

企業と担当者を炎上リスクから守る ホールで導入進む「AI投稿チェッカー」の役割とは



吉原教一郎
代表取締役社長
Yoshihara Kyoichiro

棚橋夕焼
Webブランディング事業部
マネージャー
Tanahashi Yuya

SNSを活用したマーケティングがパチンコホールに欠かせない集客手段となる一方で、企業やSNS担当者は炎上対策に注意を払わなくてはならなかった。そんな負担を軽減するITサービスが「AI投稿チェッカー」だ。ホール業界に特化してこのシステムを提供しているBLITZ Marketing(ブリッツ・マーケティング)の吉原教一郎社長と棚橋夕焼Webブランディング部マネージャーに、プロダクトに込めた思いを聞いた。(文中敬称略)

——BLITZ Marketingの事業について教えてください。

吉原 Webマーケティングにおけるアフィリエイト事業が創業時の事業でした。アフィリエイト(個人やメディア運営者)にブログやSNS、YouTube、Webサイトなどで広告主の商品を紹介してもらって成果報酬を得るビジネスモデルです。その事業を続ける中から生まれたのが現在も続いている誹謗中傷対策事業です。

——どんな事業ですか？

吉原 アフィリエイトが紹介した商品で意図せず問題が起きると、自分たちの商品ではないと説明しても納得されず、アフィリエイト本人が誹謗中傷にさらされる。そんなケースで誹謗中傷対策のサポートをしたりしていました。その中であるとき、ネットの誹謗中傷の記事を、弁護士を通じて削除できたことがあります。消えないと思いついて記事が消えたことで、同じようにネットで誹謗中傷されているアフィリエイトを助けてあげられるのではと考えるようになりました。誹謗中傷対策事業でした。

——その仕事を事業化していった？

吉原 対策を請け負っていくうちに、対応できる領域が次第に広がっていききました。ネットに書かれた記事の対策だけでなく、検索結果や、検索時に表示されるサジェスト(予測キーワード)のネガティブ対策サポートを行ったりと、取り組みの幅がさらに広がっていったのです。こうして誹謗中傷対策を、BtoBの商材として本格的に事業化しました。

——そこから派生したツールが「AI投稿チェッカー」ですか？

吉原 その前に、ネットの風評をAIが365日監視する「AIブランドモニター」というサービスがありました。設定したキーワードをAIが監視し続けて検索結果を自動で収集。ネガティブな書き込みとネガティブではない書き込みが自動で振り分け、ネガティブな書き込みがあったらアラートで通報するサービスです。

——このサービスが生まれた背景は？

吉原 すでにネット上で誹謗中傷されている会社だけではなく、誹謗中傷や炎上を予防するシステムがあれば、もっと多くの企業の役に立てると考えました。そこで、SNSに投稿する前に投稿文をチェックすると、炎上リスクを可視化できるように「AIブランドモニター」に「AI投稿チェッカー」の機能を内包したのです。

マルハン東日本からの問い合わせが契機に

——「AI投稿チェッカー」をパチンコホールに提案したきっかけは？

吉原 マルハン東日本様がAI投稿チェッカーの機能に注目してくださって、「これをパチンコ業界向けにカスタマイズしてもらえませんか」というお問い合わせをいただいたのがきっかけでした。その後、マルハン様のご意見を聞きながらシステム開発を行って、「AI投稿チェッカー」ができていきました。



——マルハン東日本にはどんな課題があったのですか？

棚橋 店舗スタッフが、炎上してしまうことを恐れるあまり、SNSに取り組みたくないという声の一部がありました。炎上しないように、全てのポストを上位役職者に確認してもらうというの現実的ではなく、スタッフの心理的安全性が担保された状況でSNSを運用できるようにということで投稿チェッカーを採用していただきました。

——マルハン東日本と一緒に開発を？

棚橋 はい。去年の2月ごろに問い合わせをいただいてから、一緒に開発を進めさせていただきました。広告宣伝ガイドラインの内容も含めて、独特の言い回しでNGになるワードの最新情報なども教えていただきながら、よりパチンコ業界の実情に沿う形で機能をカスタマイズしました。現在はマルハン東日本様以外にも10社ほどのホール企業様でご利用いただいています。

——「AI投稿チェッカー」の具体的な機能について教えてください。

棚橋 SNSの中でもっとも誹謗中傷が起きやすいXへの投稿が対象です。使い方はこうです。まず「元のメッセージ」欄に140文字以内の文章を貼り付ける。次

に投稿チェックボタンをクリック。すると5秒後ぐらいにその投稿文に対するリスクスコアが出てきます。リスクスコアが0〜20だとリスクはほぼゼロ。20ぐらいだと少し修正がかかった文章が提案され、さらにリスクが高くなると、投稿ができないようになります。例えば、リスクスコア60〜80は他人への悪口や「設定」という言葉を入れているレベル。これらは投稿ができないように投稿ボタンが無効にされます。

——リスクが中レベルの場合はどうなりますか？

棚橋 「校正後のメッセージ」をAIが提案してくれます。Xへの予約投稿機能も一緒にしているのです。この段階でチェックを済ませた上で予約投稿をすることで、より安全に運営できるようになっています。業界全体で共有されている禁止ワードは全体に反映されますが、法人様ごとにこのキーワードには気を付けてというワードがある場合は、初期段階でのヒアリングでお伺いして追加しています。人のチェックも可能な投稿管理ツールをお望みの企業様もいらつしたつたので、チェックが終わった後の保留機能や承認機能を持たせたりもしています。

——画像のリスク判定についてはどういう状況ですか？

棚橋 現在、ポスターやバナーの画像

チェック機能も実装して試験版で提供しています。

——企業ごとのニーズを聞きながら伴走していくイメージですね。

棚橋 まさにそうなんです。企業ごとのニーズを聞きながら地域ごとの特性にも対応するようにしています。10月に「広告宣伝では正勧告を行った事例」が公表されたので、現在はそれを踏まえた地域ごとのルールにも対応。なるべく地域特性に合ったリスクスコア判定ができるようにしています。

——「AI投稿チェッカー」も誹謗中傷対策も、ネット社会で浮き彫りになった社会課題を解決するプロダクトだと思っています。

吉原 ぼくは決して誹謗中傷対策をしたいわけではありません。「ネガティブは1円にもならない」という信念があるだけなんです。ネガティブな感情に基づく発言は、ちよっと仲間を作っているとき楽になるだ

続きはデジタルブックで
ご覧いただけます。

詳細はこちら▶

