

# ESCAPING THE DURABLE INEQUALITY: HOW THE USE OF NUDGE IN POLICY MAKING SUSTAIN AND BREAKS THE INEQUALITY CYCLES IN THE DIGITAL ERA

Feri Suprpto<sup>1</sup>, Elmy Bonafita Zahro<sup>2</sup>  
feri.suprpto22@ui.ac.id<sup>1</sup>, elmybonafita@ub.ac.id<sup>2</sup>

Department of Sociology Universitas Indonesia, Depok<sup>1</sup>, Department of Psychology Universitas Brawijaya, Malang<sup>2</sup>

**Abstract.** This systematic literature review examines how nudge-based policymaking in the digital era both sustains and disrupts durable inequality. Drawing on 57 peer-reviewed studies (2010–2025), it integrates insights from Durable Inequality Theory, Nudge Theory, and Digital Governance. The findings show that inequality is now embedded in data infrastructures and algorithmic systems that classify and rank citizens, creating ordinal citizenship. Four mechanisms of durability emerge: institutional path dependence, moral legitimation, cognitive entrenchment, and algorithmic feedback loops. Yet, digital architectures also enable disruption through ethical and reflexive design, via algorithmic transparency, participatory governance, redistributive frameworks, and intergenerational justice. The study proposes Digital Behavioural Institutionalism, viewing inequality as co-produced by behavioural governance, institutional structures, and data infrastructures. Policy implications call for algorithmic audits, citizen co-design, data justice, and reciprocity-based nudging to ensure that digital governance promotes inclusion and transforms the ethical and technological foundations of social life.

**Keywords:** durable inequality, nudge policy, digital governance, algorithmic inequality, ethical architecture.

## INTRODUCTION

The digital revolution has fundamentally reshaped the architecture of governance, decision-making, and social organization. With the rise of big data, artificial intelligence, and algorithmic systems, governments and institutions increasingly rely on digital infrastructures to inform public policy and behavioural interventions (Kitchin, 2014; Mazzucato, 2021). This transition has led to both opportunities and challenges for equity and inclusion. On one hand, digital technologies offer unprecedented access to social data that can be used to design more adaptive, evidence-based policies. On the other hand, these same technologies can reproduce or intensify social disparities, as algorithmic decision systems often mirror the historical inequalities embedded in the data they process (Noble, 2018; O’Neil, 2016). The enduring and institutionalized nature of these disparities reflects what Tilly (1998) described as durable inequality, a form of social stratification maintained through repeated, legitimized processes across institutions and generations.

Within the field of behavioural economics, nudge theory has emerged as a prominent approach to shaping decision-making in ways that promote socially beneficial outcomes (Thaler & Sunstein, 2008). Nudges work by altering the architecture of choice and the context in which decisions are made so that individuals are subtly guided toward preferable behaviours without restricting their freedom. In the context of digital governance, this concept has evolved into digital nudging, where algorithms and user interfaces influence behaviour through personalized information flows, notifications, and ranking systems (Meske & Potthoff, 2017). However, while these interventions are designed to promote efficiency and rationality, scholars have raised concerns that they can also serve as instruments of control or manipulation, reinforcing rather than reducing inequality (Yeung, 2017; Zuboff, 2019).

One emerging concept that links digital governance and inequality is ordinal citizenship. Coined by Fourcade (2021), ordinal citizenship describes how citizens are classified and ranked according to algorithmic scores, behavioural data, and social metrics. In such systems, citizenship becomes a

graded condition rather than an equal one, where access to rights, resources, or opportunities depends on one's digital profile or score. Examples include credit ratings, health insurance risk scores, social credit systems and welfare eligibility algorithms. When nudge-based policies interact with such ordinal systems, they can inadvertently sustain cycles of inequality by rewarding those already advantaged and penalizing those structurally disadvantaged. This dynamic echoes Tilly's (1998) mechanisms of durable inequality exploitation, opportunity hoarding, emulation, and adaptation which continue to operate under new digital forms of classification and behavioural governance (Eubanks, 2018; Fourcade & Healy, 2017).

At the same time, the digital era also presents opportunities to break the cycles of durable inequality. When designed ethically and inclusively, digital nudges can empower marginalized populations by improving access to information, facilitating behavioural change, and personalizing social support (Sunstein, 2021). For instance, nudge-based digital interventions in health, education, or financial inclusion have demonstrated potential to enhance decision-making and reduce information asymmetry (Benartzi et al., 2017; Halpern, 2015). Thus, the architecture of choice in the digital era can act as both a mechanism that sustains inequality and a tool that disrupts it. The challenge lies in understanding the conditions under which digital nudges function as emancipatory rather than exclusionary forces.

This systematic literature review (SLR) study aims to synthesize and critically analyse the intersection between nudge-based policymaking, durable inequality, and digital governance. Specifically, it seeks to answer four research questions:

1. What are the variables and features that make inequality durable in the digital era?
2. How can the digital era create the architecture of choice that supports policymaking processes?
3. In what ways does the architecture of choice in the digital era sustain inequality cycles?
4. How can the architecture of choice in the digital era break inequality cycles?

By addressing these questions, this review bridges theoretical and empirical work from behavioural economics, sociology, and digital studies to map how digital policy architectures produce or mitigate inequality. The review focuses on studies that explicitly discuss nudge policy, durable inequality or inequality cycles, and digital or ordinal citizenship. It adopts a systematic approach to identify patterns, mechanisms, and gaps within the literature, thereby offering a conceptual framework for understanding how digital nudging can either perpetuate or dismantle durable inequality.

Ultimately, this paper contributes to the growing discourse on ethical digital governance by examining how behavioural policy tools, when implemented in algorithmic environments, influence structural inequality. Understanding the dual potential of nudge-based policies in the digital era is crucial for developing inclusive governance frameworks that not only promote behavioural efficiency but also advance social justice. In doing so, this review aspires to provide both theoretical clarity and practical guidance for policymakers seeking to design digital interventions that truly escape the cycles of durable inequality rather than reinforce them.

