innovation Mission (AIM), Government of India's flagship initiative for innovation and entrepreneurship, housed at NITI Aayog. At AIM, Arnab manages INR 200 crores early-stage fund for incentivising development and commercialization of tech products. An ex-Investment Banker, Arnab had spent more than 5 years working for Deutsche Bank's New York, Hong Kong, Singapore and Mumbai desks and had executed more than USD20bn of equity, M&A and debt transactions, prior to joining NITI Aayog in 2017. Arnab is an alumnus of Indian School of Business, Columbia Business School, BIT Mesra and is also a Chartered Financial Analyst.

Baneen Karachiwala is a Public Health Researcher providing support to health projects across India, and globally. She has 13 years of experience, with a focus in the areas of strategy, knowledge management, implementation research, photojournalism and documentation. Her special interests lie in the areas of human rights, maternal and child health. She is currently based in Bangalore and is affiliated with the Ramalingaswami Centre of Equity and Social Determinants of Health, which is part of the Public Health Foundation of India, Delhi. She also is a Senior Advisor to a human rights and community media organisation, Video Volunteers based in Goa.

Chinmayi Arun founded and led the Centre for Communication Governance, the only academic research

centre in India that works on information law and policy, at NLU Delhi for five years.She is a member of the United Nations Global Pulse Data Privacy Advisory Group, and of UNESCO India's Media Freedom committee. She is also a Fellow of the Berkman Klein Centre at Harvard University, and has been a consultant to the Law Commission of India and the Indian government's multi-stakeholder advisory group for the India Internet Governance Forum in the past. Chinmayi has published work on algorithms, harmful speech, privacy and information gatekeepers. She is lead author of the India country report in Freedom House's Freedom on the Net report for 2014 and 2015, and of the India report in the Global Network of Centers' study of online intermediaries.

Daniel Buhr is Head of the Steinbeis Transfer Center for Social and Technological Innovation and Associate Professor of Policy Analysis and Political Economy at the Faculty of Economics and Social Sciences of the Eberhard Karls University Tübingen in Germany. He conducts transdisciplinary research and teaches at the interface between politics and economics with a special focus on economic and innovation policy (i.e. Industrie 4.0, IoT) as well as on social, health and care policy. On behalf of the Ministry of Social Affairs and Integration he represents the State of Baden-Württemberg at CORAL (Community of Regions for Assisted Living) and serves on various committees and expert task forces.

Deepak Khemani is a professor in the Department of Computer Science and Engineering, IIT Madras, India. He graduated with three degrees from IIT Bombay, including two

in computer science.His areas of research are broadly in artificial intelligence, with focus on memory based reasoning, knowledge representation, automated planning, constraint satisfaction, and qualitative reasoning. He has done several sponsored projects applying case based reasoning to build corporate memory systems; looking at ontologies in the domain of manufacturing; and building automated planning based applications for space and defense research.

Mahesh Venkateswaran is Chief Growth Officer at Knack (knack.it), a behavioural AI company that uses a combination of neuroscience, machine learning and gamification to identify natural talents of individuals. Knack is used by enterprises, educational institutions, Governments and social impact programs in 120+ countries. Mahesh leads growth and expansion at Knack and is remotely connected from Delhi, India. Trained as an engineer, Mahesh has worked in areas broadly spanning research, consulting, social impact, Government and technology. Prior to Knack, he was part of the leadership team at NSDC for 4.5 years.

Malavika Jayaram is the inaugural Executive Director of the Digital Asia Hub. Prior to her relocation to Hong Kong, she spent three years as a Fellow at the Berkman Klein Center for Internet & Society at Harvard University, focused on privacy, identity, biometrics and data ethics, and eight years in London, with the global law firm Allen & Overy in the Communications, Media & Technology group and as Vice President and Technology Counsel at Citigroup. While a partner at Jayaram & Jayaram in India, she was one of 10 Indian lawyers selected for The International Who's Who of

Internet e-Commerce & Data Protection Lawyers directory for 2012 and 2013. In August 2013, she was voted one of India's leading lawyers – one of only 8 women to be featured in the "40 under 45" survey conducted by Law Business Research, London.

Naveen Thayyil is a faculty at the Department of Humanities and Social Sciences at the Indian Institute of Technology, Delhi. His research and teaching interests lie in the interstices of three domains - legal and democratic theory, environmental law, and STS (Science and Technology studies). Naveen is interested in understanding law and its connection to democratic normativities, and the role of development and deployment of technology here. His interests lie not only at the level of public policy viz., issues of regulation of technology for the protection of public health, environment and related rights that seek to democratise society, but also in theorising and understanding how categories of law, technology and society shape each other.

