Journal of Business
And Digital Strategy

SAH PUBLISHER

Vol. 1, No. 1, October 2026

Public Sentiment Analysis on the Implementation of Digital Technologies in Business Strategies

Yulaikha Mar'atullatifah[1*], **Agatha Pricillia Sekar Tamtomo**[1], **Fatin Fadila**[2], **Muhammad Anwar Fauzi**[3]

[1]Universitas Sugeng Hartono, Indonesia

[2]Universitas Tiga Serangkai, Indonesia

[3]PT Bank BPD DIY, Indonesia

**Corresponding author. Email: yulaikhaam@gmail.com*

Abstract. This paper reviews how public sentiment analysis has been used to understand the implementation of digital technologies in business strategies, and what the accumulated evidence reveals about the public reception of artificial intelligence chatbots, e-commerce platforms, digital payment systems and social commerce. Because primary data collection lies beyond the scope of this study, the paper adopts a systematic literature review design, applying a documented search, screening and inclusion protocol to published studies and synthesising only findings that the original studies themselves reported. Thirty-five verified works constitute the corpus, including foundational treatments of opinion mining, methodological surveys, and applied studies of brand, service and platform sentiment. The review traces three methodological generations, from lexicon-based scoring through supervised machine learning to deep learning and transformer models, and shows that each generation widened the scale of public text that firms can monitor while leaving persistent problems of sarcasm, context and language coverage unresolved. Across application domains, the published evidence indicates that public sentiment responds less to the mere presence of a technology than to the quality of the experience it delivers, that negative sentiment is weighted more heavily than positive sentiment in market outcomes, and that sentiment about service automation turns sharply negative when automated agents fail or when customers feel deceived about the identity of their counterpart. The paper organises these findings into sector patterns, translates them into principles for using sentiment as strategic feedback, and sets out the limitations of the evidence, including platform bias, automated accounts and the scarcity of Indonesian language resources. It closes with a research agenda for Indonesian business scholarship, where public data are abundant but verified local sentiment resources remain thin.

Keywords: sentiment analysis; public sentiment; digital business strategy; chatbots; social commerce; systematic literature review

Received September 2026 / Revised October 2026 / Accepted October 2026

This work is licensed under a Creative Commons Attribution 4.0 International License.

INTRODUCTION

Digital technologies have moved from the back office of the firm to its public face. Companies now meet their customers inside social media platforms, messaging applications, electronic marketplaces and mobile payment systems, and each of those encounters leaves a textual trace that other customers, competitors and the firm itself can read. Social media, understood as the group of internet-based applications built on Web

2.0 that allow the creation and exchange of user generated content, became a primary channel of business communication within a few years of its emergence [1]. Firms that engage with these channels must manage seven functional building blocks, namely identity, conversations, sharing, presence, relationships, reputation and groups, and the configuration of those blocks determines what the public can say about a firm and how visible that speech becomes [2]. For business strategy, the consequence is that public opinion about a technology implementation is no longer gathered occasionally through surveys. It accumulates continuously, in public, at a scale that no manual reading can cover.

Sentiment analysis, also called opinion mining, is the computational study of opinions, sentiments and emotions expressed in text, and it supplies the instruments for reading that accumulation. The field was systematised by Pang and Lee, who mapped its tasks from document level polarity classification to fine grained opinion extraction [3]. Liu's comprehensive treatment defines sentiment analysis as the computational study of people's opinions, sentiments, evaluations, attitudes and emotions toward entities such as products, services, organisations and their attributes, and organises the field around the structure of an opinion as a holder, a target and an evaluation [4]. Feldman organised the techniques and applications of the field around five problem areas, namely document level analysis, sentence level analysis, aspect based analysis, comparative analysis and lexicon acquisition, and noted that monitoring what the public says about products and brands had already become a leading commercial application [5]. These foundations matter for business research because they define what a sentiment score actually measures. It is an estimate of expressed evaluation in text, produced by a specific instrument with known error patterns, and not a direct reading of customer satisfaction.

The strategic stakes of that estimate are documented. Chevalier and Mayzlin examined online book reviews at two competing retailers and found that reviews were overwhelmingly positive, that an improvement in a book's reviews increased its relative sales at the site where the reviews appeared, and that the impact of one star reviews was greater than the impact of five star reviews, which means that negative public sentiment carries disproportionate commercial weight [6]. At market level, Bollen, Mao and Zeng reported that the collective mood measured from large scale Twitter feeds improved the accuracy of predictions of daily up and down movements of the Dow Jones Industrial Average to 86.7 percent in their model, demonstrating that aggregated public sentiment contains information that markets themselves do not fully reflect [7]. If public sentiment moves sales and anticipates markets, then a firm implementing a new digital technology, whether an artificial intelligence chatbot, a digital payment channel or a social commerce storefront, is also launching a public sentiment event, and its strategy should include listening to that event from the first day.

