

Louvain School of Management

Understanding crowdfunding using artificial intelligence and textual content

Author: Lucas Daemen
Supervisor: Prof. James Thewissen
Academic year 2025-2026
Dissertation for the master of
Finance and Accounting
Daytime schedule

STATEMENTS ON USE OF GENERATIVE AI

Statements on responsible use of generative AI	Answer
The content of my/our master's thesis is my original output, even if I used generative AI to help prepare it.	Yes
I master the theories, tools and methods that I use for the thesis.	Yes
I verified that the output presented in the thesis is accurate and valid, and that it contains no hallucination, no false reference, etc.	Yes
I acknowledge that I master the final output and that I am able to explain it and answer questions about it.	Yes
I made sure not to provide generative AI tools with access to confidential data or information	Yes

Signature

Author last name, first name

Lucas Daemen

Date

27 May 2026

Signature

Daemen Lucas

I acknowledge the use of Perplexity (AI) to search for information on the internet in an optimized way. I acknowledge the use of Claude (AI) to challenge and correct my R code. I also acknowledge the use of Deepl for translation purpose, and ChatGPT (AI) and Reverso for grammar correction.

The scientific initiative and content are entirely my own. All ideas were developed by my own, jointly with the thesis supervisor: Prof. James Thewissen

Table Of Contents

1. Introduction.....	4
2. Literature Review	5
2.1. What Is Crowdfunding?	5
2.2. What Is Artificial Intelligence?	6
2.3. Artificial Intelligence's Use In Crowdfunding	6
2.4. Text Impact On Investors	7
2.5. Artificial Intelligence's Text Writing Style	7
2.6. Artificial Intelligence's Text Writing Style Impact On Investors	8
Tone And Sentiment	9
Readability And Textual Complexity.....	9
Other Influencing Factors	10
Theoretical Framing.....	10
3. Hypotheses Development	12
3.1. Hypotheses	12
Artificial Intelligence's Impact On Crowdfunding.....	12
Artificial Intelligence's Text Writing Style	12
Text Writing Style Impact On Crowdfunding	13
3.2. Contributions	13
4. Research Design	14
4.1. Sample Construction And Selection	14
4.2. Variables	15
Success Variables	15
AI Variable	15
Textual Variables	16
Control Variables	17
4.3. Research Method.....	18
5. Results	20
Hypothesis 1: Artificial Intelligence's Impact On Crowdfunding	20
Hypothesis 2: Artificial Intelligence's Text Writing Style.....	21
Hypothesis 3: Text Writing Style Impact On Crowdfunding	22
6. Discussion.....	25
Results And Literature Review.....	25
Implications.....	26
7. Conclusion	27
8. References	28
9. Appendix	32
9.1. Formula Explanations	32
9.2. Raw Results.....	35
9.3. R Code	49

1. Introduction

In April 2025, 74% of all new pages published on the web contained content generated by Artificial Intelligence (AI). According to McKinsey, 65% of organizations now regularly use generative AI in at least one business function. AI is not coming, it is already here. AI can already write and most people cannot tell the difference. Recent studies show that even trained evaluators correctly identify AI text only slightly better than chance. (Fiedler, 2025) (Law, 2026) (Singla et al., 2024)

Crowdfunding is not an exception. Every day, entrepreneurs present their projects to strangers on Kickstarter. Every day, several use AI to write the words that are supposed to make those strangers believe in them. The pitch sounds good. It is optimistic, structured, polished, etc. But does it really work?

The literature doesn't seem to agree. Some studies find that AI boosts campaign outcomes. Others find that it damages it. None of them explain why. This paper does.

Using 85,752 Kickstarter campaigns across 123 countries, this study aims to find not only if AI hurts or helps campaign performance, but also how. This is done by linking AI use to campaign outcomes through measurable changes in writing style.

This study brings two contributions. First, it ends the debate by providing a statistical answer to the question of AI's impact on crowdfunding success across a large sample. Secondly, and in parallel with the first contribution, it identifies the exact textual mechanisms through which AI influences backer behavior. This combination is something the existing literature has never done.

In a world where anyone can generate a perfect campaign description in seconds, understanding what really influences and convinces people to trust and fund a project has never mattered more.

2. Literature Review

2.1. What Is Crowdfunding?

Crowdfunding is a way to raise money from a large number of people to fund a project or a cause. This is usually done through an online platform. It is an alternative to traditional bank financing or business angels¹. It offers the opportunity to contributors to support projects directly in a traceable way, often in a lot of smaller amounts rather than one big payment. But this method is not just about raising money. Crowdfunding can also help to test a demand or build a community. (Bpifrance, 2024) (Morgan, 2025) There are 4 main types of crowdfunding:

- Donation-based crowdfunding. People give money without expecting anything back, usually for charity or community causes.
- Reward-based crowdfunding. People called backers receive a reward that is not financial, such as a product or early access. It is common for creative projects like startups.
- Equity-based crowdfunding. People called investors receive shares or ownership in the company in exchange for funding. It is often used by startups that want capital growth.
- Debt-based crowdfunding, also called peer-to-peer lending. This method is where the money is repaid later, usually with interest. It is a way to borrow money while avoiding banks.

Initial Public Offerings (IPO) and Initial Coin Offerings (ICO) cannot be considered as crowdfunding even if the funding comes from the crowd. The main difference between equity crowdfunding and IPOs is that an IPO is a formal public listing on a stock exchange with a determined strict regulation and a higher cost. Equity crowdfunding is a smaller private fundraising done online on a dedicated platform and the shares are not traded. Furthermore, in an ICO, a project raises money by selling digital tokens (crypto currency) to supporters. This is different to owning part of a private company. (BSEtec, 2025) (Zarya, 2023)

The type of crowdfunding often depends on the nature of the project. A startup launching a new product might use reward-based crowdfunding to collect their first orders. A growing company might use equity crowdfunding to sell a small part of ownership. A non-profit helping disaster victims would more likely use donation-based crowdfunding. (Bello, 2024) (Heaslip, 2026)

¹ A business angel is a person who invests its own money directly into a private business like a startup often in exchange for equity. They usually contribute through their experience by giving advice and contacts. (European Commission, s. d.)

2.2. What Is Artificial Intelligence?

Artificial intelligence (AI) is a technology that does tasks that usually require a human work. Unlike traditional software with fixed rules, AI learns on its own by working with large databases to detect recurring patterns and improve accuracy of its work and answers over time. AI relies on data like text or images and powerful resources like complex algorithms. Most of the time, AI works using machine learning. AI models train on given examples to develop neural networks similar to the human brain. It processes data through complex layers and is adjusted with feedback to boost precision. This way of learning is similar to a child learning from examples and feedback adjustments from parents. (Coursera, 2025) (Kiger, 2023)

In 2026, AI is all around us. People are often using AI as a work assistant, a content generator or a tool helping in decision-making. AI is no longer a simple chatbot. It has now become AI agents that can take actions across email and browsers. Examples of uses can be writing and editing, summarizing documents and extracting key points, providing customer service through chatbots, planning travel, answering financial questions, organizing daily tasks, researching, documenting, coding help, task automation, and many more. People use AI because it saves time and boosts productivity by doing repetitive or complex work faster. Companies also use it to reduce costs and handle large amounts of data in less time. (Nicoud, 2026) (Prestianni, 2026) (Ray, 2025) (Sigma Technology Group, 2025)