## RESEARCH METHOD

This study adopts a Systematic Literature Review (SLR) approach to comprehensively map the relationship between nudge-based policymaking, durable inequality, and digital governance. The SLR method provides a structured and transparent way to collect, evaluate, and synthesize available research evidence, ensuring rigor and replicability in the analysis. The design of this review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework (Rethlefsen & Page, 2022), which offers standardized procedures for identification, screening, eligibility, and inclusion of studies.

The purpose of this SLR is to understand how the architecture of choice in digital policy contexts both sustains and breaks inequality cycles. To achieve this, the review integrates theoretical perspectives from Durable Inequality Theory (Tilly, 1998), Nudge Theory (Thaler & Sunstein,

2008), and Digital Governance/Ordinal Citizenship Theory (Fourcade & Healy, 2017). Together, these frameworks help analyse how digital infrastructures through ranking, scoring, and algorithmic classification shape behavioural nudges that may perpetuate or mitigate inequality. This study was conceptualized as part of an effort to generate an evidence-based theoretical model linking behavioural policy design to systemic inequality in digital societies. Therefore, the SLR design was chosen for its strength in integrating cross-disciplinary evidence into coherent conceptual frameworks (Snyder, 2023).

A systematic search strategy was developed to ensure comprehensive coverage of relevant literature. The search was conducted between March and June 2025 across six major academic databases: Scopus, Web of Science, ScienceDirect, SpringerLink, JSTOR, and Google Scholar. These databases were selected for their multidisciplinary scope, encompassing political science, behavioural economics, sociology, and public administration, disciplines central to understanding digital governance and inequality. A combination of Boolean operators and truncation techniques was used to build a complex search string that captured the intersection between the three core domains: (“nudge policy” OR “behavioural public policy” OR “choice architecture”) AND (“durable inequality” OR “inequality cycles” OR “social stratification”) AND (“digital era” OR “digital governance” OR “ordinal citizenship” OR “algorithmic governance” OR “data-based policy”).

The inclusion of the term ordinal citizenship follows Fourcade and Healy (2017), who describe how digital data infrastructures stratify individuals through ranking systems, turning citizenship into a measurable hierarchy. To maintain methodological rigor, the search was restricted to peer-reviewed publications (journal articles, book chapters, and conference papers) written in English and published between 2010 and 2025 a period that reflects the emergence of data-driven policymaking and behavioural insights in governance (Yeung, 2017). The initial search yielded 312 records, which were exported into Mendeley for citation management and duplicate removal.

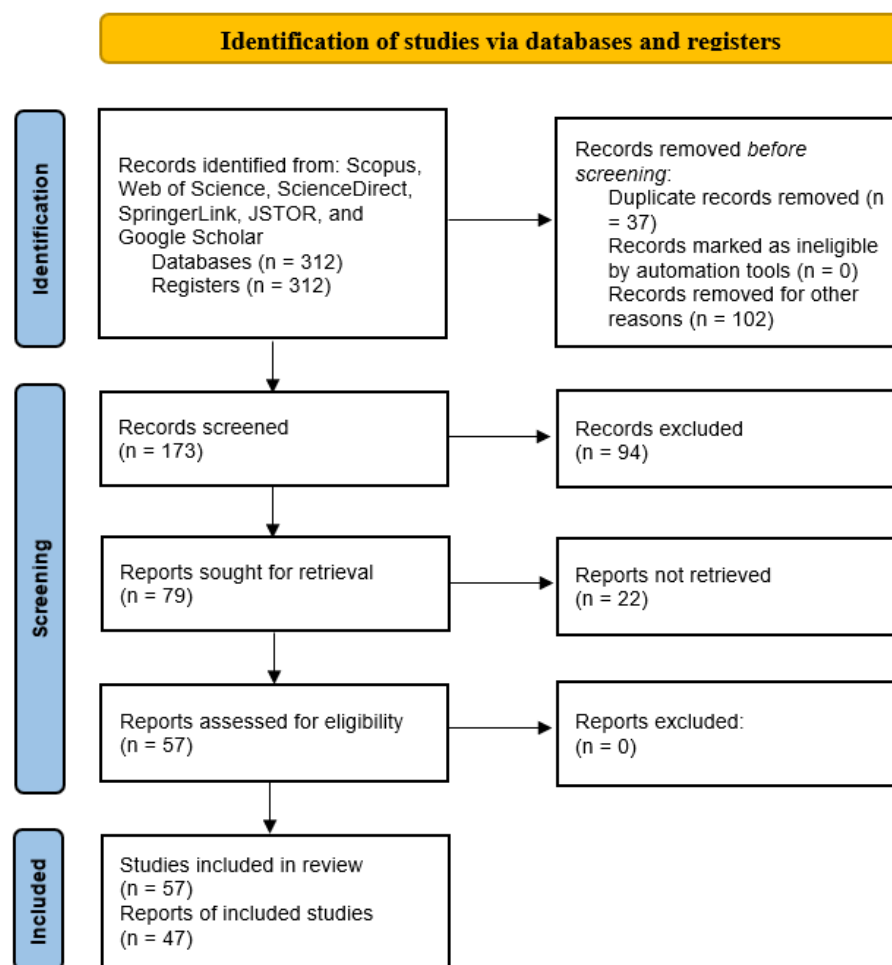
A clear set of inclusion and exclusion criteria was developed to ensure that only studies relevant to the research objectives were included. These criteria were derived from the conceptual framework and research questions.