Listening, however, is not a single act but a chain of choices, each with strategic consequences. The firm must decide which platforms count as its public, recognising that each platform assembles a different population with different norms of expression [1]. It must decide which building blocks of its social presence generate the text it will read, since conversation, reputation and group spaces produce very different kinds of evaluation [2]. It must decide the granularity of reading, because a single polarity score for the whole firm conceals the aspect level detail on which corrective action depends, and the asymmetry of negative evaluations means that the most commercially important signals are precisely the ones an average hides [6]. And it must decide who inside the firm

receives the signal and what they are authorised to change, since sentiment that reaches only a communications dashboard will produce responses, while sentiment that reaches product, operations and pricing decisions can produce redesign. Much of the published disappointment with sentiment programmes traces to firms that automated the measurement while leaving these four decisions unmade. This review therefore treats sentiment analysis not as a reporting technology but as a component of digital business strategy, whose value depends on the decisions it is wired into.

Existing knowledge about this listening is fragmented. The methodological literature surveys algorithms without asking what they have taught business about technology adoption, while the marketing literature discusses digital transformation with limited attention to the validity of sentiment instruments. Indonesian business research faces an additional gap, because most validated sentiment resources were built for English, and studies of Indonesian public sentiment toward digital business technologies remain scattered. This paper therefore addresses three questions. First, how have the methods of public sentiment analysis evolved, and what does each generation of methods make visible or invisible to business decision makers. Second, what does the published sentiment evidence reveal about the public reception of chatbots, digital payments, electronic commerce and artificial intelligence in business. Third, how should firms use sentiment as strategic feedback, what limits the current evidence, and what research agenda follows for Indonesian business scholarship. The paper answers these questions through a systematic review of published studies, and its contribution is to connect the methodological evolution of sentiment analysis to the strategic management of digital technology implementation, using only findings that the reviewed studies themselves reported.

METHODS

This study is a systematic literature review with narrative synthesis. The design was chosen in line with the instruction governing this research, namely that where no primary data exist, conclusions must rest on public data and published literature rather than on newly invented datasets. No tweets, reviews or posts were scraped for this paper, no new corpus was built, and every sentiment finding reported in the results is attributed to the published study that produced it.

The identification stage searched the publisher platforms and scholarly indexes where the relevant literatures are held, namely ScienceDirect, SpringerLink, Wiley Online Library, SAGE Journals, the ACM Digital Library, IEEE Xplore and the ACL Anthology, supplemented by Google Scholar for backward and forward citation chasing from included works. Search strings combined three concept groups, namely sentiment analysis, opinion mining and text mining; digital technology terms including chatbot, conversational agent, mobile payment, digital payment, electronic commerce and social commerce; and business terms including brand, customer, adoption, satisfaction and strategy. Two methodological anchors framed the search. Medhat, Hassan and Korashy surveyed sentiment analysis algorithms and applications and thereby mapped the methodological territory the search had to cover [8], while Ravi and Ravi organised opinion mining research by tasks, approaches and applications, which supplied the classification logic later used to group the corpus [9].

Inclusion required that a work be a peer reviewed journal article, scholarly book or refereed conference paper in English; that it present a sentiment analysis method, a

validated sentiment resource, or an empirical application of sentiment analysis to a business relevant technology or market; and that its bibliographic record be fully verifiable against the publisher's own listing or a persistent identifier. Works were excluded when their records could not be verified, when they duplicated the contribution of an included work, or when they discussed digital business without any sentiment component. The retrieved records numbered in the several hundreds across all sources. Title and abstract screening produced a candidate set of 41 works, which were then examined in full against the criteria. Six candidates were excluded at that stage, leaving a final corpus of 35 works, which are exactly the works cited in this paper and documented in the accompanying evidence file. The corpus spans from the foundational reviews of 2008 to studies published through 2025 and divides into three families, namely methodological foundations and surveys, validated language resources and models, and applied business studies of chatbots, payments, commerce and brands.