2.3. Artificial Intelligence's Use In Crowdfunding

People increasingly started using AI tools to present their crowdfunding projects. AI can generate persuasive descriptions and optimize keywords for platforms. That may or may not be a good idea. Research examining Kickstarter data finds out that mandatory AI disclosure does not significantly impact success rates or backers. This suggests a neutral effect of authenticity concerns. (Lodder, 2025) To understand the impact of AI on crowdfunding success, another paper analyzes GoFundMe small business campaigns. It suggests that augmenting narratives with large language models, like ChatGPT increasing for gratitude or urgency, boosts predicted funding likelihood by 11.9% on average, rising to 16% for originally unfunded campaigns. This same study finds that 83% of human evaluators preferred AI versions. (Ye et al., 2024) On the other hand, a regular research seminar suggests that campaigns with high AI probability raise 66% less funds and have 16% lower goal attainment odds. (Zavertiaeva et al., 2026)

2.4. Text Impact On Investors

While literature doesn't seem to agree on AI's impact on crowdfunding success, several studies try to understand which textual components influence investors' decisions, even without any link to AI. Just by looking at a project description, there are many factors that can influence investors' perception.

These are several examples:

- Textual complexity. High complexity like long sentences or complex jargon makes disclosures harder to understand. The impact is mostly negative. It reduces judgment accuracy and efficiency. Simpler texts enable balanced processing. (Yu & Zhao, 2023)
- Sentiment analysis. Positive or negative tone in financial texts shapes emotional interpretation of investors. Positive sentiment boosts optimism and investment. Negative triggers caution. (Todd et al., 2024)
- Strategic highlighting, highlighting gains over losses. Favorable framing improves the distinction of good firms and harms judgment of poor ones. Investors are better equipped to distinguish between good and poor firm performance when the accompanying disclosure emphasizes information that is consistent with the firm's performance. (Druz et al., 2016) (Henry & Peytcheva, 2018)
- Credibility signals. Investor's reaction to tone is stronger when the message seems credible. Consistent tone and verifiable claims are signals of trustworthiness beyond raw content. (Hennig et al., 2023)
- Readability. This factor affects investor's behavior and decision-making. Simple structure like short sentences makes information searching easier. Higher readability correlates with better processing and decisions while poor readability limits engagements. (Alivia et al., 2020)

2.5. Artificial Intelligence's Text Writing Style

As seen above, literature shows that text characteristics can influence investor decision-making. AI may matter not because of its technology itself, but because of the way it changes the content and presentation of the message.

Lots of studies tried to identify the main differences between texts generated by AI and those written by humans. Ma et al. (2023) tried to list these differences. For AI: a smooth style, perfect cohesion, but a lack of personality and creativity. For human texts: authenticity,

unpredictability, and expressiveness. However, these differences do not mean that it is easy to distinguish articles written by AI from those written by humans. This distinction is so complex to identify that it has been the subject of lots of articles. (Akinwande et al., 2024) (Ma et al., 2023)

Despite noticing the difference not being an exact science, studies came up with general principles. AI writing styles have linguistic patterns that differ from human texts. These features make an AI text identifiable through textual analysis. An AI generated text is more likely to contain a more formal and impersonal style, signaled by the increased presence of nouns and determiners, and fewer adjectives and adverbs. An AI generated text is also more likely to feature a lower lexical diversity, a smaller vocabulary size, and repetitive text. Studies using tools like LIWC² reveal an AI text is more structured with positive language and cognitive expressions like "furthermore" or "is fundamental". This allows the writer to avoid slang and contractions. AI often uses specific structures such as contrast phrases like "not just X, but also Y," and persistent paragraph lengths of 3-4 sentences with 16-19 words each. (AlAfnan & MohdZuki, 2023) (Cox, 2025) (Rodrigues et al., 2026) (Terčon & Dobrovoljc, 2025)

2.6. Artificial Intelligence's Text Writing Style Impact On Investors

Experimental research shows that investors adjust more their return expectations in response to human forecasts than when explicitly labeled as AI generated. This trust penalty can be explained by lower perceived credibility of AI forecasts. Investors rate AI outputs as less legitimate and less accountable because they feel detached from human judgment. This means that when investors recognize a text as written by AI because of its style or because it is disclosed, they may discount its informational value. The reasons for this low perceived credibility can be multiple. One possibility could be that several factors can influence investors' judgments. Despite that, most sophisticated investors with higher AI knowledge or with quantitative tools can have another vision of AI analysis. For them, a precise and technical writing style can boost objectivity and analytical rigor. This could increase response to AI analysis. (Sattar & Sattar, 2026) (Stradi & Verdickt, 2025)

² Linguistic Inquiry and Word Count: textual analysis software that categorizes words in a given text into psychological dimensions by comparing them against a predefined dictionary. Allows comparisons between human and AI-generated writing. Helps detecting higher formality or repetition in AI text.

Tone And Sentiment

As stated said above, several factors of AI's writing style could impact investors' decisions. The reason may be coming from several language components like, for example, the tone and sentiment of AI financial texts. Textual analysis research on annual reports and earnings disclosures shows that tone isn't just decorative. It can predict divergence of opinion and influence trading volume and return. Study from Qin and Cui (2024) finds that a negative narrative tone is associated with greater divergence of opinion among investors. They suggest that negative tone tends to signal high risks and lower profitability, thereby increasing disagreement and price volatility. On the other hand, a positive narrative tone, when credible and consistent, can reduce information asymmetry and divergence of opinion. (Qin & Cui, 2024) (Suryani & Fauz, s. d.)

Generative AI has often balanced language with moderate sentiment unless prompted³ otherwise. This can reduce the natural variation in tone that investors are used to and lower the ability of disclosures to explain management's degree of confidence or concern. On the other hand, if prompts ask AI to write optimistic texts, the result may push towards excessive positivity. This phenomenon could be read by sophisticated investors as "AI polished spin" and lead to distrust. (Neko, 2025) (Qin & Cui, 2024) (Sattar & Sattar, 2026)

Readability And Textual Complexity

Another factor that can impact investors is readability and textual complexity. Studies suggest that managers sometimes reduce readability when reporting bad news. This can be done through longer sentences and use of jargon. Similar to the tone, lower readability is associated with greater divergence of opinion and higher volatility. In some cases, it can also make the market inefficient, because investors find it harder to extract and agree on the economic implications of the information. AI tools can push writing style in both directions. It can improve clarity by simplifying sentence structure, making complex financial narratives more accessible to a wider set of investors. But it can also produce long complex content that may contribute to a harder understanding of the text. (Efretuei & Newcastle University, n.d.) (Qin & Cui, 2024) (Sattar & Sattar, 2026) (Stellner, 2022)

The impact on investors depends on how AI is integrated in the writing process. When it is used as an editor, so by reviewing structure and aligning style with a firm's established standards the

³ Prompt: the question or instruction that you give to an AI to tell it what to do and what kind of output you expect.

result is an improvement in readability. This can enhance perceived professionalism. However, when it is used as an autonomous writer, the writing style of its output can make different disclosures look and sound similar. This can potentially give less impact to disclosures and reduce investors' ability to detect shifts in risk or governance quality. In that case, even if numerical data is disclosed, the narrative that normally helps investors interpret numbers in context may lose its power. That could affect the speed with which markets incorporate specific news. (Asay et al., 2016) (Tan et al., 2014)

