

From Victorian Statistics to Big Data: Reading *Hard Times* through Critical Data Studies

Deepak

M.A (English), NET

Mail Id: deepakboora107@gmail.com

Abstract

This paper aims to analyze the growing interdisciplinary field of Critical Data Studies (CDS) to read Dickens's novel *Hard Times* (1854) and show how the novel anticipates contemporary critiques of big data culture, algorithmic government and data-driven decision-making by analyzing its critique of statistical rationalism, utilitarian quantification and the reduction of human experience to numerical data. In this study, based on the theorizing of other CDS scholars like Johanna Drucker, Safiya Noble, Cathy O'Neil and Danah Boyd the ongoing critique of political economy and the epistemological constraints and political implications of data-driven knowledge and social organisation are anticipated throughout the novel and in Dickens's representation of Bounderby's statistical management of labour and Gradgrind's fact-based education system.

Keywords: Critical Data Studies, *Hard Times*, Charles Dickens, big data, Victorian statistics, utilitarianism, data culture, algorithmic governance, quantification, political economy

1. Introduction

Charles Dickens's *Hard Times* (1854) is a novel about data. The novel highlights the utilitarian philosophy of its time-believing that knowledge can only be classified into facts, human experience can only be measured, and social issues can be resolved by the application of statistical reasoning. Thomas Gradgrind's first line is a famous one:

'Now, what I want is, Facts. Teach these boys and girls nothing but Facts. Facts alone are wanted in life. Plant nothing else, and root out everything else. You can only form the minds of reasoning animals upon Facts: nothing else will ever be of any service to them. This is the principle on which

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

I bring up my own children, and this is the principle on which I bring up these children. Stick to the Facts, sir!’ (Ch. 1)

One of the easier-to-recognize expressions of data rationalism in Victorian literature is this declaration of faith in the power of numerical data to capture and manage the complexity of human existence. In this modern day and age of "big data" – the data that is gathered, analyzed and mobilized to make decisions in areas ranging from health to criminal justice to education – the questions that Dickens raises in *Hard Times* have taken on new urgency. As in Victorian political economy, the idea of using data to create objective, value-free information about the world, the more the better, is at the heart of modern data culture.

This paper interprets *Hard Times* as a critique of statistical rationalism, and shows how the novel's criticism prefigures some of the more modern critiques of big data culture. The study traces the historical development of data rationality from the Victorian statistical movement to the present day of big data, and explores the links between the Victorian statistical practice and the current

data culture through the lens of the novel *Hard Times*.

2. Theoretical Framework: Critical Data Studies

Critical Data Studies (CDS) is an interdisciplinary approach to studying the social, political, cultural and ethical aspects of data practices. CDS emerged in the 2010s and modifies the notion that data would be neutral, accurate reflections of reality, saying that data are always created in particular social, institutional, and political contexts that have an impact on what is counted, how it is counted, and how the outcomes are understood.

Theorizing Interpretation (2014) and Graphesis (2014) by Johanna Drucker are the two theoretical pillars of CDS. Drucker is making the point that data is not ‘given’ but ‘captured’: that data production is always a process of interpretation that is shaped by the assumption, values, and priorities of the data producer(s). Her idea of ‘capta’ (from the Latin *capere*, to take), highlights the artefactual aspects of data, and thus the idea

that data are necessarily neutral representations of presenting reality is called into question.

Algorithms of Oppression (2018) by Safiya Noble and Weapons of Math Destruction (2016) by Cathy O'Neil explore how algorithms and data-driven decision-making in modern societies impact politics. As Noble illustrates, search algorithms can replicate racial and gender biases, and as O'Neil reveals, algorithmic models in education, criminal justice, and employment can perpetuate inequality while posing as being mathematical.

Six Provocations for Big Data Research (Boyd and Crawford, 2012) and The Data Revolution (Kitchin, 2014) are important texts for thinking about the epistemological aspects of data culture. Boyd and Crawford say that big data isn't very different from other data; they're just produced, curated, and interpreted by humans using certain assumptions and priorities.

3. The Victorian Statistical Movement and *Hard Times*

The period which saw *Hard Times* written was one of great interest to statistical knowledge and its application to social issues. The Victorian statistical movement was a part of a wider campaign to employ scientific approaches to the study of society, which occurred in the 1820s and 1830s. William Farr, Edwin Chadwick and Adolphe Quetelet were among the pioneers in utilizing statistical information to examine public health, crime, poverty, and education.