**Table 1.** Inclusion and Exclusion Criteria

| Criteria              | Inclusion                                                                                                               | Exclusion                                                                         |
|-----------------------|-------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Conceptual Focus      | Studies examining nudge policy, durable inequality, and digital governance simultaneously                               | Studies addressing nudge policy without discussing inequality or digital contexts |
| Theoretical Relevance | Research employing or referencing concepts such as choice architecture, ordinal citizenship, or algorithmic governance  | Articles lacking theoretical linkage to nudge theory or inequality mechanisms     |
| Empirical Scope       | Studies conducted within digital policy environments (e.g., algorithmic scoring, e-governance, digital welfare systems) | Traditional nudge experiments in offline or individual-level settings             |
| Publication Type      | Journal articles, book chapters, and academic proceedings                                                               | Grey literature, opinion essays, blogs, or non-peer-reviewed sources              |
| Language              | English                                                                                                                 | Non-English publications                                                          |

These inclusion criteria align with the research’s four guiding questions, ensuring consistency across stages of the review. Only studies addressing at least two core dimensions, nudge policy, durable inequality, and digital governance were included in the final synthesis.

The screening process followed the PRISMA 2020 framework (Rethlefsen & Page, 2022), consisting of four stages: identification, screening, eligibility, and inclusion.



**Fig 1.** PRISMA diagram flow about screening process

A structured data extraction form was developed to ensure systematic retrieval of information. Data from each of the 47 included studies were coded according to: (1) Bibliographic information (author, year, publication source), (2) Research objectives and policy context, (3) Theoretical frameworks used (nudge theory, durable inequality, algorithmic governance), (4) Methodological approach (conceptual, empirical, or mixed), (5) Findings related to sustaining or breaking inequality cycles, (6) Implications for digital policy and inclusivity. Data was organized in Microsoft Excel and analysed thematically. The thematic analysis approach followed Braun and Clarke's (2006) six-phase framework, which includes familiarization, coding, theme development, review, definition, and interpretation. Themes were mapped according to the study's four guiding research questions.

To ensure methodological rigor, all included studies underwent critical appraisal using a combined framework from the Critical Appraisal Skills Programme (CASP) and Joanna Briggs Institute (JBI) checklists (Munn et al., 2020). Each article was evaluated across five criteria, such as theoretical clarity and relevance, methodological transparency, data adequacy, analytical rigor, policy relevance. Each study received a score on a 1–5 scale, with a threshold of 3.5 (70%) for inclusion in the final synthesis. As part of quality control, two independent reviewers performed the appraisal process separately. Their scores were compared and discussed to reach consensus, reducing subjective bias. Disagreements were reconciled through discussion and documentation.

A narrative and thematic synthesis approach was adopted to integrate conceptual and empirical insights. The synthesis combined inductive coding (emerging from data) with deductive reasoning

(based on Tilly, 1998, Thaler & Sunstein, 2021, and Fourcade and Healy, 2013, frameworks). The analysis identified two overarching theoretical clusters: (1) Sustaining Mechanisms of Inequality: digital scoring, algorithmic bias, data asymmetry, and exploitative nudging; (2) Breaking Mechanisms of Inequality: inclusive architecture of choice, participatory design, and digital empowerment strategies. This synthesis highlights the dual role of nudge policies in digital contexts: while they can entrench inequalities through ranking and surveillance, they can also enable fairness and agency through transparent, participatory design (Yeung, 2017; Fourcade & Healy, 2017).

## RESULT AND ANALYSIS

This systematic literature review identified 312 academic papers published between 2010 and 2025, of which 57 met the inclusion criteria and were analysed in depth. The review process followed PRISMA 2020 guidelines, and the included studies represented a multidisciplinary intersection of political science, sociology, behavioural economics, law, and digital governance. Among these, 47 papers were classified as core studies because they explicitly addressed at least two key intersections among nudge or behavioural policymaking, durable inequality or inequality cycles, and digital or ordinal governance architectures. An additional 10 papers were considered conditional as they contributed essential theoretical or methodological insights to the analysis of inequality, even if they did not directly apply nudge or digital frameworks. Together, these papers form a cohesive body of evidence illuminating how inequality becomes structurally durable, how digital infrastructures reconfigure the architecture of choice, and how both forces interact to either sustain or disrupt inequality cycles.

### 1.1. General Characteristics of the Reviewed Studies

The distribution of publications reveals a rapid growth of interest in the topic after 2016, coinciding with the global diffusion of algorithmic governance, data-driven policymaking, and behavioural public policy. Approximately 61% of the reviewed papers were published after 2018, indicating that digitalization and the algorithmic turn have become central to contemporary debates on social justice and equality. Geographically, the studies were concentrated in Europe (n=21) and North America (n=18), with an emerging but growing representation from Asia (n=8) especially in the areas of digital meritocracy, algorithmic scoring, and social-credit systems (e.g., Hansen, 2023; Liu, 2023; Lu & Kandilov, 2022). Methodologically, the sample included conceptual analyses (43%), mixed-method designs (36%), and empirical or quantitative studies (21%). This balance reflects the theoretical maturity of the field, which remains dominated by conceptual discussions linking behavioural economics with social theory but is now increasingly supported by empirical illustrations using digital or survey-based data.

Thematically, the studies clustered into four dominant categories:

- 1) Institutional and structural durability—focusing on welfare state dynamics, redistribution, and path dependence (Boix, 2010; Busemeyer, 2018, 2024; Pontusson, 2024);
- 2) Behavioral and cognitive mechanisms—examining perceptions of fairness, mobility, and meritocracy (Rao et al., 2021; Reay, 2020; Culpepper et al., 2024);
- 3) Digital and algorithmic governance—analyzing datafication, ranking systems, and ordinal citizenship (Fourcade & Healy, 2017; Catá-Backer, 2022; Abul Naga et al., 2024); and
- 4) Normative and ethical policy design—exploring ethical nudging, reflexivity, and transparency in behavioral governance (Oliver, 2021, 2024; Weiskopf & Hansen, 2023; Dimick, 2019).