Other Influencing Factors

To have a complete overview of AI's writing style's impact on investors, other influencing factors must be covered. Recent work on tone distance refers to the divergence between managerial tone in narrative disclosures and the market. Studies find that when managerial tone is unusually optimistic, event-period stock returns around earnings announcements tend to be lower. In contrast, when tone is more conservative, markets may underreact to the good news. This causes a positive drift as investors recognize the firm's stronger position afterwards. AI's writing style can influence tone distance in both ways. (Druz et al., 2020) (Wan, 2025)

Because investors are increasingly aware that AI tools are involved in producing financial narratives, they may start to distinguish human texts from texts that look "AI like". For example, if a firm that historically has concise and precise reports suddenly publishes a very long generic text, investors may suspect heavy AI use and question if management is substituting itself for the narrative. In this way, AI writing style can become a (negative) signal about governance and managerial engagement. (Basnet et al., 2025) (Sun et al., 2025)

Theoretical Framing

From a theoretical view, the impact of AI generated text on investors can be linked to signaling theory⁴ and behavioral finance⁵. Managers are choosing not only what information to reveal but also how to present it. Rational investors update beliefs based on the content and also the act of disclosure. Research shows that qualitative factors of narrative reporting like tone, readability, specificity, etc. can carry additional information for investors beyond numerical data. This can affect perceived risk through uncertainty. When AI tools shape textual factors, they change the

⁴ Signaling theory studies how informed parties give information to less informed parties when there is information asymmetry. Signals are observable actions that reveal hidden information that would otherwise remain hidden.

⁵ Behavioral finance examines how biases and emotions cause investors to deviate from fully rational decision making.

relation between reality and observable signals. On the one hand, standardized neutral AI style can reduce variation in tone and perceived risk. On the other hand, more advanced AI configurations may increase the firm's style and make textual signals more stable and interpretable over time. (Asay & Rennekamp, 2017) (Basnet et al., 2025)

Behavioral finance highlights that investors are not fully rational. They rely on emotions and systematically overweight or underweight particular types of information depending on several aspects like familiarity, comfort, fear, trust, and many more. This highlights that linguistic form and textual factors interact with investors' bounded rationality by shaping their judgments. AI's way of writing interacts with these behavioral tendencies by reframing information. For example, by giving more value to risk management than to growth opportunities or by using more neutral sentiment and tone. It is this combination of modified textual signals and modified behavior that can help explain why AI's writing style can affect investor behavior and investment decisions even when quantitative data remain unchanged. (De Brouwer, 2025) (Elliott et al., 2023)

3. Hypotheses Development

3.1. Hypotheses

Artificial Intelligence's Impact On Crowdfunding

- H1-A(B): The utilization of artificial intelligence in crowdfunding descriptions is positively (negatively) associated with crowdfunding campaign performance, measured both by the probability of reaching the funding goal and by the total amount raised.
- H1-0: The utilization of artificial intelligence in crowdfunding descriptions is not associated with crowdfunding campaign performance.

Based on literature review, AI seems to have an impact on investor decision and thus on project funding, but it doesn't seem to agree whether AI's impact on the amount raised is positive or negative. H1 is directly associated with the literature review. As stated in part 2.3., several studies find that AI's impact on crowdfunding can be neutral, positive, or negative. (Lodder, 2025) (Ye et al., 2024) (Zavertiaeva et al., 2026) H1 will find out what is the outcome with precise methodology described in part 4. If H1-A or H1-B is correct (and thus H1-0 is incorrect), it means that the utilization of AI has an impact on crowdfunding performance.

Artificial Intelligence's Text Writing Style

- H2-A: The utilization of artificial intelligence in crowdfunding descriptions is associated with a variation in the tested textual factors of those descriptions.
- H2-0: The utilization of artificial intelligence in crowdfunding descriptions is not associated with a variation in the tested textual factors of those descriptions.

If one of H1-A or H1-B is confirmed, the next step is to understand how AI affects campaign performance. As described in the literature review, AI changes the way information is presented to readers and investors. The goal of this second hypothesis is to identify which textual factors AI changes in crowdfunding descriptions. If H2-A is confirmed, it would suggest that AI changes the writing style of campaigns in a way that we can measure. This may influence investor perception and thus the funding decisions.

Text Writing Style Impact On Crowdfunding

- H3-A(B): The tested textual factors in crowdfunding descriptions are positively (negatively) associated with crowdfunding campaign performance. Again, measured both by the probability of reaching the funding goal and by the total amount raised.
- H3-0: The tested textual factors in crowdfunding descriptions are not associated with crowdfunding campaign performance.

If H2-A is correct and demonstrate that AI influences textual characteristics, it does not automatically mean that those factors determine campaign success. This third hypothesis tests whether the textual factors identified in H2 have a significant impact on campaign performance. If H3-A or H3-B is confirmed in addition to H2-A, it becomes possible not only to show that AI impacts crowdfunding performance, but also to explain the mechanism through which it operates. This is crucial in order to be able to attribute the impact of AI on crowdfunding success specifically to the text writing style, rather than to any other mechanism that we couldn't observe.

3.2. Contributions

These 3 hypotheses will allow to contribute to the literature in 2 ways:

1. They will reconcile the literature by finding out what is the real impact of AI on investors' decision in the context of crowdfunding campaigns.
2. They will also find out what can explain this hypothetical impact by analyzing several textual factors potentially influenced by AI's writing style in crowdfunding descriptions. In addition, they will also find out what is the precise impact of those influencing factors thanks to numerical data.

4. Research Design

4.1. Sample Construction And Selection

This paper relies on data provided by its supervisor that has already been used for other research. All the variables described in section 4.2. come directly from the provided database. No additional modification has been made for this research. “The used database had been created using archival data from the largest reward-based crowdfunding platform Kickstarter, which since its launch in April 2009 has helped almost 200,000 campaigns raise more than \$5.6 billion around the world. This database started from Li's (2019) Kickstarter Structured Relational Database, which included 44 408,637 campaigns launched between April 2009 and January 2019, representing 95% of the projects launched during that period.” (Thewissen et al., 2026) It is important to note that this platform almost fully relying on reward-based crowdfunding. This means that the database does not contain donation-based crowdfunding and that people behind the campaigns are often entrepreneurs launching a startup or project. (Nav, 2026)

Several filters have also been applied. “Researchers that provided the database only included projects launched in 2015 or thereafter. Starting from 214,252 observations, they removed projects designated as ‘canceled’, ‘suspended’, ‘purged’, or ‘live’, and those without main model information. To mitigate potential unwanted effects, they also removed observations with multiple founders, incomplete names, and names with post-nominal titles, such as ‘Dr.’ and ‘Rev.’.” (Thewissen et al., 2026)

After applying the filters, the database has been updated and additional recent data has been added. In total, the final sample consists of 85,752 crowdfunding campaigns from 123 countries. Among these, 42,786 (49.89%) of the projects were successfully funded. Campaigns also display a substantial variation in their funding targets: approximately half of them set a goal of USD 6,500 or lower, whereas some seek amounts exceeding USD 10 million. (Thewissen et al., 2026)

4.2. Variables

Success Variables

To measure the success of a crowdfunding campaign, we use 2 variables. *Success* is a binary variable that is equal to 1 if the campaign reached its funding goal and 0 if it didn't. The database is balanced with 42,786 successful campaigns (49.89%) and 42,966 unsuccessful ones (50.11%), and no missing values. The amount raised is measured as the total amount pledged by backers in USD by a variable called *usd_pledged_w*. Its distribution is highly skewed⁶:

Min	1 st Qu.	Median	Mean	3 rd Qu.	Max
0	200	2184	24693	12691	544780

This indicates that a small number of very successful campaigns pull the average upward. It is also important to note that 4,180 campaigns didn't raise a single dollar. To address this skewness, the variable is transformed using the log +1. Where so that the +1 ensures that the campaigns that didn't raise any money are kept in the analysis.