The statistical movement was linked with the utilitarian philosophy of Jeremy Bentham and James Mill, who believed that the aim of social policy should be the greatest happiness of the greatest number, and that this can be done by the rational calculation of costs and benefits.

The novel's depiction of Coketown is a reflection of Dickens' distrust of statistics as a means to capture the complexity of human experience:

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

It was a town of red brick, or of brick that would have been red if the smoke and ashes had allowed it; but as matters stood it was a town of unnatural red and black like the painted face of a savage. It was a town of machinery and tall chimneys, out of which interminable serpents of smoke trailed themselves for ever and ever, and never got uncoiled. It had a black canal in it, and a river that ran purple with ill-smelling dye, and vast piles of building full of windows where there was a rattling and a trembling all day long. (Ch. 5)

The uniformity of Coketown, its regularity and its abnegation of human life to mechanical processes are all stressed. The ‘interminable serpents of smoke’ that ‘never got uncoiled’ are the inescapable, machine-like rationality of industrial production, which dehumanizes human beings as human data within a system of production and consumption.

3.1. Gradgrind’s System: Data as Epistemology

Data rationalism is most clearly expressed in Gradgrind's philosophy of education, which he states: “All children should be taught nothing but facts.”

Gradgrind's school becomes a factory for manufacturing facts, and the children are seen as ‘little vessels’ to be ‘filled with facts’:

The room was a plain, bare, monotonous vault of a schoolroom, and the speaker’s square forefinger emphasized his observations by underscoring every sentence with a line on the schoolmaster’s sleeve. The emphasis was helped by the speaker’s square wall of a forehead, which had his eyebrows for its base, while his eyes found commodious cellars in two dark caves, overshadowed by the wall. The speaker’s voice was inflexible, dry, and dictatorial. The speaker’s hair, which bristled on the skirts of his bald head, a plantation of firs to keep the wind from its shining surface. (Ch. 1)

The description of Gradgrind himself – “his square wall of a forehead”, “his square forefinger”, “his inflexible, dry, and dictatorial” voice – is an embodiment of the rationalism of the data. His body is stiff and stiff is his system of knowledge.

CDS can be seen as an epistemology of data, a belief system that holds that data are not just used to understand the world, but are the source of all legitimate knowledge, this is Gradgrind's system. Gradgrind's

philosophy is that facts are truths, that facts speak for themselves, and that facts are the key to understanding.

The novel's critique of Gradgrind's epistemology is played out in the narrative structure: a structure that on a regular basis illuminates the limits of factual knowledge to understanding human experience. Louisa is unable to express her own emotions, a sign of the failure of data-driven education to be present in Gradgrind –

'What do I know, father, of what is beautiful, of what is expressive, of what is graceful or ungraceful, of what is right or wrong-' 'My dear Louisa, you are tired.' 'Tired, father?' 'Of what?' 'I don't know of what-of everything, I think.' (Ch. 15)

Louisa's reply, 'I don't know of what-of everything, I think' is the essence of a data-driven education which is missing the skills to support its students in understanding their experience. The novel in the end illustrates the nature of data as argued by CDS scholars: incomplete, partial, and always the product of an interpretive frame.

4. Bounderby's Statistics: Data as Power

The data-rationality of Joseph Bounderby, wealthy manufacturer and Gradgrind's friend and ally, is the application of data rationality to the management of labour and justification of economic inequality. The Bounderby self-made man story, his argument that he had gone from rags to riches via hard work and determination is a statistical narrative, a story told in numbers:

'I was born in a ditch, and my mother ran away from me. I was brought up by my grandmother, the most complete wretch that ever existed. She put me into an egg-shell. I was a little child, and I was put into a pigsty. I had no shoes, no stockings, no clothes, no nothing.' (Ch. 6)

The power of Bounderby the manufacturer is based on his mastery of statistical knowledge, that is, his skills in quantifying his workers' productivity, the profitability of his factory, and the role played by his enterprise in the national economy.

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

In the eyes of CDS, Bounderby's appeal to statistics can be interpreted as an exercise in data-driven power: the power of quantitative data to create and reinforce social hierarchies. Bounderby's statistical stories help to naturalize economic inequality by attributing it to individual merit, not structural advantage.