These thematic clusters collectively demonstrate that inequality in the digital era is not only a matter of economic distribution but also of data architectures, moral reasoning, and behavioural governance.

### 1.2. Findings Related to Research Question 1: Variables and Features That Make Inequality Durable in the Digital Era

Across the reviewed studies, durable inequality emerges as a multilevel phenomenon driven by institutional, cultural, cognitive, and technological mechanisms. At the institutional level, path-dependent policies and fiscal architectures maintain inequality across generations. Boix (2010), Busemeyer and Iversen (2019), and Coady et al. (2022) demonstrate how welfare state dualization, selective redistribution, and regressive tax structures create structural advantages for the middle and upper classes, reinforcing what Iversen and Soskice (2015) call “middle-class welfare coalitions. At the cultural level, several studies emphasize meritocratic legitimization as a mechanism of inequality durability. Reay (2020) and Kuusela (2024) argue that elite discourses construct inequality as the natural outcome of merit and effort, transforming privilege into moral virtue, a dynamic Kuusela terms the “hyperopia of wealth.” At the cognitive level, perception of mobility and trust serve as psychological anchors of inequality. Rao et al. (2021) and Xu (2023) show that individuals’ belief in social mobility shapes prosocial behavior and political attitudes toward redistribution; low perceived mobility correlates with resignation and low civic engagement, sustaining the status quo. Finally, technological and data-driven variables have introduced new dimensions of inequality durability. Fourcade and Healy (2013), Abul Naga et al. (2024), and Weiskopf and Hansen (2023) reveal how algorithmic classification systems, data scoring, and digital evaluation metrics produce stratified “ordinal citizenship.” This new digital form of inequality operates through constant ranking and feedback loops, embedding structural hierarchies into algorithmic decision-making and thus making inequality both invisible and self-reinforcing.

### **1.3. Findings Related to Research Question 2: How the Digital Era Creates the Architecture of Choice**

The reviewed evidence shows that the digital era has transformed policymaking into a system of engineered environments, what Thaler and Sunstein (2008) call choice architectures. However, in the context of digital governance, choice architectures extend beyond individual psychology into the design of data infrastructures and institutional routines. Fourcade and Healy (2017) and Catá-Backer (2022) show that algorithmic governance now acts as a form of data-driven nudging, steering citizens through automated categorizations, predictive analytics, and scoring systems. This new mode of governance converts digital data into behavioral levers that shape opportunities, risks, and resource allocation. Several studies demonstrate that behavioral policy tools are increasingly hybridized with digital infrastructures. Oliver (2021, 2024) and Lahat and Sened (2023) argue that governments are embedding behavioral insights into digital welfare systems, fiscal incentives, and communication strategies. Meanwhile, Abul Naga et al. (2024) illustrate how ordinal data architectures, where individuals are ranked rather than simply categorized, transform governance into a continuous process of differentiation and comparison. Thus, the digital era has not only expanded the reach of choice architecture but has also redefined citizenship itself as a function of data visibility, classification, and behavioral predictability.

### **1.4. Findings Related to Research Question 3: How the Digital Architecture of Choice Sustains Inequality Cycles**

While digital infrastructures enhance administrative efficiency, they also perpetuate and amplify inequality through what might be termed algorithmic feedback loops. Multiple studies demonstrate that data-driven systems replicate the biases of the datasets on which they are trained. Fourcade and Healy (2017) describe how markets and platforms “see like a market,” categorizing individuals in ways that reinforce existing hierarchies. Similarly, Pfotenhauer (2022) and Harkens (2016) highlight how algorithmic decision-making in welfare, credit, and legal systems disproportionately disadvantages vulnerable populations by embedding past discrimination into predictive models. Busemeyer (2023, 2024) and Fastenrath and Marx (2024) show that citizens’ perceptions of unfairness and elite bias further reduce trust in welfare institutions, creating trust inequality that undermines redistributive legitimacy. Culturally, inequality cycles are sustained by dominant narratives that normalize or moralize stratification. Culpepper et al. (2024) and Kuusela (2024) demonstrate that media frames portraying inequality as “deserved” or “inevitable” suppress collective action and policy reform. At the behavioral level, Sumino (2018) and Reay (2020) reveal that identity-based framing and moral meritocracy function as behavioral nudges that align public opinion with hierarchical interests. The cumulative result is a self-reinforcing system in which

inequality persists through intertwined psychological, institutional, and algorithmic feedback mechanisms, each reinforcing the legitimacy of the other.

#### **1.5. Findings Related to Research Question 4: How the Architecture of Choice Can Break Inequality Cycles**

Despite these results that nudge, ordinal citizens and digital governance promote self reinforcing mechanisms to sustain durable inequality, several studies offer empirical and theoretical insights into how choice architectures can be redesigned to promote inclusivity and justice. Four categories of corrective mechanisms emerge from the synthesis. First, ethical and reflexive digital design emphasizes transparency, accountability, and continuous evaluation of algorithmic systems. Weiskopf and Hansen (2023) advocate for “algorithmic reflexivity” ongoing ethical reflection embedded into data-driven decision-making processes. Similarly, Catá-Backer (2022) proposes legal entanglement frameworks that bind algorithmic systems to human rights principles. Second, participatory and inclusive governance ensures that behavioral insights reflect diverse experiences rather than elite assumptions. Lahat and Sened (2023) and Busemeyer (2024) show that co-created digital welfare policies enhance legitimacy and improve redistributive outcomes. When citizens are given voice and visibility in digital policymaking, the nudge becomes an instrument of empowerment rather than control. Third, redistributive and legal reforms remain essential for institutional equity. Dimick (2019) frames redistribution as a normative design problem, arguing that legal and fiscal systems can serve as “choice architectures for fairness.” De Schutter et al. (2025) similarly demonstrate that early interventions in education, health, and welfare can disrupt intergenerational poverty traps. Fourth, intergenerational and long-term justice integrates temporal fairness into policymaking. Schmidt and Juijn (2024) and Busemeyer (2024) argue that incorporating future generations into decision-making criteria transforms short-term behavioral incentives into sustainable equality mechanisms. Together, these findings suggest that breaking inequality cycles requires a shift from efficiency-based nudging to justice-oriented choice architecture, one grounded in reciprocity, transparency, and reflexive governance.