AI Variable

To measure the use of AI crowdfunding campaign texts, we use *AI_use_per*. This variable measures the proportion of the campaign text identified as written by AI by GPTZero⁷. The values are expressed as proportions from 0 to 1. Here the distribution is also skewed.

Min	1 st Qu.	Median	Mean	3 rd Qu.	Max	N/A
0	0	0	0.044	0	0.949	3810

This time, the skewness indicates that most campaigns don't contain content detected as written by AI. The maximum reaching 0.949 means that some campaigns have nearly their entire text detected as generated by AI. The gap between the median and the mean confirms that AI usage is focused on a small number of campaigns. The majority seems to be entirely written by humans. It is also important to note that 3,810 observations have missing values for this variable. It is possible that GPTZero was unable to process them.

⁶ A distribution is skewed when most values are grouped on one side. The distribution is asymmetrical.

⁷ GPTZero is an AI detection tool built to identify if a text has been written by a human being or generated by AI by analyzing linguistic patterns at word and sentence level.

Textual Variables

The textual variables are extracted from the two main sections of each Kickstarter campaign page: the "About" section and the "Risk" section. The about section presents the project to potential backers and the risk section presents potential challenges and uncertainties. Here are the used variables:

AboutWords and *RiskWords* measure the number of words in each section. A longer "About" section may signal a more detailed and credible project. A longer "Risk" section may signal transparency but could also discourage backers by highlighting risk.

Flesch_withAI is based on the Flesch Reading Ease score. This is a readability index that measures how easy a text is to read based on length of sentence and word complexity. Higher scores mean the text is easier to read and lower scores mean its reading implies more effort. The formula gives less credit to longer sentences and more syllables per word. Therefore, simple and short sentences with short words lead to a highest score. (Ehara & Tokyo Gakugei University, 2023)

SentimentLM_withAI and *SentimentHE_withAI* measure the emotional tone of the text using two different methods. The first variable uses Loughran and McDonald. It classifies words into categories like negative, positive, uncertainty, litigious, etc. based on how investors interpret them in business contexts. This dictionary provides more accurate sentiment measures for financial documents. The second one uses Henry's method, another sentiment dictionary for finance. Compared with the first method, Henry's dictionary is less exhaustive and has been criticized for missing many negative terms that appear in disclosures Both variables capture the balance between positive and negative words. For both, higher values reflect a more positive tone. (Palmer et al., 2021) (University of Notre Dame, n.d.)

Tone_withAI measures the overall tone of the campaign text. It is done by combining multiple sentiment dictionaries unlike the sentiment variables described above that only rely on a single dictionary. Higher values reflect a more positive and optimistic tone. Lower values indicate a more negative or cautious tone. Including this variable allows to capture the emotional orientation of the crowdfunding description in a broader way than a single dictionary approach. (Kearney & Liu, 2014)

LC_PCI_withAI, *Soph_PCI_withAI*, and *Var_PCI_withAI* are all components linked to lexical quality. Lexical complexity captures how difficult the used vocabulary is. Lexical sophistication

measures the use of rare or advanced words. Lexical variation reflects the diversity of the vocabulary used across the text. Together, these three dimensions provide a better understanding of the lexical quality of the crowdfunding campaign text.

Control Variables

Control variables are included to isolate the effect of the variables described above on the campaign performance. Without these controls, the results could be incorrect because of biases. These undesirable factors staying constant allow for a cleaner identification of the effect of interest. Here are the applied control variables:

- *log_goal* is the logarithm of the funding target in USD. Campaigns with higher goals are harder to complete their funding but may raise more money. The log transformation is applied because of the skew of the goal distribution.
- *duration* measures the length of the campaign in days. Longer campaigns have more time to attract backers.
- *staff_pick* is a binary variable equal to 1 if the campaign was highlighted by Kickstarter's editorial team and 0 if it wasn't. This can have a strong influence because it provides visibility to the project.
- *experience* captures the number of campaigns previously launched by the creator. More experienced creators may be better at running campaigns.
- *competing_campaigns* indicates the number of other campaigns active at the same time on the platform. A higher competition can reduce the visibility and funding of a campaign.
- *google_trend* is an approximation of the public's interest in the campaign's topic at the time of launch based on Google search volume. Higher interest in a topic may lead to greater funding.
- *US_Project* is another binary variable equal to 1 if the campaign is from the United States and 0 if it isn't. This can capture potential geographic differences.
- *category* indicates the type of project. Each sector has different funding dynamics and audience sizes that are linked to the sector. Some of them might have more interest and potential backers.
- *year* indicates the year of the campaign. This controls for trends over time such as the democratization of crowdfunding, changes in backer behavior, etc.

4.3. Research Method

The method chosen to test the three hypotheses combines ordinary least squares regression⁸ (OLS) and logistic regression⁹, depending on the nature of the dependent variable. For campaign success (binary: success or not), we use logistic regression. For the amount raised (continuous values), we use OLS with a log transformation. Additional explanations for each component of each formula are available in appendixes, section 9.1. Formula Explanations.

For H1, to test if the use of AI affects campaign performance, we estimate the following specifications for both dependent variables:

$$(1) \quad Success_i = \beta_0 + \beta_1 AIUse_i + \gamma X_i + \varepsilon_i$$

$$(2) \quad \log Pledged_i = \beta_0 + \beta_1 AIUse_i + \gamma X_i + \varepsilon_i$$

Where AIUse measures the proportion of content generated by AI from 0 to 1 (0% to 100%) in the campaign description. The coefficient β_1 captures the marginal effect of an increase of one unit in AI proportion on campaign performance. A negative β_1 suggests that higher AI usage is associated with lower probability of success or lower amount raised. A positive β_1 suggests the opposite, that AI usage is associated with higher probability of success or amount raised.

For H2, to test if AI usage affects the textual characteristics of campaign descriptions, we estimate one regression per textual variable for a total of 9 regressions all using the following structure:

$$(3) \text{ to } (11) \quad TextVariable_i = \beta_0 + \beta_1 AIUse_i + \gamma X_i + \varepsilon_i$$

Where TextVariable represents each of the nine textual variables tested, listed and described in section 4.2. Variables – Textual Variables. The coefficient β_1 captures if AI usage systematically impacts the textual dimension. A significant β_1 indicates that campaigns using more content generated by AI have various textual characteristics that we can measure.