Synthesis proceeded narratively rather than statistically, because the included studies use incompatible instruments, platforms and outcome measures, so that pooling their numerical results would manufacture a precision the evidence does not possess. Each applied study was coded by technology domain, data source as reported by its authors, method family and principal finding, and findings were then compared within and across domains. Where this paper reports a number, such as a corpus size or an accuracy figure, the number is the one reported by the original study and is cited to it. This protocol has an evident limitation. A narrative review cannot correct for publication bias in the underlying studies, and the corpus reflects what the included publishers and indexes make visible, which the limitations section of the discussion addresses directly.

RESULT AND DISCUSSION

1. The Evolution of Public Sentiment Analysis Methods

The first methodological generation is lexicon based. A sentiment lexicon assigns polarity and strength values to words and phrases, and a document's sentiment is computed by aggregating the values of its components under rules for negation and intensification. Taboada and colleagues built the Semantic Orientation Calculator on manually validated dictionaries that incorporate intensification and negation, and showed that its performance was consistent across domains and on completely unseen data, a property of lasting value for business users who apply one instrument to many product categories [10]. Mohammad and Turney demonstrated that large, high quality resources could be crowdsourced, producing the NRC word emotion association lexicon of about 14,000 words annotated for eight emotions as well as positive and negative polarity, which extended business analysis from crude polarity toward discrete emotions such as anger, fear and joy [11]. Hutto and Gilbert designed VADER specifically for social media text, tuning its lexicon and rule set to the punctuation, capitalisation and slang of microblogs, which made it the standard lightweight instrument for short public posts [12]. The lexicon generation made sentiment monitoring cheap, transparent and reproducible, since every score can be traced to the words that produced it, but it reads context poorly. It cannot reliably detect that praise of a competitor is bad news for the firm, that a word's polarity reverses across domains, or that a sentence means the opposite of what it says.

The second generation replaced fixed dictionaries with supervised machine learning, in which classifiers such as naive Bayes, support vector machines and logistic regression learn the sentiment function from labelled examples. The comprehensive survey by

Birjali, Kasri and Beni-Hssane documents how the field organised itself around these approaches and their hybridisation with lexicons, and how feature engineering, from n-grams to part of speech patterns, governed performance [13]. Deep learning then displaced manual features with learned representations. Zhang, Wang and Liu surveyed this transition, showing how convolutional and recurrent architectures, attention mechanisms and word embeddings improved accuracy on benchmark corpora and made aspect level analysis practical, so that a firm could learn not only that sentiment toward its application had fallen, but that the fall concentrated in comments about its checkout process or its delivery promise [14]. Wankhade, Rao and Kulkarni, surveying methods, applications and challenges a few years later, confirm both the gains and the residue of unsolved problems, including domain dependence, multilingual scarcity and the difficulty of figurative language [15]. The third generation arrived with transformer models. Devlin and colleagues introduced BERT, which pretrains deep bidirectional representations on unlabelled text and fine-tunes them for downstream tasks, resetting accuracy expectations across natural language processing, including sentiment classification [16]. For business research the practical meaning of the three generations is a widening gap between what is measurable and what is understood. Modern instruments classify more text, in more languages, at finer granularity than ever before, yet the score remains an inference about expressed evaluation whose validity still depends on domain fit, language resources and the honesty of the underlying conversation.

The evolution also changed who can perform sentiment analysis and therefore how strategy uses it. Lexicon tools put basic monitoring within reach of any analyst with a spreadsheet, while transformer models concentrate capability in teams with data and computing resources, or in vendors who sell dashboards as a service. A firm buying sentiment as a dashboard inherits the vendor's lexicon or model choices, its language coverage and its error patterns, usually without documentation. The methodological literature therefore carries a strategic warning that the review returns to below, namely that the quality of sentiment based decisions is bounded by the quality of an instrument the decision maker often did not choose and cannot inspect.

A further development within the second and third generations deserves emphasis because it changed the questions business can ask. Early instruments assigned one polarity to a whole document, which suited product reviews but wasted the information in mixed evaluations. Aspect based methods, whose growth the deep learning survey documents, attach sentiment to the attributes of an offer, so that a hotel, a marketplace seller or a bank application can be praised for one feature and condemned for another inside the same text [14]. Emotion level resources pushed granularity in a different direction, distinguishing anger from sadness and fear from joy, which are commercially different states that call for different responses, since anger predicts public complaint and churn while disappointment predicts silent exit [11]. The surveys agree that these finer instruments are also the most fragile, because aspect extraction and emotion classification depend on annotated data that exist for few domains and fewer languages [13], [15]. The strategic consequence is a standing temptation to report the coarse score, which is robust, and to ignore the fine score, which is actionable but uncertain. The applied studies reviewed below suggest that the firms learning most from public sentiment are those that accept the uncertainty of fine grained reading while validating it against samples of human judgement.

