Book review: Lives of data: essays on computational cultures from India

Lives of Data is an intellectual response to the birth and evolution of data societies as instruments of knowledge and everyday infrastructures. The book, in a pioneering effort, charts the development and immersion of data infrastructures across socio-economic and institutional contexts in India through the articulation of social, political, economic, cultural worlds – and "data worlds". Lives of data offers a unique contemporary snapshot of what is at once India's data world and the digital every day. It does the above by connecting historical journeys of small and big data, their embodiments in model building and in their impacts on human socio-economic sustenance.

Editor Sandeep Mertia focusses on "relationality" and "lineages, affinities, and relations of data in context-sensitive ways" to centre and ground questions and debates in the 14 essays that make up this volume. Beginning the collection of essays with a historical detour of computing practices through an interrogation of the Indian Statistical Institute and Mahalanobis, the theme of "relationality" explores interconnected social constructions of the "possibilities" of computing technology, organisational cultures, everyday practices and management of professional labour. This introductory chapter is also a plea to situate history and past visions of computing practices in specific contexts of knowledge production. The second chapter is also a historical segue into "vernacular histories of computation and data" of an advertising distribution network since the 1950s. The generation and processing of data by Blaze advertising illumine backstage cultures of cinema advertising and the chaotic hybrid nature of data formats and institutional cultures they are embedded in. Author Mendonca defines the "object of Software" beyond a techno-linguistic categorisation to the myriad forms of data generation in distributive networks of an advertisement agency. A third historical essay concerns the semiotic valency of data models, in this case the Duckworth–Lewis–Stern data model in the game of cricket and their transformational impacts in the thinking of cricketing communities. Author Arumugam argues that the success of the DLS model is also the success of probabilistic/predictive data models capturing the imagination of a cricket-watching public and investing value in a culture of knowing through numbers. The semiotic valency of data resides in the fairness it can produce in predictions by simultaneously eradicating unfairness and inaccuracy in model-building.

A discussion of Aadhaar technologies in chapters 4, 7 and 8 focus on capturing, storing and ultimately creating the Indian citizen primarily through digital data infrastructures. Author Ranjit Singh alludes to "thinning" of individual identity to fit data capture and processing model and to the dictates of both machine readability and the state's organisational practice. Achieving thinness of data is directly attributable to the predefined rules of data record keeping and consequences for an individual bearing the Aadhaar-driven identity. Singh also draws attention to data infrastructures as a dynamic shifting entity as they layer and "imbricate" over time – especially in the infrastructural, administrative and social practices establishing an Aadhaar-vetted identity. Ways to study data trajectories within a socio-technical system is to study "moments of data imbrication" within the system. Aadhaar is primarily seen as an epicentre of tensions between a voluntary and a compulsory national identity scheme and entanglements between the legislature and judiciary bodies of the Indian state, arising out of implementing Aadhaar as constitutive of a citizen identity. The shifting meanings of Aadhaar as a socio-technical system and process

not only politicise but also change conditions surrounding the procurement of a national ID. Another instance of Aadhaar as an unstable e-infrastructure is its stand in for everyday gateways to several essential mandatory services and its operationalisation as a "genuine material" infrastructure. The latter leads to failures in everyday data protocols and human lives dependent on such data – what author Mudliar refers as the chasm between "broken" data and broken bodies – How can broken data be repaired and dealt with humanely? Interestingly Chapter 9 on the building of data through field work addresses the above problem of broken data and bodies. The chapter primarily offers qualitative data as a key missing link in support of policy building and the challenges thereof field researchers face in documenting, processing and translating large-scale data collection. Whilst the essay elaborates upon tools for rigorous, accurate field data collection, it suggests technological solutions to circumvent challenges in field data collection, processing and validation. Data sources are also brought into question wherein corruption of data occurs at the stage of collection even before the socio-technical system can indulge in acts of commission and omission in building data infrastructures.