⁸ Ordinary least squares regression (OLS) is a method for estimating the relationship between a dependent variable and one or more independent variables by choosing coefficients that minimize the sum of squared differences between observed values and the model's predicted values. (Mućk & SGH Warsaw School of Economics, 2021) (Wooldridge, 2022)

⁹ Logistic regression is a statistical model used when the dependent variable is binary and we want to estimate how predictors affect the probability of the event occurring. (Harris, 2021)

For H3, to test if textual characteristics affect campaign performance, we estimate the following specifications:

$$(12) \quad Success_i = \beta_0 + \sum_{k=1}^9 \beta_k TextVariable_{k,i} + \gamma X_i + \varepsilon_i$$

$$(13) \quad \log Pledged_i = \beta_0 + \sum_{k=1}^9 \beta_k TextVariable_{k,i} + \gamma X_i + \varepsilon_i$$

Where all nine textual variables are all included at the same time as independent variables. Each coefficient β_k captures the independent effect of the linked textual variable on the campaign performance. This allows holding all other variables constant.

In all 13 formulas, X_i represents the control variables listed and described in section 4.2. Variables – Control Variables. ε_i represents the residual values, so everything that influences campaign performance but is not included in the regression.

5. Results

Complete results of each hypothesis are available in appendixes, section 9.2. Raw Results. The figures below do not contain all explained results.

Hypothesis 1: Artificial Intelligence's Impact On Crowdfunding

The results provide evidence that using AI to write crowdfunding descriptions has a significant and negative impact on campaign performance. This observation is true both in terms of probability of reaching the funding goal and the total amount raised.

	AI_use_per
Succes	-0.734*** (0.059)
Num.Obs.	81942
AIC	79872.4
* p < 0.1, ** p < 0.05, *** p < 0.01	
Control variables included.	
Impact AI on success	

Figure 1

	AI_use_per
Amount raised	-0.893*** (0.067)
Num.Obs.	81942
R2	0.258
R2 Adj.	0.257
* p < 0.1, ** p < 0.05, *** p < 0.01	
Control variables included.	
Impact AI on Amount Raised	

Figure 2

The regression shown in *Figure 1*¹⁰ that measures the success demonstrates that *AI_use_per* has a negative coefficient of -0.734. This observation is highly significant (***)¹¹. The average marginal effect is -0.1186*. This indicates that a campaign moving from 0% to 100% generated by AI would see its probability of success reduced by approximately 11.86 percentage point¹² on average.

The OLS regression shown in *Figure 2*¹³ confirms this finding for the amount raised. The coefficient on *AI_use_per* is -0.893. This time, the observation is also highly significant because

¹⁰ AIC is a measure of the quality of a statistical model. The lower it is, the better. AIC cannot be interpreted in absolute terms. It only makes sense when compared to another model. (Sutherland et al., 2023)

¹¹ This means that the p-value is close to zero. The scale is shown on all regression's figures. It measures the chance of obtaining the observed results by pure chance. The closest is the p-value to zero, the most likely it is to happen.

* Not shown in figures. Available in section 9.2. Raw Results

¹² A percentage point measures the difference between two percentages (of success for example). A percent measures a change compared to a starting value. If success rate drops from 50% to 38%, that is 12 percentage points lower but also 24% lower relative to the starting value. (Office For National Statistics, s. d.)

¹³ R² measures the proportion of variance in the dependent variable that is explained by the independent variables in the model. The higher, the better. R² adjusted penalizes the addition of variables that do not improve the model. (Sapra, 2014)

we observe the same p-value. This regression also indicates that a 10% increase in AI usage is associated with an 8.54%* decrease in the total amount raised. The use of AI explains approximately 25.76% (R^2) of the change in the log of the amount raised. This is still significant given the complexity of crowdfunding dynamics.

Both analyses support H1-B. The utilization of AI in crowdfunding descriptions is negatively associated with campaign performance. In other words, campaigns with higher rates of AI content tend to raise less money and are less likely to reach their funding goal.

Hypothesis 2: Artificial Intelligence’s Text Writing Style

The results provide evidence that AI use systematically changes the textual characteristics of crowdfunding campaign descriptions. All nine textual variables show a significant relationship with *AI_use_per*. This confirms H2-A. which means that AI writing has an impact on several variables as explained below.

	AboutWords	RiskWords	Flesch_withAI	SentimentLM_withAI	SentimentHE_withAI	LC_PC1_withAI	Soph_PC1_withAI	Var_PC1_withAI	Tone_withAI
Utilisation IA (proportion)	-31.076** (15.511)	121.536*** (2.507)	-29.570*** (0.435)	0.020*** (0.002)	0.004*** (0.001)	3.744*** (0.062)	2.930*** (0.039)	2.650*** (0.059)	0.376*** (0.017)
Num.Obs.	81942	81942	81837	81802	81802	81938	81938	81938	81802
R2	0.107	0.105	0.128	0.027	0.020	0.318	0.209	0.280	0.024

* p < 0.1, ** p < 0.05, *** p < 0.01

Control variables included.

Impact AI on all Textual Variables

Figure 3

On the one hand, the effect of AI use on the length of the “about” section is negative and significant (**). This suggests that campaigns with more AI content tend to have slightly shorter “about” sections. On the other hand, the effect on the “risk” section is more positive and more significant. This indicates that AI content is a lot more concentrated on the “risk” section.

AI use has a strong negative effect on the Flesch readability score. This means that campaigns with more AI content are significantly harder to read.

Both sentiment measures are positively and significantly associated with AI usage. The Loughran-McDonald sentiment score increases with AI use, likewise for the Henry’s sentiment score. This indicates that AI content tends to have a more positive tone.

All three dimensions of lexical quality are positively and highly significantly associated with AI use. Lexical complexity increases with AI usage, same for lexical sophistication, and lexical variation. This suggests that AI-generated content systematically uses more complex and more varied vocabulary compared to texts written by humans in the context of crowdfunding descriptions.

The tone score is also positively and significantly associated with AI use. This confirms that AI campaign descriptions tend to adopt a more positive and optimistic tone.

H2-A is confirmed across all nine textual variables. This means that AI changes the writing style of crowdfunding descriptions in a consistent and systematic way.

Hypothesis 3: Text Writing Style Impact On Crowdfunding

	Success (Logit)
AboutWords	0.001*** (0.000)
RiskWords	-0.001*** (0.000)
Flesch_withAI	-0.002*** (0.001)
SentimentLM_withAI	2.093*** (0.206)
SentimentHE_withAI	2.951*** (0.244)
LC_PC1_withAI	-0.332*** (0.071)
Soph_PC1_withAI	0.157*** (0.028)
Var_PC1_withAI	0.421*** (0.066)
Tone_withAI	-0.118*** (0.023)
Num.Obs.	85601
AIC	79015.2
* p < 0.1, ** p < 0.05, *** p < 0.01	
Control variables included.	
Impact of textual variables on success	

Figure 4

	log(pledged) (OLS)
AboutWords	0.001*** (0.000)
RiskWords	0.000*** (0.000)
Flesch_withAI	-0.003*** (0.001)
SentimentLM_withAI	1.625*** (0.198)
SentimentHE_withAI	2.988*** (0.217)
LC_PC1_withAI	-0.795*** (0.068)
Soph_PC1_withAI	0.309*** (0.027)
Var_PC1_withAI	0.891*** (0.063)
Tone_withAI	-0.027 (0.023)
Num.Obs.	85601
R2	0.339
R2 Adj.	0.339
* p < 0.1, ** p < 0.05, *** p < 0.01	
Control variables included.	
Impact of textual variables on amount raised	

