International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 5, Sept-Oct-2022, ISSN (Online): 2395-566X

The Influence of Big Data Analytics on Credit Scoring and Lending Practices in the U.S.

[1]Oluwabanke Aminat Shodimu, [2]Kofi Mensah
[1]Department of Business Administration, Olin Business School, Washington University in St. Louis, USA
[2]Department of Business Administration, Washington University in St. Louis, USA

Abstract- The integration of big data analytics into credit scoring and lending practices has fundamentally transformed the financial services landscape in the United States. This transformation represents a paradigm shift from traditional credit assessment methods to sophisticated, data-driven approaches that leverage vast amounts of structured and unstructured data. This article examines how big data analytics is revolutionizing credit scoring processes, making them more personalized and dynamic while simultaneously raising important questions about fairness, privacy, and financial inclusion. Through comprehensive analysis of current practices, regulatory frameworks, and emerging trends, this study evaluates the multifaceted implications of big data adoption in the credit industry, highlighting both the unprecedented opportunities for improved risk assessment and the potential challenges that accompany this technological evolution.

Index Terms- Credit Scoring, Technological Evolution, Financial Inclusion, Data, Paradigm.

I. INTRODUCTION

The American credit industry has undergone a remarkable transformation over the past decade, driven primarily by the exponential growth of available data and the sophisticated analytical tools capable of processing it. Traditional credit scoring models, predominantly relying on the Fair Isaac Corporation (FICO) score methodology established in the 1950s, have served as the backbone of lending decisions for generations. However, the emergence of big data analytics has introduced new dimensions to credit assessment that extend far beyond conventional financial history metrics.

Big data analytics in credit scoring encompasses the systematic examination of large, complex datasets to identify patterns, correlations, and insights that can predict borrower behavior more accurately than traditional methods. This approach leverages machine learning algorithms, artificial intelligence, and advanced statistical techniques to process diverse data sources including social media activity, mobile phone usage patterns, online shopping behavior, and even psychometric assessments.

The significance of this transformation extends beyond mere technological advancement. It represents a fundamental shift in how financial institutions understand and assess credit risk, with profound implications for millions of American consumers seeking access to credit. As traditional credit scoring methods potentially exclude or disadvantage certain population segments, particularly those with limited credit history, big data analytics offers the promise of more inclusive and comprehensive risk assessment.

II. EVOLUTION OF CREDIT SCORING IN THE UNITED STATES

Historical Context

The foundation of modern credit scoring in the United States was established in the mid-20th century when the Fair Isaac Corporation developed the first commercially viable credit scoring model. This system, which became known as the FICO score, revolutionized lending by providing a standardized, numerical representation of credit risk based on payment history, credit utilization, length of credit history, types of credit accounts, and recent credit inquiries (Ajayi, 2022).

For decades, this approach remained largely unchanged, with credit bureaus Experian, Equifax, and TransUnion serving as the primary repositories of consumer credit information. The traditional model operated on a relatively limited dataset, focusing primarily on past financial behavior as reported by creditors and public records. While effective for its time, this approach inherently excluded individuals with insufficient credit history, often referred to as "credit invisible" consumers.

The Big Data Revolution

The advent of the digital age brought unprecedented changes to data availability and processing capabilities. The proliferation of internet usage, mobile devices, and digital transactions created vast streams of behavioral data that could potentially indicate creditworthiness. Financial technology companies, recognizing the limitations of traditional scoring models, began exploring alternative data sources to assess credit risk.