Osama Manzar is a global leader on the mission of eradicating information poverty from India and global south using digital tools. He is a social entrepreneur, author, columnist, impact speaker, angel investor, mentor, and sits on several government and policy committees in India and on international organisations working in the areas of Internet, access, and digital inclusion. With over 20 years experience, Osama has worked in the areas of journalism, new media, software enterprise and created the Digital Empowerment Foundation that works in India to digitally empower the

masses with a footprint of 200 locations and interventions in more than 10 countries, mostly in South Asia.

Satish Sangameswaran is Principal Program Manager at Microsoft Research India and is based out of their India research lab in Bangalore. He has over two decades of experience in the technology sector with earlier stints in Intel, Compaq, etc. apart from Microsoft. HIs charter includes managing Outreach initiatives for Microsoft Research India, with a focus on creating and nurturing partnerships for collaborative research across academia, industry, government and the social sector. Satish is a believer in the potential of technology to create a better world.

Sean Blagsvedt is the Founder and former CEO of Babajob.com, India's largest digital job marketplace for the aspiring labor, with over 8.5 million users and acquired by Quikr in June 2017. Currently, Sean is a Principal at Jaaga.in, Chairman at TheADAO.org and advises Unifize, Harambee and Quikr. Sean moved to Bangalore in 2004 as the 3rd founding member of Microsoft Research India and earlier was a Program Manager in the UX teams of Office and Windows. He also was a White House intern with President Clinton's Internet Policy Czar, Ira Magaziner.

Sumandro Chattapadhyaya is a Research Director at the Centre for Internet and Society (CIS), India, where he leads the Researchers at Work programme. He undertakes and supervises academic, policy, and applied research on a series of topics including open data and open research, e-governance and digital ID, digital media and humanities, and emerging

digital economies. He is a Co-investigator at the Development Implications of Digital Economies (DIODE) Strategic Research Network; and is a co-founder of India Open Data Association and a member of DataMeet.

Usha Ramanathan works on the jurisprudence of law, poverty and rights. She researches, writes and speaks on issues that include the nature of law, Bhopal Gas Disaster, mass displacement, eminent domain, manual scavenging, civil liberties including the death penalty, beggary, criminal law, custodial institutions, the environment, judicial process. She has been tracking, and engaging with, the UID project and has written, and debated extensively, on the subject. She was a member of the Expert Group on Privacy set up in the Planning Commission of India which gave in its report in October 2012; a committee (2013-14) set up in the Department of Biotechnology to review the Draft Human DNA Profiling Bill 2012; the Committee set up by the Prime Minister's Office (2013-14) to study the socio-economic status of tribal communities which gave its report to the government in 2014.

Patrick Ruether is Country Representative & Head of Office of FES in India. Among his key portfolios are feminism and digitalisation. Previously to his posting in India, Patrick was in charge of the FES business network in Berlin. During this time he focused on the social inclusion and social innovations in the context of the Industry 4.0 and digitalisation. Patrick is also an expert on the European Union having lived in Brussels for a couple of years. Now he enjoys his life in Delhi since 2015.

Sehaj Malik is Advisor for the Socio-Economic Transformation program at FES. She leads the Foundation's work on policy, strategy, and liaison within the socio-economic sphere in India. Prior to FES, she worked with the Ministry of Finance and the UNDP on political bargaining and international negotiations within the BRICS and G20 forums. Sehaj has previously worked as a macroeconomist in Roubini Global Economics in London and Bruegel in Brussels.

Urvashi Aneja is Founding Director of Tandem Research. She works on the governance and sociology of emerging technology; southern partnerships for humanitarian and development assistance; and the power and politics of global civil society. Urvashi is also Associate Fellow at Chatham House. She has a PhD from the Department of Politics & International Relations, University of Oxford.

Vikrom Mathur is the Founding Director of Tandem Research. Vikrom is an anthropologist of science and technology. His diverse research interests include the governance of emerging technologies, social and cultural dimensions of technological transitions, political and social contingencies on the production of scientific knowledge about Nature, cultural perceptions of environmental risk, dynamics between science and policy, and Cultural Theory. He has a PhD from the Institute of Science, Society, and Innovation at the University of Oxford. Vikrom is a Senior Fellow of the Observer Research Foundation & Associate Fellow of the Stockholm Environment Institute.

Tandem Research is an interdisciplinary research collective generating policy insights at the interface of technology, society, and sustainability. We believe in finding iterative solutions to real world problems through evidence based enquiry and public engagement.
tandemresearch.org

The Friedrich-Ebert-Stiftung (FES) is a non-profit German foundation committed to the values of democracy and social justice. Widely acknowledged by the German and Indian governments for being an important actor in the promotion of dialogue, FES has been active in India since the late 1970s and established its official country office in 1983 in New Delhi.
fes-india.org