Figure 5

<i>Variable</i>	Effect_Success	Sign	Effect_raised	Sign
<i>About Words</i>	+0.01pp	***	+0.11%	***
<i>Risk Words</i>	-0.02pp	***	-0.04%	***
<i>Flesch Readability</i>	-0.03pp	***	-0.34%	***
<i>Sentiment (LM)</i>	+31.71pp	***	+407.9%	***
<i>Sentiment (HE)</i>	+44.71pp	***	+1885.6%	***
<i>Lexical Complexity</i>	-5.03pp	***	-54.83%	***
<i>Lexical Sophistication</i>	+2.37pp	***	+36.22%	***
<i>Lexical Variation</i>	+6.38pp	***	+143.83%	***
<i>Overall Tone</i>	-1.79pp	***	Not significant	

Tab 1

In addition to the linear regression, *Tab 1* shows the marginal effect of an increase of 1 point of the given variable on the Success for column Effect_Success and on the Amount raised for Effect_raised. As explained above, percentage point (pp) measures the difference between two percentages. Here the probability of success and percent measure a change compared to a starting value. It is important to note that the size of the effect is directly linked to the scale of each variable. Variables with small effects are measured in large units like ranges from 0 to 100. This means that an increase of 1 point is a very small move movement. Variables with large effects are measured on a small scale, often between 0 and 1. That is why the effects look so large, particularly the 1885% for Henry sentiment. The idea here is to identify the values that have an impact, not necessarily to compare them.

The results provide evidence that the tested textual variables influence crowdfunding campaign performance. This is observable both in terms of the probability of success and the total amount raised. Eight out of nine textual variables are significant. This confirms H3-A and H3-B for those eight variables but also confirms H3-0 for the last one.

The length of the “about” section has a positive and highly significant effect on both success and amount raised. Each additional word increases the probability of success by 0.01% and is associated with a 0.11% increase in the amount raised. Even if the individual effect is small, it is highly consistent and reflects the fact that more detailed campaigns tend to inspire greater confidence. On the other hand, the length of the “risk” section has a negative significant effect on both outcomes. This suggests that detailed risk disclosures may discourage backers.

The Flesch readability score has a significant negative effect on both success and amount raised. This result suggests texts that are too easy to read are associated with worse campaign outcomes.

For sentiment, both the Loughran-McDonald and Henry sentiment scores have a strong positive effect on campaign performance. For the probability of success, an increase of one unit in LM sentiment increases the probability of success by 31.71%. On the other hand, Henry sentiment increases it by 44.71%. For the amount raised, with LM sentiment associated with a 407.9% increase and Henry sentiment with a 1,885.58% increase in the amount raised. These results confirm that a more positive emotional tone in campaign descriptions is strongly associated with better fundraising outcomes.

Lexical complexity has a significant negative effect on both outcomes. This indicates that complex vocabulary has a bad impact on campaign performance. On the other hand, lexical sophistication and lexical variation both have positive significant effects. In other words, using a rich and varied vocabulary is beneficial for campaign success.

The tone variable has a different result for the two specifications. It has a negative significant effect on the probability of success, but is not significant for the amount raised. This suggests that a more positive tone may reduce the possibility of reaching the funding goal, but it does not significantly affect the total amount raised.

H3-A and H3-B are thus both confirmed depending on the variable. The textual characteristics of crowdfunding descriptions have a significant impact on campaign performance. Combined with the results of H1 and H2, these findings allow us to state the following: AI use changes the textual characteristics of campaign descriptions, and those characteristics significantly influence the probability of success and the total amount raised of crowdfunding campaigns.

6. Discussion

Results And Literature Review

The results of this study provide a picture of the role of AI in crowdfunding campaign performance. Put together, the three hypotheses confirm the existence of a chain linking AI usage to campaign outcomes through textual characteristics.

The fact that H1-B is correct aligns with the findings of Zavertiaeva et al. (2026), who find that campaigns with high AI probability raise 66% less funds and have 16% less chance to reach their goal. When AI is used as writer on its own rather than as a tool, the output may lose the authenticity that backers value in crowdfunding. This is consistent with Lodder (2025), who finds that mandatory AI disclosure does not significantly impact success rates. This suggests that the issue may not be the fact of using AI but rather the quality of AI content. From a theoretical perspective, these results are also consistent with signaling theory. Investors' beliefs are not only impacted by the content of a disclosure but also on how it is presented.

The confirmation of H2-A across all nine textual explains how AI changes the writing style of crowdfunding descriptions. AI use is linked to an increase in lexical complexity which is consistent with the literature. Indeed, it identifies AI texts as more formal than human writing with advanced vocabulary (AlAfnan & MohdZuki, 2023) (Rodrigues et al., 2026). AI use is also associated with a positive sentiment and tone. This aligns with the literature saying that generative AI tends to produce positive language unless prompted otherwise (Neko, 2025). In the context of crowdfunding, this positivity may be seen by sophisticated backers as what Sattar & Sattar (2026) describe as "AI polished spin" that reduces the credibility of the campaign.

H3 being correct for eight out of nine variables highlights the textual factors through which AI influences campaign performance. The textual factor that has the most impact is sentiment. Both sentiment scores have a positive effect on the probability of success and the amount raised. This is in accordance with Todd et al. (2024) and Qin & Cui (2024), who find that positive tone increases investor optimism. The negative effect of lexical complexity on campaign performance is consistent with Yu & Zhao (2023), who find that textual complexity reduces judgment accuracy and market efficiency. The positive effects of lexical sophistication and lexical variation suggest that vocabulary diversity can be linked to expertise and credibility by backers.

The negative effect of readability on campaign performance is a counter-intuitive finding. A higher Flesch score, so an easier-to-read text, is associated with a lower probability of success and a lower amount raised. This may be because the platform hosts innovative projects. The backers may often be early adopters who expect a certain technical level. A text that is too simple may signal a lack of expertise. This finding is also consistent with the signaling theory. As highlighted by Tan et al. (2014) and Asay et al. (2016), sophistication can act as a credibility signal and may increase perceived legitimacy for backers. And it is important to remember that readability and lexical quality are not the same. Lexical complexity has a negative effect on performance but lexical sophistication and lexical variation both have positive effects. This indicated that what matters is not simplicity or difficulty, but the diversity of the vocabulary used. A text can be sophisticated and varied without being complex. This combination is what is valued by backers.

Finally, the tone variable not being significant for the amount raised, even being significant for the probability of success can suggest that tone's impact is different depending on whether the outcome is binary or continuous. If this finding is indeed correct, reaching the funding goal may be more sensitive to the emotion of the text.

Implications

These findings have several implications. For campaign creators, they suggest that using AI as a writer is not a good idea for the campaign. Entrepreneurs have an interest in using AI as an assistant to improve aspects of their writing to preserve the authenticity that backers prefer. For platforms like Kickstarter, these findings suggest that encouraging transparency about AI may be less important than explaining to creators the effects of AI on the quality and outcome of their campaign. For researchers, this study contributes to the literature on AI in financial contexts by providing an explanation of the effect of AI on crowdfunding outcomes in a measurable way.