The novel exposes the Bounderby self-made man narrative as a sham, and, in effect, exposes the statistical basis upon which he stands as an authority as well. If Bounderby's statistical story is bogus, so are also the numbers-based rationales for economic inequality.

In Bounderby's interactions with his employees, too, one can see the logic of modern-day work place surveillance and algorithmic management. The employees in Bounderby's factory are treated as productive components whose value is determined by their productivity; their actions are checked and controlled; and their personal requirements and lives are sacrificed to the statistical imperative of efficiency and productivity.

5. From Victorian Statistics to Big Data: Continuities and Transformations

Through the lens of CDS, some striking parallels can be seen between Victorian statistical thinking and what is happening now in data culture. These characteristics of the modern data culture, which Dickens lampooned in *Hard Times*, actually existed in his own time in the Victorian statistical movement: Data viewed as objective, value neutral knowledge about the world, believing that "more data means better decisions," treating data as a self-evident representation of the world, using data to legitimate and reproduce existing power structures.

But there are also important differences between the Victorian statistics and the modern-day big data. Data collection and analysis today is unimaginable to the Victorians: digital systems are gathering and analysing billions of data points every day, at a speed and scale which is beyond human comprehension. The models of analysis of

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

these data are much more powerful than the Victorian statistical tables.

The basic issues that Dickens asks in *Hard Things* are relevant to the current culture of data. Are there other parts of human experience that can only be felt? Is there always a better approach to making decisions based on data than on intuition and empathy or moral judgment? Who wins and who loses with data-driven decision making?

From Victorian statistics to big data, there are also new queries regarding the link between data and democracy. *Hard Times* highlights a tension between the data-driven perspective of Gradgrind and the novel's humanist perspective, as well as a larger debate over the place of quantitative knowledge in democratic societies. This debate has grown more acute in the modern day, as governments, businesses and civil society are relying more and more on information and data to make decisions.

As with facts and statistics in the Victorian era, the notion of 'dataism' - the idea that data are the basis for understanding and acting in the world - is a dominant ideology in today's societies. Dataism can be

described by Van Dijck (2014) as 'a pervasive ideology that privileges quantifiable information over other forms of knowledge.' As *Hard Times* shows, this dataist ideology is not a new phenomenon but a further elaboration of the data rationalism that Dickens was criticizing in the Victorian era.

5.1. Sissy Jupe: Imagination vs. Fact

The most overt alternative to data rationalism in the novel is the character of Sissy Jupe. The daughter of a circus worker, Sissy is enrolled in Gradgrind's school but cannot give him the statistical definitions that he requires. On being asked what a horse is, she answers in a pictorial way, not a literal sense:

'Please, sir, I cannot define it, if you will ask me about it. I mean, I cannot tell you what I would call it, because I don't know. The horse is a quadruped, a graminivorous animal, with forty teeth-' (Ch. 2)

Sissy's failure to give the horse a factual definition isn't any evidence of her

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

lack of intelligence; it's proof of her refusal to confine the living breathing horse to a set of statistics. Her alternative way of knowing (derived from imagination, empathy, and direct experience) stands for humanist values which she opposes to the philosophy of facts upheld by Gradgrind.

It would be easy to view this as a shift in values that contradicts the scientific method, yet Sissy's success with Gradgrind's system – the value of empathy and imagination that Gradgrind's own children are lacking – implies that the values of empathy and imagination are not only a supplement to data-driven knowledge, but are necessary to its acquisition. This insight is a contemporary empowerment of arguments within the CDS that data-driven decision making needs to be complemented with qualitative, contextual and humanistic forms of knowledge.

Rachael, the empathic nurse tending to Stephen Blackpool, is another voice in the rationalization of human cost involved in data-driven rationality. Rachael's attitude towards Stephen, her rejection of being considered a statistical category, and her

belief in the worth of human relationships are an opposing view to the data driven logic of Coketown. Her character implies that values, such as caring, solidarity and mutual aid, which cannot be measured, are fundamental elements of a just society.

5.2. Stephen Blackpool: The Human Cost of Data

Stephen Blackpool embodies the human cost of decision-making based on data, in its most extreme form. The "statistical logic of the factory system" catches Stephen, a respectable working man in Bounderby's factory, who is unable to get a job, a divorce, or out of the poverty that is bound up in the system. The meeting between Stephen and Bounderby is significant:

'I ha' read, in a aw, that a' the poor were to be blotted out.' *'Blotted out?'* said Bounderby. *'I mean, sir,'* Stephen replied, *'that they were to be done away wi'.* *One time, they were to be blotted out altogether.'* (Ch. 15)

The reference to 'blotting out' the poor is illustrative of the 'governance by the numbers' approach that reduces the poor to a

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

data issue which can be addressed and resolved. When poverty is viewed statistically, as allowed by the Poor Law Amendment Act of 1834, the poor are seen as something to be managed, optimized and, if possible, eliminated. This is the human cost that Stephen has experienced.