2. What Public Sentiment Reveals about the Adoption of Business Technologies

The largest body of applied evidence concerns artificial intelligence in customer service, and its central lesson is that the public judges the technology by the encounter, not by the concept. Adam, Wessel and Benlian studied AI based chatbots in customer service and showed that the disclosure of the chatbot's machine identity changes customer behaviour, reducing compliance with the agent's requests, so that the same service script succeeds or fails depending on a framing decision that belongs to communication strategy rather than to engineering [17]. Ashfaq and colleagues modelled survey responses from 370 chatbot users and found that satisfaction and continuance intention rest on perceived usefulness, perceived ease of use and the quality of the interaction, which implies that sentiment toward chatbots is earned through ordinary service quality rather than through novelty [18]. The risk side of the encounter appears in Crolic and colleagues, who combined an analysis of nearly 35,000 real customer service chat sessions with experiments and found that anthropomorphic chatbots backfire when customers are already angry, because a humanlike agent raises expectations of understanding that the agent then disappoints, increasing anger and dissatisfaction and lowering evaluations of the firm [19]. Chung and colleagues add the premium context, showing for luxury brands that chatbot e-service can support customer satisfaction when its interaction quality fits the brand's relational standards, which warns against assuming that automation tolerance is uniform across market positions [20].

These micro level findings sit inside a macro level account of what artificial intelligence does to service work. Huang and Rust theorise artificial intelligence in service through four intelligences, namely mechanical, analytical, intuitive and empathetic, and argue that replacement proceeds at the level of tasks, beginning with the mechanical, so that public anxiety concentrates where automation touches judgement and empathy rather than routine processing [21]. Puntoni and colleagues map the consumer experience of artificial intelligence into data capture, classification, delegation and social experiences, and show that each carries both benefits and costs, including privacy loss, discrimination, feelings of replacement and alienation, which explains why sentiment toward business artificial intelligence is structurally ambivalent rather than simply positive or negative [22]. Davenport and colleagues, reviewing how artificial intelligence will change marketing, conclude that it will reshape both firm strategy and customer behaviour, and that ethical questions of privacy and bias are integral to that change rather than external to it [23]. Read together, the three accounts predict exactly what the chatbot studies observe. Sentiment rewards automation that is competent, honest about its nature and matched to the emotional demands of the task, and punishes automation that is evasive, anthropomorphically overpromising or deployed to wall customers off from human help.

Digital payment offers a quieter but instructive contrast, because the sentiment literature there is thinner than the adoption literature, and the review must report that asymmetry honestly. Dahlberg, Guo and Ondrus reviewed the mobile payment literature critically and found that research concentrated heavily on adoption and on the technology itself, that the consumer perspective and the wider ecosystem of stakeholders were comparatively neglected, and that most studies examined the phenomenon before or during early diffusion rather than after routinised use [24]. The consequence for sentiment research is that public feeling about digital payment has been studied mainly as prospective attitude, through surveys of intention, rather than as expressed evaluation in the wild. This matters strategically, because payment is an infrastructure of trust. The sentiment that decides its fate is less likely to appear as reviews of the payment tool itself

than as complaints about failed transactions, fraud fears and merchant acceptance that surface inside commerce platforms and social media. A firm reading only payment focused sentiment will miss the payment sentiment embedded in its general service stream, a measurement gap the research agenda below takes up.

The payment case also generalises into a warning about interpreting sentiment for infrastructural technologies. Technologies that succeed become invisible, and invisible technologies attract public comment mainly when they fail, so the sentiment stream of a mature payment system is dominated by exception reports rather than by the silent majority of successful transactions. The same negativity skew appears wherever a digital service becomes routine, and it interacts with the classification and delegation costs that consumers experience when systems sort and act on them without explanation [22]. Managers who compare the sentiment of a new, much discussed technology with the sentiment of an established, failure dominated one will misread novelty enthusiasm as superiority and operational maturity as decline. Dahlberg, Guo and Ondrus's call for research on payment ecosystems rather than on isolated adoption decisions points to the remedy, namely to study sentiment across the whole chain of merchants, platforms, issuers and users among whom trust in a transaction is actually produced [24].