Chapter 5 undertakes an epistemological turn in the enquiry of data as an object of academic interrogation and the disaggregation and re-contextualising of literary objects as digital media objects. The digital archive, the lab, the processes of curation and re-representing of cultural artefacts can now be de-constructed and rendered for re-reading in digital formats. Does interpretation of texts change as methods to annotate technologise? Can understanding a film change because methods to de-construct them change? These questions ponder the relationship between the object and method and implication for "lives of data".

Lily Irani's chapter on hackathon as data in action raises questions that have plagued disciplines like anthropology, particularly applied domains of knowledge, about scholarship and the ethics of intervention. Hackathons are viewed as labour bringing data to life, invoking passion in technocrats to care for and create software-to-come. Chapter 6 essentially looks at hackathon processes as extremely political, "a model for making Indians into entrepreneurial citizens". Hackathons are viewed as aggregating creative crafty labour in the hope for software ingenuity and excellence coupled with the dream of future investments for improved career-life chances and securing a future extending the life of complex networks of digital technologies.

Chapters 10 and 11 deal with affordances and challenges of making data open and by an extension that is accessible. Two things seem to have implications on openness of data – first are tools that bring data into existence and the second is politics. At the level of tools that bring data to life, organisational, editorial and summarisation apparatuses may pose technical challenges to making data open – a simple example being the endurance of the PDF as a data repository format carrying the inscription of built-in authorial power. Speaking about the development sector, author Narula identifies challenges in the initial stages of sourcing/collecting data and the process of this data being "crunched, dissected, and sometimes tortured" to produce a report. Openness is to make the above processes visible, accessible and traceable. Author Godhwani speaks of open documents as moral documents like government budgets that are moral and value laden because of their direct impacts on lives of citizens. The author argues that data are built on moral premise and rest on collective and social values. By this token, the life of data in documents and reports can have moral intentions! Chapter 13 leads to data that are bureaucratic and untidy awaiting not only a cleaning but also a cleansing process. Author Solanki, speaking about Data Nagar, a government information management system that deals with the vexed edict of knowledge equals data in the invoking of life in data regimes! The untidiness of data is not only evident in the spread sheets hosting them but also in the earlier stages of collection and recording of data – what Solanki calls "data life-sources".

Talking of data formats, Chapter 12 and author Ravel speak about the everyday data keeping or accounting, taking hybrid material data formats and the persistence of materiality. From cabbies in the UK remembering thousands of street names, which is a kind of materiality turning redundant, Uber driver's *Hisab/Kitab*, account housekeeping and of income data on paper, there is a stubborn co-existence of local forms of data formats. Lives of data take multiple forms of co-existence!

Lives of data refer to multiple active and dynamic processes by which data are cleaned, cleansed, trusted, possessed and presented. The rendering of data is a socio-technical process, by which data are imagined, politicised, moralised, trusted and live in a relationship with a variety of data producers and analysts. This volume of essays raises intriguing methodological questions approaching data as a cultural-economic object – we get to know data as inter-disciplinary, socio-technical, moral, political, biased, alive and to simultaneously study [or research] data as commodity and as a powerful relational entity. Production and processing of data in the hands of the nation-state or multinational corporates for end goals must invoke a variety of approaches and methodological tools. The methodological question in the life of data is still an inconclusive one.

The book pitches "data" as a powerful inclusionary apparatus pre-empting a means and conditions of production debate – for example, situating the production of fake news as a democratically generated circulation of content or a top-down invasion of free speech. Additionally, the book provokes questions of fluidity, transformability, performativity or contradictory aspects of an object or a person belonging to the scheme of invoking lives of data as "identity". I believe the historical sensibility and flavour of the book are keys to unravel several unsettled debates on the nature of data worlds – they chart a journey, map a trend and sight aspects of future data worlds.

Nimmi Rangaswamy
Kohli Centre on Intelligent Systems, IIIT, Hyderabad, India