Stephen can't get a divorce, because his wife is an alcoholic and has left him, but the law won't allow for a divorce. Another kind of data driven injustice is Stephen's inability to get a divorce, due to his wife being an alcoholic and abandoning him, but the law doesn't allow divorce. The legal system is as much of a system as the factory system, and functions on rules, rather than on the actual details of a person's life. Stephen's suffering is unsaved by any statistic, it is an experience of injustice that cannot be calculated and accepted by the systems of logic and data.

The ultimate cost of data-driven governance is the death of Stephen, who died after being trapped in a mine shaft following a workers' protest, even though he was innocent. His death is not an isolated incident, but a tragedy, a life ruined by the

combination of economic exploitation, legal injustice and political manipulation. This death of Stephen, therefore, is a compelling indictment of the data-driven logic that seeks to categorize human lives.

5.3. The Role of Rachael: Care vs. Data

The image of Rachael, the caring nurse attending to the man who performs the rational calculations of Stephen Blackpool provides an alternative view of the human price of rational calculation. Rachael's concern for Stephen, her unwillingness to be categorized and reduced to a number, and her belief in the worth of one-to-one human contact is an alternative to the data-driven logic that rules Coketown. Rachael tells about her relationship with Stephen:

*'I have known him from a child.
We were both poor, and both worked
hard, and both were often enough
without the common necessities of
life. He was always kind to me, and I
was always kind to him.'* (Ch. 16)

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

The narrative on Rachael's relationship with Stephen ('kind to me' and to him) is a knowledge based not on statistical calculation but on care, solidarity and mutual aid. Her relationship with Stephen is not a statistic, but an experience of human connection that is out of the reach of the statistical tools of the factory system.

Rachael's character implies that the values of care, solidarity and mutual aid are crucial values of a just society, which are not quantifiable or measurable by statistical approaches. This is in line with the feminist critiques of data culture today that emphasize the need to address the values of care and relationality in the process of data-driven decision making. The novel's portrayal of Rachael thus anticipates the current debate about the incorporation of care ethics into data-driven governance.

The novel's closing pages, which represent an alternative model of a more "humane" sociality, may be interpreted as a condemnation of not only the special kinds of data rationality portrayed in the novel, but also of the general epistemological paradigm that gives preference to quantification over

understanding, facts over wisdom, and statistical regularity over personal experience. Dickens's ideas about society valuing imagination, empathy, and care over factual information can be interpreted as being the beginning of the current debate on the need for data-driven decisions to be backed up by knowledge that is qualitative, contextual, and humanistic.

6. The Human Cost of Data: *Hard Times* and Contemporary Data Ethics

Both *Hard Times* and the current CDS emphasize the "human" side of decision making. Dickens' depiction of the honest but ruined Stephen Blackpool, for instance, predicted modern reports of workers who are displaced, exploited, or excluded from work due to algorithmic decisions. In not finding work, not getting a divorce, or not breaking the cycle of poverty, Stephen's situation is a case of how a system designed to be rational is also morally irrational. Stephen's situation is a case of how a system designed to be rational is also morally irrational, in not

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

finding work, in not getting a divorce, or in not escaping the poverty that traps him.

The character of Sissy Jupe, who is barred by Gradgrind from entering the school, because she can't provide statistical definitions, and who stands for the values of imagination, empathy, and care, can be interpreted as representing the opposite of data rationalism, which Dickens and contemporary CDS scholars are promoting. In the Victorian era, as now, Sissy's resistance to being encapsulated by a number, to being reduced to a data point, speaks of the power of stories, emotions, and relationships.

The compassionate woman, Rachael, who helps Stephen Blackpool, provides another view of the human cost of rationality, driven by data. Rachael's concern for Stephen, her defiance against being categorized, and her belief in the importance of personal relationships offers an alternative to the data-driven logic which prevails in Coketown.