Social commerce completes the adoption picture. Hajli showed that social commerce constructs, including forums and communities, ratings and reviews, and recommendations and referrals, build social support and thereby increase consumers' intention to buy, which makes visible public evaluation a component of the transaction mechanism itself rather than commentary upon it [25]. Busalim and Hussin, reviewing 110 studies of social commerce published between 2010 and 2015, mapped a rapidly growing field still dominated by adoption and intention models [26]. Zhang and Benyoucef's review of consumer behaviour in social commerce likewise organised the evidence around trust, social influence and the stages of the purchase decision [27]. For sentiment analysis the implication is direct. In social commerce, expressed sentiment is simultaneously an input to other consumers' decisions, an output of the firm's service performance and a signal the firm can monitor, so the same text stream serves as market research, sales force and early warning system at once.

3. Sector Patterns in Public Sentiment

Across sectors, a consistent finding is that public sentiment differentiates competitors and aspects of the offer in ways that aggregate satisfaction scores conceal. He, Zha and Li combined social media competitive analysis with text mining in the pizza restaurant industry, drawing on Twitter and Facebook content, and showed that sentiment comparison across rival chains identifies the specific strengths and weaknesses on which competition is actually being fought, turning public text into a benchmarking instrument [28]. At brand level, Ghiassi, Skinner and Zimbra built a hybrid system combining n-gram analysis with a dynamic artificial neural network and applied it to a corpus of more than ten million tweets about major consumer brands, demonstrating that brand sentiment on social media is measurable at industrial scale with instruments tuned to the platform's language [29]. Mostafa reached a complementary result with a lexicon of about 6,800 seed adjectives applied to a random sample of 3,516 tweets, finding that sentiment toward the global brands studied was generally positive, a reminder that public brand talk skews favourable in ordinary times and that deterioration, not the level itself, is often the actionable signal [30].

Hospitality shows the research and revenue value of aspect level reading. Xiang and colleagues applied big data and text analytics to hotel reviews on Expedia and showed that the textual content of reviews reveals what drives guest satisfaction more informatively than numerical ratings alone, because guests explain in words which attributes of the stay produced their evaluation [31]. Electronic commerce research generalises the point. Salehan and Kim used sentiment mining within a big data analytics design to predict the performance of online consumer reviews, establishing that the sentiment characteristics of a review help determine the response it receives from other consumers, so sentiment shapes the information environment of the marketplace and not merely its measurement [32]. Fang and Zhan, working with product review data from Amazon, compared sentiment analysis methods on real marketplace text and confirmed both the feasibility of review mining at scale and its sensitivity to method choice [33]. The pattern across these sectors is consistent. Sentiment toward a technology mediated offer is granular, comparative and consequential, attaching to delivery, service recovery, price fairness and interface quality rather than to the technology in the abstract, and it feeds back into the decisions of other buyers, which is why Chevalier and Mayzlin's asymmetry, in which negative evaluations outweigh positive ones, recurs throughout the applied literature [6].

Sector comparison also exposes where sentiment evidence is strong and where it is thin. Consumer brands, hospitality and retail platforms are heavily studied because their public text is abundant and their outcomes, in sales, ratings and review helpfulness, are observable. Financial services, healthcare and business to business technology, where the implementation of digital technology is equally strategic, produce far less public text per transaction and far fewer sentiment studies in the corpus. The strategic reading of a sector's silence is therefore ambiguous. It may indicate satisfaction, or merely privacy, regulation and channel structure. Firms in low text sectors that benchmark their sentiment programmes against retail experience will systematically overestimate how completely public sentiment represents their customers.

4. Sentiment as Strategic Feedback for Business Strategy

The reviewed evidence supports treating public sentiment as a formal feedback system inside digital business strategy, performing four functions. The first is sensing. Continuous sentiment monitoring detects the public reception of a technology launch, a price change or a service failure earlier than periodic surveys can, and Bollen, Mao and Zeng's market result shows that aggregated public mood can carry leading information even about outcomes far outside the firm's own channels [7]. The second function is diagnosis. Aspect based analysis converts a falling score into an agenda, localising discontent in onboarding, payment failure, delivery or support, as the hospitality and electronic commerce studies demonstrate [31], [32]. The third function is benchmarking, since the same public stream that carries evaluations of the firm carries evaluations of its rivals, and comparative sentiment analysis of the kind He, Zha and Li performed converts that stream into competitive intelligence [28]. The fourth function is accountability, because a public, time stamped sentiment record makes it harder for an organisation to declare a digital transformation successful while its users report otherwise, and easier for boards to ask which specific experience the transformation improved.