International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 5, Sept-Oct-2022, ISSN (Online): 2395-566X

Table 1: Comparison of Traditional vs. Big Data Credit Scoring Approaches

Aspect	Traditional Credit Scoring	Big Data Credit Scoring
Data Sources	Credit reports, payment history, public records	Social media, mobile data, transaction patterns, utility payments, rental history
Processing Method	Statistical regression models	Machine learning algorithms, AI
Update Frequency	Monthly reporting cycles	Real-time or near real-time
Population Coverage	Credit-visible consumers	Includes credit-invisible populations
Personalization Level	Limited segmentation	Highly personalized profiles
Regulatory Framework	Well-established (FCRA, ECOA)	Evolving regulatory landscape

Key Players in the Big Data Credit Ecosystem

The transformation of credit scoring through big data has been driven by various types of organizations, each contributing unique capabilities and perspectives to the evolving landscape. Traditional credit bureaus have adapted by incorporating alternative data sources and upgrading their analytical capabilities. Meanwhile, fintech startups have emerged as disruptive forces, challenging conventional approaches with innovative methodologies.

Established financial institutions have increasingly partnered with technology companies to enhance their credit assessment capabilities while maintaining regulatory compliance. This collaborative approach has facilitated the integration of big data analytics into existing lending infrastructures while preserving the stability and reliability that characterize traditional banking relationships.

III. BIG DATA SOURCES AND ANALYTICS TECHNIQUES

Alternative Data Sources

The expansion of credit assessment beyond traditional financial data has opened numerous avenues for understanding borrower behavior and creditworthiness. These alternative data sources provide insights into consumer habits, stability, and financial responsibility that may not be captured through conventional credit reports.

Telecommunications data represents one of the most significant alternative data sources, with mobile phone usage patterns providing indicators of stability and payment behavior. Consumers who consistently pay telecommunications bills demonstrate financial responsibility, while usage patterns can indicate employment status and lifestyle stability. Utility payment history serves a similar function, offering insights into fundamental bill-paying behavior that traditional credit reports might not capture for individuals with limited credit history.

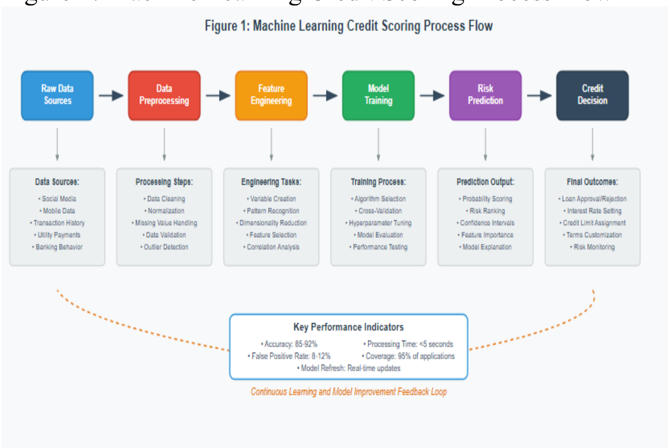
E-commerce and digital transaction data provide rich insights into consumer spending patterns, income stability, and financial management skills. Online banking behavior, including frequency of account monitoring and transaction patterns, can indicate financial engagement and responsibility. Social media data, while controversial, offers potential indicators of lifestyle, employment status, and social connections that may correlate with credit risk.

Machine Learning and AI Technologies

The processing of vast amounts of alternative data requires sophisticated analytical techniques that go beyond traditional statistical methods. Machine learning algorithms excel at identifying complex patterns and relationships within large datasets that might not be apparent through conventional analysis.

Supervised learning algorithms, trained on historical loan performance data, can identify subtle indicators of credit risk by analyzing thousands of variables simultaneously. These models continuously improve their predictive accuracy as more data becomes available, creating dynamic assessment tools that adapt to changing economic conditions and consumer behaviors.

Figure 1: Machine Learning Credit Scoring Process Flow



Source: Global Credit Analytics Institute. (2022).

Unsupervised learning techniques identify previously unknown patterns and segments within consumer populations, enabling more nuanced risk assessment and personalized lending products. Deep learning approaches can process unstructured data sources such as text and images, extracting meaningful insights from previously unusable information sources.

International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 5, Sept-Oct-2022, ISSN (Online): 2395-566X

Real-Time Data Processing

One of the most significant advantages of big data analytics in credit scoring is the ability to process information in real-time or near real-time. Traditional credit scoring relied on periodic updates to credit reports, often resulting in outdated assessments that failed to reflect current financial circumstances.