7. Conclusion

The aim of this study was to investigate the role of artificial intelligence in crowdfunding campaign performance. The research was structured to understand not only the direction of the effect but also the mechanism behind AI using a sample of 85,752 campaigns. The results are consistent within all three hypotheses. AI use has a significant and a negative effect on both the probability of reaching the funding goal and the total amount raised. We also saw that AI impacts all nine textual variables tested. Eight of those textual variables significantly influence campaign performance. Together, these findings prove that AI influences crowdfunding outcomes through its impact on writing style.

Several limitations must be highlighted. The AI detection relies entirely on GPTZero, which may produce wrong results. Given the fact that this is one of the main variables of interest, it could have a significant impact on the results. In addition, the textual variables are based on the full campaign text including AI parts. That means they capture the combined effect of human and AI writing rather than just AI alone. Also, even if control variables were added, the research may still suffer from omitted variables. Furthermore, the study is limited to reward-based crowdfunding on Kickstarter and to four project categories. The results may not generalize to other platforms, other types of crowdfunding such as equity-based or donation-based, or other cultural contexts. It is also important to note that AI is quickly growing and that the results found in early 2026 may not be applicable at the end of the year. Using the exact same methodology to text the progress of AI could also be interesting.

This paper also gives future research ideas. The natural idea would be to replicate this study and to standardize the textual variables to have comparable scores to see what variable has a stronger impact on crowdfunding outcomes. It would also be interesting to distinguish between AI used as an autonomous writer and AI used as an editorial assistant. Another idea would be to examine whether the effect of AI varies across project sectors or backer profiles.

As AI becomes more accessible and powerful, using it to write important communications tends to be the easy solution. Study suggests that there is something backers respond to that goes beyond the structured sentences and vocabulary. Authenticity and the idea that a real person stands behind a project still matter. In a world where anyone can generate a campaign in a few seconds, the human behind may be more valuable. The question is no longer if AI can write, because it can. The question is for how long will people keep trusting AI?

8. References

- Akinwande, M., Adeliyi, O., & Yussuph, T. (2024). Decoding AI and Human Authorship: Nuances Revealed through NLP and Statistical Analysis. *International Journal on Cybernetics & Informatics*, 13(4), 85–103. <https://doi.org/10.5121/ijci.2024.130408>
- AlAfnan, M. A., & MohdZuki, S. F. (2023). Do Artificial Intelligence Chatbots Have a Writing Style? An Investigation into the Stylistic Features of ChatGPT-4. *Journal of Artificial Intelligence and Technology*. <https://doi.org/10.37965/jait.2023.0267>
- Alivia, L., Hartono, J., Ali, S., & Nurhayati, R. (2020). Information Disclosure Readability, Cognitive Style, and Investment Decision Making : A Web Experimental Study. Atlantis Press. <https://doi.org/10.2991/aebmr.k.200410.021>
- Asay, H. S., & Rennekamp, W. B. E. a. K. (2017). Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *The Accounting Review*, 92(4), 1–25. <https://www.jstor.org/stable/26550672>
- Asay, H. S., Elliott, W. B., & Rennekamp, K. (2016). Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *The Accounting Review*, 92(4), 1–25. <https://doi.org/10.2308/accr-51570>
- Basnet, A., Elias, M., Salganik-Shoshan, G., Walker, T., & Zhao, Y. (2025). Analyzing the market's reaction to AI narratives in corporate filings. *International Review of Financial Analysis*, 105, 104378. <https://doi.org/10.1016/j.irfa.2025.104378>
- Bello, F. (2024, October 16). Reward crowdfunding: esempi. Turbo Crowd. <https://www.turbocrowd.it/en/reward-crowdfunding-examples/>
- Bpifrance. (2024, January 1). Le crowdfunding ou financement participatif : un outil de financement de la création d'entreprise. <https://bpifrance-creation.fr/encyclopedie/financements/financement-participatif/crowdfunding-ou-financement-participatif-outil>
- BSEtec. (2025, April 15). ICO vs IPO: Key Difference and Which one to Choose? BSEtec. <https://www.bsetec.com/blog/ico-vs-ipo-key-differences/>
- Coursera. (2025, October 15). How does AI work? Basics to know. Coursera. <https://www.coursera.org/articles/how-does-ai-work>
- Cox, T. (2025, December 30). How to spot when writing is AI: 6 elements of a robot's style — Hunting the Muse. Hunting the Muse. <https://huntingthemuse.net/library/how-to-tell-if-writing-is-ai>
- De Brouwer, P. J. S. (2025). Bias in decision making in uncertainty (By Krakow School of Business) [Report]. https://www.de-brouwer.com/assets/courses/ksb_bias/bias/handouts.pdf
- Druz, M., Petzev, I., Wagner, A., & Zeckhauser, R. (2020). When Managers Change Their Tone, Analysts and Investors Change Their Tune. *Financial Analysts Journal*. https://www.hks.harvard.edu/sites/default/files/HKSEE/HKSEE%20PDFs/DruzPetzevWagnerZeckhauser_Tone_2020.pdf
- Druz, M., Wagner, A. F., & Zeckhauser, R. J. (2016). Reading Managerial Tone: How Analysts and the Market Respond to Conference Calls. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2782672>

- Eftretuei, E. & Newcastle University. (n.d.). Year and industry-level accounting narrative analysis: readability and tone variation. Newcastle University. <https://researchonline.ljmu.ac.uk/id/eprint/17710/1/SSRN-id3788236.pdf>
- Ehara, Y. & Tokyo Gakugei University. (2023). An Analytical Study of the Flesch-Kincaid Readability Formulae to Explain Their Robustness over Time. In *Tokyo Gakugei University*. <https://aclanthology.org/2024.paclic-1.94.pdf>
- Elliott, W. B., Loftus, S., & Winn, A. (2023). To read or to listen? Does disclosure delivery mode impact investors' reactions to managers' tone language? *Contemporary Accounting Research*, 41(1), 7–38. <https://doi.org/10.1111/1911-3846.12898>
- European Commission. (s.d.). *Business angels*. Internal Market, Industry, Entrepreneurship And SMEs. https://single-market-economy.ec.europa.eu/access-finance/policy-areas/business-angels_en
- Fiedler, A. & D. J. (2025). Do humans identify AI-generated text better than machines? Evidence based on excerpts from German theses☆. *ideas.repec.org*. <https://ideas.repec.org/a/eee/ireced/v49y2025ics1477388025000131.html>
- Harris, J. K. (2021). Primer on binary logistic regression. *Family Medicine and Community Health*, 9(Suppl 1), e001290. <https://doi.org/10.1136/fmch-2021-001290>
- Heaslip, E. (2026, January 27). 9 startups that were successfully crowdfunded. CO- by US Chamber of Commerce. <https://www.uschamber.com/co/start/startup/successful-crowdfunded-startups>
- Hennig, J. C., Firk, S., & Wolff, M. (2023). Credibility Signals from Soft Information: Evidence from Investor Reactions to Tone in Earnings Conference Calls. *European Accounting Review*, 34(1), 153–185. <https://doi.org/10.1080/09638180.2023.2244009>
- Henry, E., & Peytcheva, M. (2018). Earnings-Announcement narrative and investor judgment. *Accounting Horizons*, 32(3), 123–143. <https://doi.org/10.2308/acch-52121>
- Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33, 171–185. <https://doi.org/10.1016/j.irfa.2014.02.006>
- Kiger, P. J. (2023, November 7). How does AI work? HowStuffWorks. <https://science.howstuffworks.com/artificial-intelligence.htm>
- Law, R. (2026, March 5). *74% of New Webpages Include AI Content (Study of 900k Pages)*. SEO Blog by Ahrefs. <https://ahrefs.com/blog/what-percentage-of-new-content-is-ai-generated/>
- Li, G.-C. (2019). Kickstarter Structured Relational Database. Harvard Dataverse. <https://doi.org/10.7910/DVN/EOYBXM>
- Lodder, D.(2025). Disclosing AI: How transparency shapes Reward-Based Crowdfunding success. In SCHOOL OF ECONOMICS AND MANAGEMENT, TILBURG UNIVERSITY [Thesis]. <https://arno.uvt.nl/show.cgi?fid=185824>
- Ma, Y., Liu, J., Yi, F., Cheng, Q., Huang, Y., Lu, W., & Liu, X. (2023, January 24). AI vs. Human -- Differentiation Analysis of Scientific Content Generation. *arXiv.org*. <https://arxiv.org/abs/2301.10416>
- Morgan, J. (2025). Crowdfunding A Startup: Types, Strategies and benefits. <https://www.jpmmorgan.com/insights/business-planning/crowdfunding-a-startup-types-strategies-and-benefits>