Exercising these functions well imposes disciplines that the literature makes visible. Sentiment must be read as a distribution over aspects and segments rather than as a single index, because averages conceal the angry minority whose evaluations, as the one star asymmetry shows, do disproportionate commercial damage [6]. It must be triangulated

against behavioural data, since expressed sentiment and revealed behaviour diverge in known ways, for example when users praise a chatbot's speed in surveys yet abandon it in practice. It must be governed, because listening at scale is also surveillance at scale, and the costs that Puntoni and colleagues identify in data capture and classification apply to the firm's own listening apparatus as much as to its customer facing systems [22]. Finally, it must be connected to decision rights. Sentiment that no manager is empowered to act upon degrades into decoration, while sentiment routed into product, service and communication decisions becomes what Davenport and colleagues describe for artificial intelligence generally, a capability that augments rather than replaces managerial judgement [23]. The firms in the reviewed studies that extracted value from sentiment shared this trait. They treated the score as the beginning of an inquiry, not as its conclusion.

Operationally, the feedback system works best as a repeating cycle rather than as a monitoring wall. A launch or policy change is preceded by a baseline reading, so that later movement has a comparison point. During the launch window, aspect level readings are reviewed on a short cycle by the team that owns the experience, with pre-agreed thresholds that trigger human review of the underlying texts, because a threshold crossing is a prompt to read, not a verdict. After stabilisation, the cadence lengthens and the emphasis shifts to benchmarking against competitors exposed to the same public environment, following the comparative logic of the pizza industry study [28]. Two failure modes recur in practice and both are visible in the literature's warnings. The first is metric capture, in which teams optimise the sentiment score itself, for example by soliciting reviews from satisfied customers at moments of delight, until the score stops describing the public and starts describing the solicitation programme, a distortion already implicit in the finding that review populations are self selected and overwhelmingly positive [6]. The second is automation of response without automation of understanding, in which firms reply at scale to detected negativity with templated apologies while the operational cause persists, converting a listening system into a public relations filter. Both failure modes share one root, the treatment of sentiment as an output to be managed rather than as evidence about an experience to be improved.

5. Limitations of the Current Evidence

Four limitations qualify everything the corpus reports. The first is language. Sentiment instruments remain far stronger in English than in other languages, and the Indonesian case illustrates the gap concretely. The InSet lexicon, built for Indonesian sentiment analysis in microblogs, contains 3,609 positive and 6,609 negative words weighted from minus five to plus five, and its evaluation showed both that a dedicated Indonesian resource is feasible and how much smaller such resources remain than their English counterparts [34]. Transformer resources are newer still. Koto and colleagues released IndoBERT, pretrained on more than 220 million words of Indonesian, together with the IndoLEM benchmark of seven evaluation tasks, and reported state of the art results on most tasks, yet a general purpose language model is only the foundation for business sentiment instruments, not the instrument itself [35]. Indonesian public text compounds the difficulty with pervasive code mixing with English and regional languages, informal spelling, abbreviations and slang that age quickly, so models validated on formal Indonesian or on English understate their error rates on real Indonesian business conversation.

The second limitation is the text itself. Sarcasm and irony invert literal polarity, comparisons distribute sentiment across two targets, and much public evaluation is carried by images, video and emoji that text pipelines ignore, problems the methodological surveys consistently list as unsolved [13], [15]. The third limitation is the platform. Each platform's population differs by age, income and urbanisation, its algorithms amplify some voices and suppress others, and its policies on automated accounts change over time, so sentiment measured on one platform is a sample of that platform's public, not of the firm's market. Automated and coordinated accounts can manufacture apparent sentiment around launches, controversies and financial assets, and few of the applied studies in the corpus audit their corpora for such contamination. The fourth limitation belongs to this review itself. It synthesises published, predominantly English language studies, retrieved from major publishers and indexes, so it inherits their publication bias toward positive and significant results, and its narrative method cannot quantify how large any reported effect is across contexts. The honest conclusion is that public sentiment is a powerful but biased instrument, closer to a listening post on a busy street than to a census, and strategy should price its evidence accordingly.