It's worth noting that the modern movement for data ethics, which calls for transparency, accountability, and fairness in the use of data for decision-making shares

many of the values Dickens outlines in *Hard Times*. All of these calls, for democratic oversight of data practices, for transparency and accountability of algorithmic models, and for human impacts over statistical efficiency of algorithmic decisions, are all grounded in humanist values that Dickens is countering in the face of Gradgrind's philosophy of facts.

The novel's closing pages, which represent the collapse of the industrial order in Coketown and an alternative model of rational order, could be interpreted not just as a condemnation of the particular kinds of data rationality presented in the novel, but also of the epistemological system that values quantification over understanding, facts over wisdom, and statistical regularity over personal experience.

7. Conclusion

The aim of this paper is to maintain that Dickens's *Hard Times* anticipates key issues in the Critical Data Studies of the present day by providing a critical analysis of data rationalism. The novel's account of Gradgrind's 'fact' education, Bounderby's

**International Journal of Multidisciplinary Innovations & Studies
(IJMIS)**

'statistical' control of labour, and the general description of the 'fact' of human experience in numbers, reveals the continuities between Victorian statistical thinking and the present day culture of data. The central themes of Dickens in his work, *Hard Times*, -data and knowledge, quantification and understanding, statistical rationality and human values - are as relevant today as they were in Victorian times. Moving from Victorian statistics to the era of big data has changed the nature of the data-driven decision-making process and greatly expanded the scale and scope of such decisions, but it has not solved the fundamental conflict between data rationalism and humanist values that Dickens pointed to. Industries that are displaced, communities that are excluded, experiences that are never measured: the human price of algorithms, data-driven policies, and data-driven decision-making is very much an issue in today's societies. Future work could be the extension of this work into other Victorian novels, such as by an investigation of the statistical thinking of writers like Thomas Hardy, George Eliot and Elizabeth Gaskell.

International Journal of Multidisciplinary Innovations & Studies (IJMIS)

References

- Beer, D. (2016). *Metric Power*. London: Palgrave Macmillan.
- Boyd, D. and Crawford, K. (2012). 'Six Provocations for Big Data Research.' In *Proceedings of the 2012 ACM Workshop*. New York: ACM, pp. 1-8.
- Cheney-Lippold, P. (2017). *We Are Data: Algorithms and the Making of Our Digital Selves*. New York: New York University Press.
- Dickens, C. (2008). *Hard Times*. Ed. D. Craig. London: Penguin. (Original work published 1854.)
- Drucker, J. (2014). *Graphesis: Visual Forms of Knowledge Production*. Cambridge: MIT Press.
- Drucker, J. (2014). *Theorizing Interpretation*. Minneapolis: University of Minnesota Press.
- Gillespie, T. (2020). 'The Politics of Platforms.' In *The Datafied Society*. Ed. M. Schaefer and K. van Es. Amsterdam: Amsterdam University Press, pp. 35-50.
- Hacking, I. (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.
- Kitchin, R. (2014). *The Data Revolution*. London: Sage.
- Lupton, D. (2018). *Digital Health: Critical and Cross-Disciplinary Perspectives*. London: Routledge.
- Mayer-Schonberger, V. and Cukier, K. (2013). *Big Data*. Boston: Houghton Mifflin Harcourt.
- Noble, S. U. (2018). *Algorithms of Oppression*. New York: New York University Press.
- O'Neil, C. (2016). *Weapons of Math Destruction*. New York: Crown.
- Porter, T. M. (1986). *The Rise of Statistical Thinking, 1820-1900*. Princeton: Princeton University Press.
- Schaefer, M. and van Es, K. (2020). *The Datafied Society*. Amsterdam: Amsterdam University Press.
- Seaver, N. (2022). 'Big Data and the Historical Imagination.' *Victorian Studies*, 64(2), 245-268.

**International Journal of Multidisciplinary Innovations & Studies
(IJMIS)**

Thompson, M. (2021). 'Dickens, Statistics, and the Politics of Fact.' *Nineteenth-Century Literature*, 75(4), 456-480.

van Dijck, J. (2014). 'Datafication, Dataism and Its Discontents.' *Internet and Society*, 2(1), 5-18.

Wallace, J. (2019). '*Hard Times* and the Politics of Quantification.' In *The Cambridge Companion to Dickens*. 2nd ed. Ed. J. Jordan. Cambridge: Cambridge University Press, pp. 156-172.

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. New York: PublicAffairs.