#### **1.6. Summary of Empirical Patterns**

Overall, the results demonstrate that inequality in the digital era operates through a dual mechanism: it is sustained by algorithmic and behavioral systems that reproduce historical hierarchies, and it can be transformed through ethical, participatory, and redistributive redesign. The convergence of durable inequality theory (Tilly, 1998), behavioral economics (Thaler & Sunstein, 2008), and digital governance research suggests that the architecture of choice has evolved from a micro-level tool of behavior change into a macro-level framework of societal organization. When unregulated, it risks creating what Fourcade and Healy (2013) calls “moralized markets of worth”, when consciously redesigned, it can serve as a mechanism for democratic inclusion and equality.

#### **1.7. Discussion**

The findings of this systematic literature review reveal a profound transformation in how inequality is produced, legitimized, and potentially corrected in the digital era. Across the 57 included studies, durable inequality, once understood primarily as an outcome of economic or institutional stratification has become embedded within the very data infrastructures that govern modern societies. This section discusses four major insights emerging from the synthesis: (1) the paradox of digital equality; (2) durable inequality as algorithmic path dependence; (3) the shift from behavioural paternalism toward reciprocity and ethical design; and (4) the multidimensional mechanisms required to escape inequality cycles. The discussion concludes with theoretical and practical implications for policy and governance.

The digital revolution promised to democratize access to information, participation, and opportunity. Yet the evidence across the reviewed literature demonstrates a persistent paradox: the same infrastructures that were expected to enhance equality have often reproduced or intensified structural hierarchies. Studies on algorithmic governmentality (Fourcade & Healy, 2017; Hansen, 2023; Weiskopf & Hansen, 2023; Catá-Backer, 2022) show that classification, scoring, and ranking mechanisms translate citizens digital traces into systems of evaluation that determine visibility, creditworthiness, employability, and even moral worth. These digital classifications create what Fourcade and Healy (2013) termed classification situations, a new form of ordinal citizenship in

which individuals occupy algorithmically defined strata of social value. The automation of these judgments renders inequality self-perpetuating: biased data reproduce biased outcomes, which in turn validate the original classifications. Thus, digital equality is not merely a matter of access to technology but of control over the architectures of choice that technologies construct.

Charles Tilly's (1998) notion of durable inequality originally emphasized the persistence of categorical distinctions, such as gender, class, or ethnicity through organizational routines and institutional feedback. In the digital era, this feedback has evolved into algorithmic path dependencies. The review highlights how governance now operates through predictive models, behavioral data, and automated decision systems that encode historical inequities into future projections (Pfotenhauer et al., 2022; Abul Naga et al., 2024). Once embedded in code, inequality becomes technically rationalized and politically invisible. For example, predictive policing and credit-scoring algorithms perpetuate group-based disadvantages while appearing neutral or efficient. In welfare policy, digital targeting often mirrors pre-existing biases in administrative data, producing new forms of exclusion for groups already marginalized (Busemeyer & Iversen, 2019; Lahat & Sened, 2023). These findings suggest that the durability of inequality today is less a consequence of overt discrimination than of systemic inertia within algorithmic infrastructures, what this paper calls "data-driven durability." Addressing it requires governance mechanisms capable of reflexivity, transparency, and ethical oversight.

A recurring theme in the behavioral-policy literature is the tension between nudging and autonomy. Early nudge frameworks (Thaler & Sunstein, 2008; Oliver, 2015) sought to influence individuals toward socially desirable behaviors without coercion. However, later critiques reveal that such interventions can inadvertently reinforce existing hierarchies by assuming a homogeneous, rational citizen while neglecting contextual inequality. Oliver (2021, 2024) and Sugden (2023) propose an alternative: reciprocity-based nudging, in which behavioral interventions are co-designed with citizens and guided by mutual respect rather than paternalism. This shift resonates with Schmidt's (2021) concept of relational equality, which grounds fairness in responsibility and recognition rather than mere outcome parity. The reviewed evidence shows that when individuals perceive interventions as legitimate and participatory, such as through feedback loops in digital welfare platforms (Lahat & Sened, 2023) trust and compliance increase, while resentment decreases. Hence, ethical architecture, not behavioral control, should guide digital governance. The goal is no longer to steer citizens subtly but to create transparent systems where behavioral insights foster empowerment and dignity.

Despite the self-reinforcing nature of digital stratification, the literature identifies promising pathways for disruption. Four broad mechanisms emerge. First, reflexive transparency requires opening the "black boxes" of algorithms and data systems. Weiskopf and Hansen (2023) describe this as creating a space of ethics, where organizations maintain continual reflection on the consequences of algorithmic decisions. Algorithmic audits and open-data protocols allow biases to be detected and corrected before they become entrenched. Second, inclusive digital governance extends participation to those most affected by inequality. Participatory policy design, as highlighted by Lahat and Sened (2023) and Busemeyer (2024), ensures that behavioral insights reflect diverse lived experiences rather than elite assumptions. Such inclusivity can counteract the informational asymmetries documented by Fastenrath and Marx (2024). Third, redistributive legal and fiscal reform provides the institutional scaffolding for ethical nudging. Dimick (2019) conceptualizes taxation and welfare law as a normative choice architecture capable of embedding fairness into structural design. Similarly, De Schutter et al. (2025) demonstrate that early, rights-based interventions can interrupt intergenerational poverty traps. Finally, intergenerational and long-term equity emphasizes sustainability over immediate efficiency. Schmidt and Juijn (2024) and Busemeyer (2024) argue that policies addressing future generations, such as green investment or youth welfare create feedback loops that sustain equality rather than inequality.