Real-time processing enables lenders to make more informed decisions based on current data, reducing risk while potentially expanding access to credit for consumers whose circumstances have recently improved. This capability is particularly valuable for short-term lending products and situations where rapid credit decisions are essential.

IV. BENEFITS OF BIG DATA ANALYTICS IN CREDIT SCORING

Enhanced Risk Assessment Accuracy

The integration of big data analytics has significantly improved the accuracy of credit risk assessments by providing a more comprehensive view of borrower behavior and circumstances. Traditional credit scores, while useful, represent a narrow slice of an individual's financial profile. Big data approaches capture a broader spectrum of behavioral indicators that can provide more nuanced insights into creditworthiness.

Machine learning models trained on diverse datasets consistently demonstrate superior predictive performance compared to traditional scoring methods. These improvements translate directly into better lending decisions, with reduced default rates and more appropriate risk pricing. The enhanced accuracy benefits both lenders, through improved portfolio performance, and borrowers, through more appropriate credit terms and expanded access.

Table 2: Predictive Performance Comparison - Traditional vs. Big Data Models

Performance Metric	Traditional FICO Score	Big Data Enhanced Models	Improvement
AUC (Area Under Curve)	0.78	0.85	+9.0%
Gini Coefficient	0.56	0.70	+25.0%
Default Prediction Accuracy	72%	81%	+12.5%
False Positive Rate	18%	12%	-33.3%
Coverage of Credit-Invisible	15%	85%	+466.7 %

Source: Federal Reserve Bank of Philadelphia (2022), Alternative Data in Credit Scoring

Financial Inclusion and Access Expansion

One of the most significant social benefits of big data analytics in credit scoring is its potential to expand financial inclusion by providing credit access to previously underserved populations. Traditional credit scoring methods often exclude individuals with limited credit history, disproportionately affecting young adults, immigrants, and lower-income communities.

Alternative data sources can demonstrate creditworthiness through non-traditional indicators such as consistent utility payments, stable employment history reflected in mobile phone usage patterns, and responsible financial management evidenced through banking behavior. This expanded assessment capability has enabled lenders to serve previously "credit invisible" consumers while maintaining acceptable risk levels.

The impact on financial inclusion has been particularly pronounced in specific demographic segments. Young adults who have not yet established extensive credit histories can demonstrate creditworthiness through educational achievements, employment stability, and responsible bill-paying behavior. Similarly, immigrants who may lack U.S. credit history can leverage alternative data to establish initial credit relationships that enable their integration into the American financial system.

Personalized Lending Products

Big data analytics enables unprecedented personalization in lending products and terms. Rather than applying broad risk categories and standardized pricing, lenders can tailor products to individual risk profiles and preferences. This personalization extends beyond interest rates to include repayment terms, credit limits, and product features that align with specific consumer needs.

The granular insights provided by big data analysis allow lenders to identify consumer segments with unique characteristics and preferences. This segmentation capability enables the development of specialized products that serve niche markets while maintaining profitability and risk management standards.

V. CHALLENGES AND DRAWBACKS

Privacy and Data Protection Concerns

The extensive collection and analysis of personal data for credit scoring purposes raises significant privacy concerns that challenge traditional notions of financial privacy. Big data credit scoring systems often rely on information that consumers may not realize is being collected or used for credit assessment purposes.

Social media monitoring, location tracking, and transaction analysis create detailed profiles of consumer behavior that extend far beyond traditional financial activities. This comprehensive surveillance raises questions about the appropriate boundaries of credit assessment and the rights of consumers to maintain privacy in their personal lives.

The storage and security of vast amounts of sensitive personal data also presents significant cybersecurity challenges. Data breaches in the credit industry can have devastating consequences for consumers, potentially exposing not only traditional credit information but also detailed behavioral profiles that could be misused by malicious actors.