6. An Agenda for Indonesian Business Research

Indonesia combines one of the world's largest social media populations with digital platforms that mediate commerce, payment and public conversation for tens of millions of users, yet the reviewed corpus contains very few verified sentiment studies of Indonesian business technology implementation. The agenda that follows from this review has five elements. First, resource building deserves priority, extending the foundation laid by the InSet lexicon [34] and by IndoBERT [35] toward validated, openly documented sentiment instruments for Indonesian business domains, with test sets drawn from marketplace reviews, application store comments and service chats, and with explicit handling of code mixing and regional language admixture. Second, Indonesian researchers should study the technologies Indonesians actually adopt at scale, including social commerce conducted through messaging and short video platforms, digital wallets and app based transport and food services, applying the aspect based designs that proved most informative in hospitality and electronic commerce [31], [33]. Third, the chatbot findings invite local replication, since the disclosure and anthropomorphism effects reported by Adam, Wessel and Benlian and by Crolic and colleagues were established in Western service contexts, and Indonesian norms of politeness, hierarchy and indirect complaint may moderate them in either direction [17], [19]. Fourth, payment sentiment should be studied where it actually occurs, inside commerce and service streams rather than in payment specific channels, correcting the adoption centred imbalance that Dahlberg, Guo and Ondrus identified in the international literature [24]. Fifth, Indonesian business research should build the governance dimension from the start, studying consent, transparency and data protection in sentiment listening itself, so that local practice does not reproduce the privacy costs that the international literature now criticises [22]. Across all five elements, the methodological rule of this review applies with full force. Indonesian sentiment research will be credible exactly to the extent that it reports its instruments, its corpora and its exclusions openly, and attributes every finding to the data that produced it.

CONCLUSION

This systematic review of thirty-five published studies shows that public sentiment analysis has matured through lexicon, machine learning and transformer generations into a practical instrument for reading how publics receive the digital technologies that business strategies deploy, and that the evidence it has produced converges on a clear strategic lesson, namely that publics judge chatbots, payment systems, commerce platforms and artificial intelligence not by their technical novelty but by the honesty, competence and emotional fit of the experiences they deliver, with negative evaluations carrying disproportionate weight in market outcomes. For firms, sentiment is therefore best governed as a continuous feedback system for sensing, diagnosis, benchmarking and accountability, read at aspect level and triangulated against behaviour. For Indonesian business research, the priority is to build validated local language resources and to study sentiment where Indonesian digital commerce actually happens, under the same discipline this review has followed, in which no finding is reported without the study and data that produced it.

REFERENCES

[1] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of social media," *Business Horizons*, vol. 53, no. 1, pp. 59-68, 2010, doi: 10.1016/j.bushor.2009.09.003.

[2] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre, "Social media? Get serious! Understanding the functional building blocks of social media," *Business Horizons*, vol. 54, no. 3, pp. 241-251, 2011, doi: 10.1016/j.bushor.2011.01.005.

[3] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1-135, 2008.

[4] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge, U.K.: Cambridge University Press, 2015.

[5] R. Feldman, "Techniques and applications for sentiment analysis," *Communications of the ACM*, vol. 56, no. 4, pp. 82-89, 2013, doi: 10.1145/2436256.2436274.

[6] J. A. Chevalier and D. Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research*, vol. 43, no. 3, pp. 345-354, 2006, doi: 10.1509/jmkr.43.3.345.

[7] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1-8, 2011, doi: 10.1016/j.jocs.2010.12.007.

[8] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093-1113, 2014, doi: 10.1016/j.asej.2014.04.011.

[9] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowledge-Based Systems*, vol. 89, pp. 14-46, 2015, doi: 10.1016/j.knosys.2015.06.015.

[10] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267-307, 2011, doi: 10.1162/COLI_a_00049.

[11] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word-emotion association lexicon," *Computational Intelligence*, vol. 29, no. 3, pp. 436-465, 2013, doi: 10.1111/j.1467-8640.2012.00460.x.

[12] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. 8th Int. AAAI Conf. Weblogs and Social Media (ICWSM)*, 2014, pp. 216-225, doi: 10.1609/icwsm.v8i1.14550.

[13] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems*, vol. 226, Art. no. 107134, 2021, doi: 10.1016/j.knosys.2021.107134.

[14] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 4, Art. no. e1253, 2018, doi: 10.1002/widm.1253.

[15] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731-5780, 2022, doi: 10.1007/s10462-022-10144-1.

[16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North American Chapter Assoc. Comput. Linguistics, Human Language Technologies (NAACL-HLT)*, 2019, pp. 4171-4186.

[17] M. Adam, M. Wessel, and A. Benlian, "AI-based chatbots in customer service and their effects on user compliance," *Electronic Markets*, vol. 31, no. 2, pp. 427-445, 2021, doi: 10.1007/s12525-020-00414-7.