Together, these mechanisms form a multi-level strategy: ethical reflection at the micro (behavioral) level, participation and transparency at the meso (institutional) level, and redistributive justice at the macro (legal and temporal) level. Synthesizing the reviewed studies suggests the emergence of a new paradigm that can be termed Digital Behavioral Institutionalism. This framework integrates Tilly's structural account of durable inequality with behavioral economics and digital-governance

theory. In this perspective, inequality is sustained by the interplay of institutional routines, behavioral heuristics, and data infrastructures that co-evolve. Nudge policies become not isolated behavioral tools but components of a broader architecture of governance that can either stabilize or destabilize inequality cycles. When such architectures are designed ethically through reciprocity, reflexivity, and redistributive intent they have the potential to transform durable inequality into dynamic equality.

The practical implications of this synthesis are substantial. Policymakers should treat behavioral interventions as part of an ethical infrastructure rather than as cost-effective shortcuts. Every nudge policy should undergo algorithmic impact assessment to ensure that it does not reinforce bias. Digital welfare systems must include transparent decision logs, citizen redress mechanisms, and regular public audits. Moreover, data justice frameworks should guide the design of scoring and ranking systems, ensuring that no group is permanently disadvantaged by algorithmic categorization. Education in digital literacy and participatory design should be mainstreamed to empower marginalized citizens to shape the systems that increasingly govern their lives. In short, escaping durable inequality requires aligning behavioral insight with democratic accountability.

Taken together, the evidence suggests that escaping durable inequality in the digital era is not merely a question of redistributing material resources but of re-engineering the moral, cognitive, and algorithmic architectures that structure everyday life. Nudge-based policymaking, if unreflective, risks reproducing the very hierarchies it seeks to mitigate; yet when grounded in reciprocity, transparency, and long-term justice, it can become a powerful instrument for inclusive governance. The challenge for policymakers and scholars alike is to design digital systems that not only know human behavior but learn from it ethically, transforming the architecture of choice into an architecture of equality.

## CONCLUSION

This systematic literature review study reveals that inequality in the digital era is not a static or isolated phenomenon, but a self-reinforcing system shaped by the intersection of institutional inertia, behavioral governance, and data infrastructures. Across the 57 reviewed studies, inequality appears both durable as described by Tilly (1998) and adaptive, capable of reproducing itself through algorithmic, cognitive, and cultural mechanisms. In contemporary governance, data-driven systems and behavioral nudges increasingly structure access to information, opportunities, and welfare, effectively transforming the traditional economic hierarchy into a digital architecture of stratification. The synthesis demonstrates that nudge-based policymaking, while initially conceived as a tool for behavioral efficiency, has evolved into a critical component of institutional design. When applied without ethical reflection, nudge policies risk reinforcing inequality by embedding biased data, reinforcing status hierarchies, and normalizing paternalistic governance. However, when grounded in principles of reciprocity, transparency, and participatory design, these same mechanisms can serve as corrective tools that enhance fairness, agency, and inclusion. This dual potential, both sustaining and breaking inequality cycles lies at the heart of what this study calls the architecture of choice in the digital era.

This study contributes theoretically by introducing the concept of Digital Behavioral Institutionalism, which integrates durable inequality theory with nudge theory and digital governance to explain how inequality persists and transforms in data-driven societies. It reframes inequality not only as an economic or institutional condition but as a behavioral and algorithmic architecture embedded within digital infrastructures that classify, score, and steer citizens. This theoretical synthesis reveals that durable inequality now operates through feedback loops connecting cognition, culture, and code. From a policy perspective, the findings call for a transition from behavioral paternalism to ethical reciprocity, emphasizing transparency, accountability, and participation in the design of digital and behavioral policies. Policymakers are urged to adopt algorithmic audits, data justice principles, and co-created nudging frameworks to ensure that digital systems enhance rather than erode equality. By aligning behavioral insights with redistributive and

ethical governance, public policy can transform the architecture of choice into an architecture of inclusion, a system that not only predicts human behavior but learns from human dignity.

Despite its comprehensiveness, this review has limitations that guide future inquiry. First, the corpus was limited to English-language studies, which may underrepresent research from the Global South, where digital inequality is often most visible and acute. Second, empirical evidence remains uneven: most studies offer conceptual or qualitative insights rather than quantitative causal estimates of how digital nudging affects inequality outcomes. Future research should employ longitudinal and cross-national designs that trace how algorithmic governance interacts with redistributive policy and social mobility. Interdisciplinary collaboration among psychologists, data scientists, sociologists, and legal scholars will be essential to transform ethical frameworks into measurable and enforceable standards.