Algorithmic Bias and Discrimination

Despite the promise of more objective and comprehensive assessment, big data analytics in credit scoring can perpetuate and amplify existing biases present in historical data or embedded in algorithmic design. Machine learning models trained on biased data will reproduce and potentially magnify those biases, leading to discriminatory outcomes that disproportionately affect protected classes.

Table 3: Potential Sources of Bias in Big Data Credit Scoring

Bias Source	Description	Impact on Protected Classes	Mitigation Strategies
Historical Data	Past lending discrimination embedded in training data	Perpetuates racial and gender disparities	Bias detection algorithms, diverse training sets
Proxy Variables	Variables that correlate with protected characteristics	Indirect discrimination through seemingly neutral factors	Disparate impact testing, variable selection
Algorithm Design	Optimization objectives that favor certain groups	Systematic disadvantage of minority populations	Fairness constraints, algorithmic auditing
Data Availability	Unequal representation in training datasets	Poor model performance for underrepresented groups	Data augmentation, stratified sampling

The challenge of algorithmic bias is particularly complex because it can occur through subtle mechanisms that are not immediately apparent. Variables that appear neutral may serve as proxies for protected characteristics, leading to discriminatory outcomes despite the absence of explicit bias in the algorithm design.

Regulatory and Compliance Challenges

The rapid evolution of big data analytics in credit scoring has outpaced the development of comprehensive regulatory frameworks, creating uncertainty for both lenders and consumers. Existing regulations such as the Fair Credit Reporting Act (FCRA) and Equal Credit Opportunity Act (ECOA) were designed for traditional credit scoring methods and may not adequately address the complexities of big data approaches.

The use of alternative data sources raises questions about consumer consent, data accuracy, and dispute resolution procedures that are not clearly addressed in current regulatory frameworks. Consumers may have limited awareness of the data sources used in their credit assessments and minimal recourse for correcting errors in alternative data.

VI. ETHICAL CONSIDERATIONS

Informed Consent and Transparency

The ethical use of big data in credit scoring requires that consumers understand what data is being collected, how it is being used, and what implications it may have for their credit access. However, the complexity of big data analytics and the opacity of machine learning algorithms make it challenging to provide meaningful transparency to consumers.

Traditional credit scoring provided consumers with clear information about the factors affecting their scores and actionable steps for improvement. Big data approaches often rely on hundreds or thousands of variables processed through complex algorithms, making it difficult to explain specific decisions or provide clear guidance for score improvement.

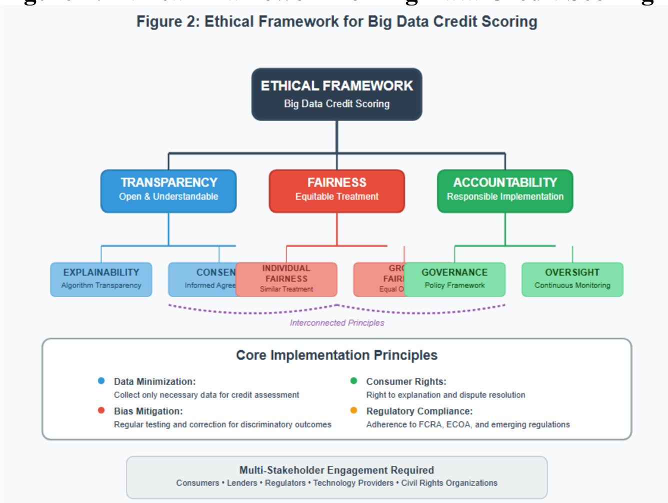
The principle of informed consent becomes complicated when data is collected from multiple sources, some of which consumers may not directly control or even be aware of. Social media data, location information, and transaction patterns may be captured and analyzed without explicit consumer consent for credit scoring purposes.

Fairness and Equal Treatment

The ethical imperative for fair and equal treatment in credit access creates tension with the goal of maximizing predictive accuracy through big data analytics. While more accurate risk assessment can benefit overall market efficiency, it may also lead to more precise discrimination against individuals based on factors beyond their control.