[18] M. Ashfaq, J. Yun, S. Yu, and S. M. C. Loureiro, "I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents," *Telematics and Informatics*, vol. 54, Art. no. 101473, 2020, doi: 10.1016/j.tele.2020.101473.

[19] C. Crolic, F. Thomaz, R. Hadi, and A. T. Stephen, "Blame the bot: Anthropomorphism and anger in customer-chatbot interactions," *Journal of Marketing*, vol. 86, no. 1, pp. 132-148, 2022, doi: 10.1177/00222429211045687.

[20] M. Chung, E. Ko, H. Joung, and S. J. Kim, "Chatbot e-service and customer satisfaction regarding luxury brands," *Journal of Business Research*, vol. 117, pp. 587-595, 2020, doi: 10.1016/j.jbusres.2018.10.004.

[21] M.-H. Huang and R. T. Rust, "Artificial intelligence in service," *Journal of Service Research*, vol. 21, no. 2, pp. 155-172, 2018, doi: 10.1177/1094670517752459.

[22] S. Puntoni, R. W. Reczek, M. Giesler, and S. Botti, "Consumers and artificial intelligence: An experiential perspective," *Journal of Marketing*, vol. 85, no. 1, pp. 131-151, 2021, doi: 10.1177/0022242920953847.

[23] T. Davenport, A. Guha, D. Grewal, and T. Bressgott, "How artificial intelligence will change the future of marketing," *Journal of the Academy of Marketing Science*, vol. 48, no. 1, pp. 24-42, 2020, doi: 10.1007/s11747-019-00696-0.

[24] T. Dahlberg, J. Guo, and J. Ondrus, "A critical review of mobile payment research," *Electronic Commerce Research and Applications*, vol. 14, no. 5, pp. 265-284, 2015, doi: 10.1016/j.elerap.2015.07.006.

[25] N. Hajli, "Social commerce constructs and consumer's intention to buy," *International Journal of Information Management*, vol. 35, no. 2, pp. 183-191, 2015, doi: 10.1016/j.ijinfomgt.2014.12.005.

[26] A. H. Busalim and A. R. C. Hussin, "Understanding social commerce: A systematic literature review and directions for further research," *International Journal of*

Information Management, vol. 36, no. 6, pp. 1075-1088, 2016, doi: 10.1016/j.ijinfomgt.2016.06.005.

[27] K. Z. K. Zhang and M. Benyoucef, "Consumer behavior in social commerce: A literature review," *Decision Support Systems*, vol. 86, pp. 95-108, 2016, doi: 10.1016/j.dss.2016.04.001.

[28] W. He, S. Zha, and L. Li, "Social media competitive analysis and text mining: A case study in the pizza industry," *International Journal of Information Management*, vol. 33, no. 3, pp. 464-472, 2013, doi: 10.1016/j.ijinfomgt.2013.01.001.

[29] M. Ghiassi, J. Skinner, and D. Zimbra, "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6266-6282, 2013, doi: 10.1016/j.eswa.2013.05.057.

[30] M. M. Mostafa, "More than words: Social networks' text mining for consumer brand sentiments," *Expert Systems with Applications*, vol. 40, no. 10, pp. 4241-4251, 2013, doi: 10.1016/j.eswa.2013.01.019.

[31] Z. Xiang, Z. Schwartz, J. H. Gerdes Jr., and M. Uysal, "What can big data and text analytics tell us about hotel guest experience and satisfaction?" *International Journal of Hospitality Management*, vol. 44, pp. 120-130, 2015, doi: 10.1016/j.ijhm.2014.10.013.

[32] M. Salehan and D. J. Kim, "Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics," *Decision Support Systems*, vol. 81, pp. 30-40, 2016, doi: 10.1016/j.dss.2015.10.006.

[33] X. Fang and J. Zhan, "Sentiment analysis using product review data," *Journal of Big Data*, vol. 2, Art. no. 5, 2015, doi: 10.1186/s40537-015-0015-2.

[34] F. Koto and G. Y. Rahmaningtyas, "InSet lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," in *Proc. Int. Conf. Asian Language Processing (IALP)*, 2017, pp. 391-394, doi: 10.1109/IALP.2017.8300625.

[35] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," in *Proc. 28th Int. Conf. Comput. Linguistics (COLING)*, 2020, pp. 757-770, doi: 10.18653/v1/2020.coling-main.66.