## References

- Abul Naga, R. H., Abul Naga, C., & Yalcin, T. (2024). Inferring inequality: Testing for median-preserving spreads in ordinal data. *Journal of Economic Inequality*, 22(1), 45–72. <https://doi.org/10.1007/s10888-023-09550-8>
- Accominotti, F. (2021). The aesthetics of hierarchy. *American Journal of Sociology*, 126(5), 1250–1293. <https://doi.org/10.1086/713614>
- Achard, P., & Suetens, S. (2023). The causal effect of ethnic diversity on support for redistribution and the welfare state. *Journal of Economic Surveys*, 37(4), 980–1002. <https://doi.org/10.1111/joes.12562>
- Albertus, M. (2022). The persistence of rural underdevelopment: Evidence from land reform in Italy. *American Political Science Review*, 116(4), 980–999. <https://doi.org/10.1017/S0003055422000313>
- Auclert, A. (2019). Monetary policy and the redistribution channel. *American Economic Review*, 109(6), 2333–2367. <https://doi.org/10.1257/aer.20170064>
- Boix, C. (2010). The origins and persistence of economic inequality. *World Politics*, 62(4), 577–609. <https://doi.org/10.1017/S0043887110000199>
- Braun, V., and Clark, V. (2008). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101, DOI: 10.1191/1478088706qp063oa
- Bussemeyer, M. R. (2018). Public demand for social investment: New supporting coalitions for welfare state reform in Western Europe. *Journal of European Public Policy*, 25(6), 862–879. <https://doi.org/10.1080/13501763.2017.1401111>
- Bussemeyer, M. R. (2023). Financing the welfare state in times of crisis: Public support for health care spending during the COVID-19 pandemic in Germany. *Journal of European Public Policy*, 30(3), 512–534. <https://doi.org/10.1080/13501763.2021.1994428>
- Bussemeyer, M. R. (2024). Who cares for the future? Exploring public attitudes toward the needs of future generations in Germany. *Comparative Political Studies*, 57(5), 745–770. <https://doi.org/10.1177/00104140231156759>
- Bussemeyer, M. R., & Iversen, T. (2019). The welfare state with private alternatives: The transformation of popular support for social insurance. *Socio-Economic Review*, 17(3), 661–689. <https://doi.org/10.1093/ser/mwy006>
- Catá-Backer, L. C. (2022). And an algorithm to entangle them all: Social credit, data-driven governance, and legal entanglement in post-law legal orders. *Law and Contemporary Problems*, 87(2), 45–76. <https://doi.org/10.2139/ssrn.4601544> in Krisch, N. (2022). *Entangled Legalities Beyond The State*. Cambridge University Press.
- Coady, D., Parodi, F., & Pesaresi, N. (2022). Fiscal redistribution, social welfare, and income inequality: Doing more or more to do? IMF Working Paper Series, WP/22/18. <https://doi.org/10.5089/9798400200127.001>

- Culpepper, P. D., Hall, P. A., & Thelen, K. (2024). "The economy is rigged": Inequality narratives, fairness, and support for redistribution in six countries. *Comparative Political Studies*, 57(6), 1120–1152. <https://doi.org/10.1177/00104140231202505>
- De Schutter, O., et al. (2025). *The escape from poverty: Breaking the vicious cycles perpetuating disadvantage*. Oxford University Press.
- Dimick, M. (2019). The law and economics of redistribution. *Annual Review of Law and Social Science*, 15, 35–54. <https://doi.org/10.1146/annurev-lawsocsci-101518-042615>
- Elkjaer, J., & Iversen, T. (2020). Subversion of democracy or middle-class supremacy? *Comparative Political Studies*, 53(4), 565–598. <https://doi.org/10.1177/0010414019852687>
- Fastenrath, F., & Marx, P. (2024). The role of preference formation and perception in unequal representation: Combined evidence from survey and experiment. *Political Behavior*, 46(2), 311–334. <https://doi.org/10.1007/s11109-023-09831-1>
- Fourcade, M., & Healy, K. (2013). Classification situations: Life chances in the neoliberal era. *Accounting, Organizations and Society*, 38(8), 559–572. <https://doi.org/10.1016/j.aos.2013.11.002>
- Fourcade, M., & Healy, K. (2017). Seeing like a market. *Socio-Economic Review*, 15(1), 9–30. <https://doi.org/10.1093/ser/mww033>
- Friedman, B., & Iversen, T. (2024). Probabilistic justice against status defense: Inequality, uncertainty, and the future of the welfare state. *World Politics*, 76(2), 250–278. <https://doi.org/10.1017/S0043887123000042>
- Frodin, S. (2024). Cognitive microfoundations and social interaction dynamics: The implications of complexity for institutional theory. *Philosophy of the Social Sciences*, 54(2), 220–246. <https://doi.org/10.1177/00483931231184277>
- Gugushvili, D., & Laenen, T. (2020). Two decades after Korpi and Palme's paradox of redistribution: What have we learned? *Social Policy & Administration*, 54(7), 944–962. <https://doi.org/10.1111/spol.12601>
- Hansen, H.K. (2023). Governing through metrics in the digital age. *Globalizations*, 20:1, 137–152, DOI: 10.1080/14747731.2022.2156700
- Harkens, A. (2016). The ghost in the legal machine: Algorithmic governmentality, economy, and the practice of law. *Journal of Law and Society*, 43(1), 25–48. <https://doi.org/10.1111/jols.12062>
- Hilmar, T., & Sachweh, P. (2022). Poison to the economy? Un-taxing the wealthy in the German Federal Parliament, 1996–2016. *Socio-Economic Review*, 20(4), 1111–1135. <https://doi.org/10.1093/ser/mwab015>
- Iversen, T., & Soskice, D. (2001). An asset theory of social policy preferences. *American Political Science Review*, 95(4), 875–893. <https://doi.org/10.1017/S0003055400400079>
- Iversen, T., & Soskice, D. (2015). Democratic limits to redistribution: Inclusionary versus exclusionary coalitions in the knowledge economy. *World Politics*, 67(2), 185–225. <https://doi.org/10.1017/S0043887114000350>
- Jonas Pontusson. (2024). The comparative politics of inequality and redistribution in liberal democracies. *Annual Review of Political Science*, 27, 85–108. <https://doi.org/10.1146/annurev-polisci-051321-043636>
- Kuusela, H. (2024). The hyperopia of wealth: The cultural legitimization of economic inequalities by top earners. *Socio-Economic Review*, 22(1), 1–25. <https://doi.org/10.1093/ser/mwad018>
- Lahat, L., & Sened, I. (2023). Behavioral knowledge for policy design: The connection between time uses behaviours and decision-making. *Social Policy & Administration*, 57(3), 511–527. <https://doi.org/10.1111/spol.12867>
- Lamont, M., & Pierson, P. (2019). Inequality generation and persistence as multidimensional processes. *Daedalus*, 148(3), 5–17. [https://doi.org/10.1162/daed\\_a\\_01749](https://doi.org/10.1162/daed_a_01749)