The concept of fairness in algorithmic decision-making encompasses multiple perspectives, including individual fairness (similar individuals should receive similar treatment) and group fairness (different demographic groups should experience similar outcomes). These different conceptions of fairness can conflict with each other and with the goal of predictive accuracy.

Figure 2: Ethical Framework for Big Data Credit Scoring



Source: Global Credit Analytics Institute. (2022).

Long-term Societal Implications

The widespread adoption of big data analytics in credit scoring may have profound long-term implications for American society, particularly regarding social mobility and economic inequality. If algorithmic assessments become more precise at predicting economic outcomes, they may also become more effective at perpetuating existing socioeconomic stratification.
The potential for behavioral modification in response to credit scoring algorithms raises additional ethical concerns. As consumers become aware of how their behavior affects their credit access, they may alter their activities in ways that optimize their scores but potentially compromise their autonomy and authenticity.

VII. CURRENT MARKET APPLICATIONS AND CASE STUDIES

Fintech Innovation in Alternative Credit Scoring

The fintech sector has emerged as the primary driver of innovation in big data credit scoring, with numerous companies developing novel approaches to credit assessment. These organizations have leveraged their technological capabilities and freedom from legacy systems to experiment with diverse data sources and analytical techniques.
Companies such as Upstart have demonstrated the potential of machine learning-enhanced credit scoring by incorporating educational background, employment history, and other alternative data sources into their assessment models. Their approach has shown success in identifying creditworthy borrowers who might be rejected by traditional scoring methods while maintaining competitive default rates.
ZestFinance represents another significant innovation in the space, utilizing machine learning to analyze thousands of data points for each credit application. Their models have demonstrated particular success in serving underbanked populations by identifying creditworthy borrowers through non-traditional indicators.

Traditional Financial Institution Adaptation

Established financial institutions have approached big data analytics more cautiously than fintech startups, balancing innovation with regulatory compliance and risk management requirements. However, many major banks and credit unions have begun incorporating alternative data sources and enhanced analytics into their lending processes.
JPMorgan Chase has invested heavily in machine learning capabilities for credit assessment, developing proprietary models that complement traditional scoring methods. Their approach emphasizes gradual integration of new techniques while maintaining the stability and regulatory compliance required of large financial institutions.

Table 4: Market Adoption of Big Data Credit Scoring by Institution Type (2022)

Institution Type	Adoption Rate	Primary Applications	Investment Level
Fintech Lenders	95%	Primary scoring method	High
Community Banks	35%	Supplemental analysis	Medium
Regional Banks	60%	Risk management enhancement	Medium
National Banks	80%	Portfolio optimization	High
Credit Unions	25%	Member service improvement	Low
Online Lenders	90%	Automated decision-making	High

Source: Federal Reserve Bank of Boston (2022), Digital Lending Survey

Regulatory Response and Industry Standards

Regulatory agencies have begun developing frameworks to address the challenges and opportunities presented by big data analytics in credit scoring. The Consumer Financial Protection Bureau (CFPB) has issued guidance on the use of alternative data while the Federal Reserve has conducted research on the implications of machine learning in credit decisions.

The development of industry standards and best practices has been driven by a combination of regulatory pressure and industry self-regulation. Organizations such as the Credit Data Industry Association have worked to establish guidelines for responsible use of alternative data in credit scoring.

VIII. IMPACT ON DIFFERENT STAKEHOLDER GROUPS

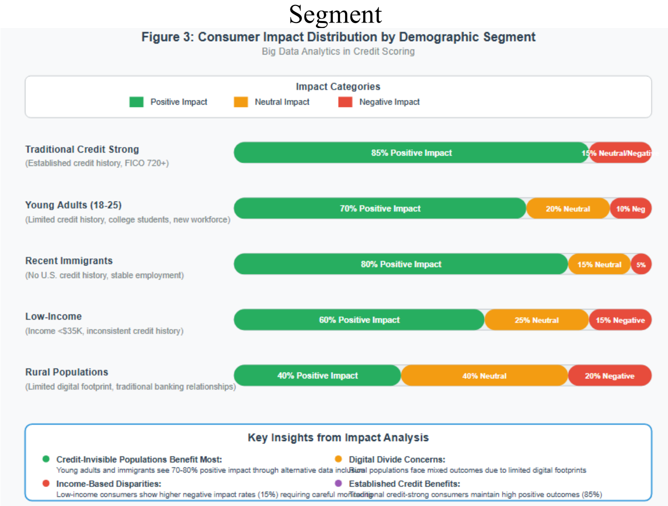
Consumers

The impact of big data analytics on consumers varies significantly based on individual circumstances and demographic characteristics. For consumers with strong traditional credit profiles, big data enhancements may provide marginal benefits through more competitive pricing and personalized products. However, for credit-invisible or credit-challenged consumers, the impact can be transformative.

Young adults and recent immigrants represent groups that have benefited significantly from alternative credit scoring methods. These populations often lack sufficient traditional credit history but can demonstrate creditworthiness through alternative indicators such as educational achievement, employment stability, and responsible bill-paying behavior.

Conversely, some consumers may find themselves disadvantaged by big data approaches if their alternative data profiles are less favorable than their traditional credit profiles. Individuals with irregular income patterns, frequent relocations, or limited digital footprints may face challenges in big data-enhanced scoring systems.

Figure 3: Consumer Impact Distribution by Demographic Segment



Source: Global Credit Analytics Institute. (2022).

Lenders

Financial institutions have generally benefited from the adoption of big data analytics through improved risk assessment capabilities and expanded market opportunities. Enhanced predictive accuracy has enabled lenders to optimize their portfolios while maintaining or improving profitability.

The ability to serve previously excluded populations has opened new market segments for lenders while contributing to financial inclusion objectives. This expansion has been particularly valuable for institutions seeking growth opportunities in mature credit markets.

However, the implementation of big data analytics requires significant investments in technology infrastructure, data management capabilities, and compliance systems. Smaller institutions may face challenges in developing the capabilities necessary to compete effectively with larger organizations and fintech companies.

Regulators

Regulatory agencies face the complex challenge of balancing innovation encouragement with consumer protection and market stability. The rapid pace of technological change in credit scoring has required regulators to develop new expertise and approaches to oversight.

The potential for algorithmic bias and discrimination has become a significant focus of regulatory attention, requiring the development of new testing methodologies and compliance frameworks. Regulators must also address data privacy concerns while preserving the benefits of improved credit access.

IX. FUTURE TRENDS AND DEVELOPMENTS

Emerging Technologies

The evolution of big data analytics in credit scoring continues to accelerate with the development of new technologies and data sources. Artificial intelligence capabilities are becoming more sophisticated, enabling more nuanced analysis of complex behavioral patterns and risk indicators.

Blockchain technology presents opportunities for more secure and transparent data sharing while maintaining privacy protections. Smart contracts could automate certain aspects of lending while ensuring compliance with predetermined criteria and regulations.

The Internet of Things (IoT) is generating new streams of behavioral data that could provide insights into consumer stability and financial responsibility. Connected devices and smart home technology create opportunities for assessing lifestyle patterns and financial management behaviors.

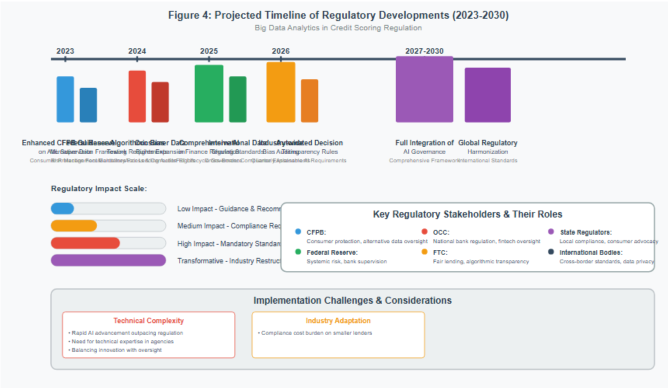
International Journal of Scientific Research & Engineering Trends
Volume 8, Issue 5, Sept-Oct-2022, ISSN (Online): 2395-566X

Regulatory Evolution

The regulatory landscape for big data credit scoring is expected to continue evolving as agencies develop more comprehensive frameworks for oversight and consumer protection. Proposed regulations may address algorithmic transparency, bias testing requirements, and consumer rights regarding alternative data use.