- Lamont, M., Beljean, S., & Clair, M. (2014). What is missing? Cultural processes and causal pathways to inequality. *Socio-Economic Review*, 12(3), 573–608. <https://doi.org/10.1093/ser/mwu011>
- Liu, A. H. (2023). Meritocracy as authoritarian co-optation: Political selection and upward mobility in China. *Comparative Political Studies*, 56(2), 210–235. <https://doi.org/10.1177/00104140221106428>
- Lu, Y., & Kandilov, I. (2022). Perceptions of upward social mobility and life satisfaction in China: The moderating role of internet use. *Social Indicators Research*, 164(3), 865–884. <https://doi.org/10.1007/s11205-022-02982-0>
- Massey, D. S., et al. (2014). Understanding inequality through the lens of cultural processes. *Social Science Research*, 47, 1–15. <https://doi.org/10.1016/j.ssresearch.2014.03.012>
- Munn, Z., Barker, T., Moola, S., Tufanaru, C., Stern, C., McArthur, A., Stephenson, M., and Aromataris, E. (2020). Methodological quality of case series studies: an introduction to the JBI critical appraisal tool. *JBI Evidence Synthesis*, DOI: 10.11124/JBISRIR-D-19-00099
- Ninan, J., & Sergeeva, N. (2022). Battle of narratives: Interaction between narratives and counter-narratives in megaprojects. *Project Management Journal*, 53(4), 380–396. <https://doi.org/10.1177/87569728221078416>
- Oliver, A. (2015). Nudging, shoving and budging: Behavioural economic policy frameworks. *Public Administration*, 93(3), 700–716. <https://doi.org/10.1111/padm.12165>
- Oliver, A. (2017). Nudges, shoves and budges: Behavioural economic policy frameworks. *Journal of Economic Policy Reform*, 20(3), 290–305. <https://doi.org/10.1080/17487870.2016.1216415>
- Oliver, A. (2021). Reciprocity and the art of behavioural public policy. *Behavioural Public Policy*, 5(4), 395–411. <https://doi.org/10.1017/bpp.2019.48>
- Oliver, A. (2024). Reflecting on reflection: Prospect theory, our behaviours, and our environment. *Behavioural Public Policy*, 8(2), 250–266. <https://doi.org/10.1017/bpp.2022.34>
- Oliver, A. (2024). On the importance of attempting to know the causes of things: A reply to Sunstein. *Behavioural Public Policy*, 8(1), 155–160. <https://doi.org/10.1017/bpp.2022.25>
- Pfotenhauer, S. M. (2022). The politics of scaling: Science and engineering in global policy. *Research Policy*, 51(6), 104554. <https://doi.org/10.1016/j.respol.2022.104554>
- Porter, E., & Tan, Y. (2023). Democracy and digitization: An introduction. *Democratization*, 30(4), 520–534. <https://doi.org/10.1080/13510347.2023.2177134>
- Reay, D. (2020). The perils and penalties of meritocracy: Sanctioning inequalities and legitimating prejudice. *British Journal of Sociology of Education*, 41(6), 783–798. <https://doi.org/10.1080/01425692.2020.1755227>
- Rao, M., et al. (2021). Perception of class mobility moderates the relationship between social class and prosocial behaviour. *Social Psychological and Personality Science*, 12(8), 1563–1573. <https://doi.org/10.1177/1948550620981712>
- Rethlefsen, M. & Page, M. (2022). PRISMA 2020 and PRISMA-S: common questions on tracking records and the flow diagram. *Journal of the Medical Library Association*. DOI: <dx.doi.org/10.5195/jmla.2022.1449>
- Sachweh, P. (2012). The welfare state and equality: Stratification realities and aspirations in three welfare regimes. *European Societies*, 14(1), 35–56. <https://doi.org/10.1080/14616696.2011.568252>
- Schmidt, A. (2021). From relational equality to personal responsibility. *Philosophy & Public Affairs*, 49(4), 387–421. <https://doi.org/10.1111/papa.12191>
- Schmidt, A., & Juijn, E. (2024). Economic inequality and the long-term future. *Futures*, 156, 103241. <https://doi.org/10.1016/j.futures.2023.103241>

Siqueira, A. C., & de Souza, F. (2017). Powered and disempowered by numbers: Data issues in global governance. *Global Policy*, 8(2), 225–238. <https://doi.org/10.1111/1758-5899.12416>

Snyder, H. (2023). Designing the literature review for a strong contribution. *Journal of Decision Systems*, DOI: 10.1080/12460125.2023.2197704

Sugden, R. (2023). Voluntariness and the bounds of cost–benefit analysis. *Economics and Philosophy*, 39(1), 29–51. <https://doi.org/10.1017/S0266267121000230>

Sumino, T. (2018). Identity priming and public opinion on income inequality. *Political Psychology*, 39(3), 519–535. <https://doi.org/10.1111/pops.12427>

Sunstein, C. (2024). Choice engines and paternalistic AI. *Humanities and Social Sciences Communication*. <https://doi.org/10.1057/s41599-024-03428-0>

Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: The final edition*. Yale University Press.

Tilly, C. (1998). *Durable inequality*. Princeton University Press.

Waitkus, N., & Wallaschek, S. (2022). Legitimate wealth: How wealthy business owners are portrayed in the press. *Journal of Economic Issues*, 56(2), 478–497. <https://doi.org/10.1080/00213624.2022.2057165>

Weiskopf, R., & Hansen, H. K. (2023). Algorithmic governmentality and the space of ethics: Examples from people analytics. *Organization*, 30(1), 3–25. <https://doi.org/10.1177/13505084211061894>

Xu, K. (2023). The negative effects of inequality on social capital. *Journal of Economic Surveys*, 37(2), 403–426. <https://doi.org/10.1111/joes.12568>