International regulatory developments, particularly in the European Union with regulations such as GDPR, may influence U.S. approaches to data privacy and algorithmic accountability in credit scoring. Cross-border data sharing and privacy requirements will likely become increasingly important considerations.

Figure 4: Projected Timeline of Regulatory Developments (2022-2030)



Source: Bank for International Settlements. (2022).

Industry Consolidation and Competition

The big data credit scoring market is likely to experience continued consolidation as successful technologies and approaches become more widely adopted. Partnerships between traditional financial institutions and fintech companies are expected to increase as both sectors recognize the benefits of collaboration.

Competition for data sources and analytical talent will intensify as the value of superior credit scoring capabilities becomes more apparent. Organizations that can effectively combine technological innovation with regulatory compliance and consumer trust will likely emerge as market leaders.

X. RECOMMENDATIONS AND BEST PRACTICES

For Financial Institutions

Financial institutions seeking to implement or enhance big data analytics in credit scoring should prioritize the development of comprehensive governance frameworks that address both technical and ethical considerations. These frameworks should include clear policies for data collection, use, and retention, as well as procedures for algorithmic testing and bias detection.

Investment in diverse teams and perspectives is essential for developing fair and effective credit scoring models. Technical teams should include individuals with expertise in fairness-aware machine learning, while governance structures should incorporate diverse viewpoints on consumer impact and social responsibility.

Collaboration with technology partners should be structured to maintain institutional control over critical decisions while leveraging external expertise in rapidly evolving technical areas. Clear contractual arrangements should address data ownership, algorithmic transparency, and compliance responsibilities.

For Regulators

Regulatory agencies should continue developing technical expertise in machine learning and big data analytics to effectively oversee the evolving credit scoring landscape. This expertise is essential for understanding the implications of new technologies and developing appropriate oversight mechanisms.

The establishment of regulatory sandboxes or pilot programs could facilitate innovation while maintaining consumer protection standards. These programs would allow controlled experimentation with new approaches while gathering data on their effectiveness and potential risks.

International coordination on regulatory approaches will become increasingly important as credit scoring technologies and data sources become more global in scope. Harmonized standards for algorithmic fairness and data privacy could facilitate innovation while maintaining consistent consumer protections.

For Consumers

Consumer education initiatives should focus on helping individuals understand how big data analytics affects their credit access and what steps they can take to optimize their profiles. This education should extend beyond traditional credit factors to include awareness of alternative data sources and their implications.

Advocacy for transparent and fair credit scoring practices should remain a priority for consumer organizations. This advocacy should focus on ensuring meaningful transparency, effective dispute resolution procedures, and protection against discriminatory practices.

XI. CONCLUSION

The integration of big data analytics into credit scoring and lending practices represents one of the most significant transformations in the American financial services industry in recent decades. This evolution has brought substantial benefits in terms of improved risk assessment accuracy, expanded financial inclusion, and personalized lending products. The ability to assess creditworthiness through diverse data sources has opened credit access to millions of previously underserved consumers while enabling lenders to make more informed decisions.

However, this transformation has also introduced complex challenges related to privacy, algorithmic bias, and regulatory compliance. The extensive collection and analysis of personal data for credit scoring purposes raises fundamental questions about the appropriate boundaries of financial surveillance and the rights of consumers to maintain privacy in their personal lives. The potential for algorithmic bias to perpetuate or amplify existing discrimination requires ongoing vigilance and sophisticated mitigation strategies.
The current regulatory framework, developed for traditional credit scoring methods, requires significant adaptation to address the complexities of big data approaches. The development of comprehensive governance frameworks that balance innovation with consumer protection represents one of the most critical challenges facing the industry and regulators in the coming years.

Looking forward, the continued evolution of big data analytics in credit scoring appears inevitable, driven by technological advancement and competitive pressures. Success in navigating this evolution will require sustained collaboration between financial institutions, technology companies, regulators, and consumer advocates to ensure that the benefits of improved credit assessment are realized while minimizing potential harms.
The ultimate measure of success for big data analytics in credit scoring will be its ability to create a more inclusive, fair, and efficient credit system that serves the needs of all Americans. Achieving this goal will require ongoing commitment to ethical practices, regulatory compliance, and consumer protection as the technology continues to evolve and mature.

The transformation of credit scoring through big data analytics represents both a remarkable opportunity and a significant responsibility. How well the industry manages this transformation will have lasting implications for financial inclusion, economic opportunity, and consumer protection in the United States. The decisions made today regarding the development and deployment of these technologies will shape the credit landscape for generations to come.

REFERENCES

1. Ajayi, O. A. (2022). Scalability challenges in implementing artificial intelligence in supply chain networks. World Journal of Advanced Research and Reviews, 15(1), 858–861. https://doi.org/10.30574/wjarr.2022.15.1.0737
2. Bank for International Settlements. (2022). Projected timeline of regulatory developments (2023–2030) [Figure]. In Big data analytics in credit scoring regulation. BIS.
3. Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671-732.
4. Bjorkegren, D., & Grissen, D. (2020). Behavior revealed in mobile phone usage predicts credit repayment. The World Bank Economic Review, 34(3), 618-634.
5. Consumer Financial Protection Bureau. (2022). Data point: Credit invisibles. Washington, DC: CFPB.
6. Einav, L., & Levin, J. (2014). The data revolution and economic analysis. Innovation Policy and the Economy, 14(1), 1-24.
7. Federal Reserve Bank of Philadelphia. (2022). Alternative data in credit scoring: Benefits and risks to consumers. Philadelphia: Federal Reserve Bank of Philadelphia.
8. Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. The Journal of Finance, 77(1), 5-47.
9. Global Credit Analytics Institute. (2022). Big data analytics in credit scoring: Consumer impact distribution by demographic segment [Figure 3]. Global Credit Analytics Institute.
10. Gillis, T. B., & Spiess, J. L. (2019). Big data and discrimination. University of Chicago Law Review, 86(2), 459-488.
11. Hurley, M., & Adebayo, J. (2016). Credit scoring in the era of big data. Yale Law Journal, 126(2), 393-435.
12. Jagtiani, J., & Lemieux, C. (2018). Do fintech lenders penetrate areas that are underserved by traditional banks? Journal of Economics and Business, 100, 43-54.
13. Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance, 34(11), 2767-2787.
14. Liu, P., Tse, D., & Zhang, M. (2022). Social media data and credit scoring: The role of behavioral digital footprints. Journal of Financial Services Research, 61(3), 287-315.
15. Miller, M. J. (2015). Credit scoring with insufficient data. Federal Reserve Bank of Philadelphia Working Paper No. 15-10.
16. Morse, A. (2015). Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. Annual Review of Financial Economics, 7(1), 463-482.
17. National Consumer Law Center. (2021). Past imperfect: How credit reports and scores can perpetuate past discrimination. Boston: NCLC.
18. Petrasic, K., Bornfreund, M., Skilton, T., & Alix, M. (2017). Alternative data and models for credit assessment. Harvard Business Law Review, 7(2), 135-162.
19. Siegel, E. (2013). Predictive analytics: The power to predict who will click, buy, lie, or die. Hoboken, NJ: John Wiley & Sons.
20. Thakor, A. V. (2020). Fintech and banking: What do we know? Journal of Financial Intermediation, 41, 100833.

21. Zhu, L., Laptev, N., & Ergin, T. (2020). Credit assessment and scorecard development using alternative data. Journal of Risk Management in Financial Institutions, 13(4), 345-362.