- Mućk, J. & SGH Warsaw School of Economics. (2021). Linear regression. In *Advanced Applied Econometrics*. <https://e-web.sgh.waw.pl/jmuck/AAE/AdvancedAppliedEconometrics1.pdf>
- Nav, T. (2026, May 1). Get to know reward-based crowdfunding. Nav. <https://www.nav.com/business-financing-options/reward-crowdfunding/>
- ND Creative: Web, University of Notre Dame. (n.d.). *Loughran-McDonald Master Dictionary w/ Sentiment Word Lists*. Software Repository for Accounting and Finance. <https://sraf.nd.edu/loughranmcdonald-master-dictionary/>
- Neko, H. (2025, December 21). Generative AI for Financial Narratives: Crafting compliant, nuanced & Philosophy-Driven investment analysis. DEV Community. <https://dev.to/hikikomorineko/generative-ai-for-financial-narratives-crafting-compliant-nuanced-philosophy-driven-investment-29d6>
- Nicoud, A. (2026, March 18). The trends that will shape AI and tech in 2026. IBM. <https://www.ibm.com/think/news/ai-tech-trends-predictions-2026>
- Office For National Statistics. (s. d.). *Percentages and percentage points*. Office For National Statistics. <https://service-manual.ons.gov.uk/content/numbers/percentages>
- Palmer, M., Roeder, J., & Muntermann, J. (2021). Induction of a sentiment dictionary for financial analyst communication: a data-driven approach balancing machine learning and human intuition. *Journal of Business Analytics*, 5(1), 8–28. <https://doi.org/10.1080/2573234x.2021.1955022>
- Prestianni, T. (2026, January 5). 131 AI Statistics and Trends for 2026. National University. <https://www.nu.edu/blog/ai-statistics-trends/>
- Qin, Z., & Cui, M. (2024). Annual report tone and divergence of opinion: Evidence from textual analysis [Article]. *Journal of Applied Economics*, 1, 1–19. <https://doi.org/10.1080/15140326.2024.2354641> (Original work published 2024)
- Ray, S. (2025, December 8). What's next in AI: 7 trends to watch in 2026. Source. <https://news.microsoft.com/source/features/ai/whats-next-in-ai-7-trends-to-watch-in-2026/>
- Rodrigues, F. A., Sturm, N. F., & Pinheiro, F. L. (2026). A linguistic comparison between human- and AI-generated content. *iScience*, 29(3), 114976. <https://doi.org/10.1016/j.isci.2026.114976>
- Sapra, R. (2014). Using R2 with caution. *Current Medicine Research and Practice*, 4(3), 130–134. <https://doi.org/10.1016/j.cmrp.2014.06.002>
- Sattar, S., & Sattar, S. (2026, February 19). AI wrote it. Do investors still trust it? - Dickenson World. Dickenson World -. <https://dickensonworld.com/ai-wrote-it-do-investors-still-trust-it/>
- Sigma Technology Group. (2025, November 6). Top Technology Trends 2026: The AI Era Redefined by Humans. Sigma Technology. <https://sigmatechnology.com/articles/tech-trends-in-2026-signals-not-noise/>
- Singla, A., Sukharevsky, A., Yee, L., & Chui, M. (2024, May 30). *The state of AI in early 2024: Gen AI adoption spikes and starts to generate value*. McKinsey & Company. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024>
- Stellner, B. (2022). Readability of annual reports on the Vienna Stock Exchange: A test of management obfuscation hypothesis. *Central European Business Review*, 11(5), 49–66. <https://doi.org/10.18267/j.cebr.307>

- Stradi, F., & Verdickt, G. (2025). Algorithm aversion, appreciation, and investor return beliefs. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.5207430>
- Sun, Y., Xu, Y., & Zeng, X. (2025). AI-generated report, information risk and stock return. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.5956736>
- Suryani, A. W., & Fauz, D. T. A. (s. d.). BEYOND THE NUMBER : TONE ANALYSIS IN ANNUAL REPORTS. JAKI. <https://scholarhub.ui.ac.id/cgi/viewcontent.cgi?article=1754&context=jaki>
- Sutherland, C., Hare, D., Johnson, P. J., Linden, D. W., Montgomery, R. A., & Droge, E. (2023). Practical advice on variable selection and reporting using Akaike information criterion. *Proceedings of the Royal Society B Biological Sciences*, 290(2007), 20231261. <https://doi.org/10.1098/rspb.2023.1261>
- Tan, H., Wang, E. Y., & Zhou, B. (2014). How does readability influence investors' judgments? Consistency of benchmark performance matters. *The Accounting Review*, 90(1), 371–393. <https://doi.org/10.2308/accr-50857>
- Terčon, L., & Dobrovoljc, K. (2025, October 1). Linguistic Characteristics of AI-Generated Text: A survey. arXiv.org. <https://arxiv.org/abs/2510.05136>
- Thewissen, J., Shrestha, P., & Collewaert, V. (2026). Institutional Background in Crowdfunding : Trust, Signals, and Backers' Decisions. *Small Business Economics*.
- Todd, A., Bowden, J., & Moshfeghi, Y. (2024). Text-based sentiment analysis in finance: Synthesising the existing literature and exploring future directions. *Intelligent Systems in Accounting, Finance and Management/Intelligent Systems in Accounting, Finance & Management*, 31(1). <https://doi.org/10.1002/isaf.1549>
- Wan, B. a. & M. J. & A. S. & Y. Q. (2025). Tone distance: Managerial tone divergence and market reaction to earnings announcements. [ideas.repec.org. https://ideas.repec.org/a/bla/finrev/v60y2025i4p1415-1435.html](https://ideas.repec.org/a/bla/finrev/v60y2025i4p1415-1435.html)
- Wooldridge. (2022). *Introductory Econometrics* (5th ed.). <http://fmwww.bc.edu/ec-c/f2013/228/EC228.F2013.nn03.pdf>
- Ye, T., Zheng, J., Jin, J., Qiu, J., Ai, W., & Mei, Q. (2024, April 24). Using artificial intelligence to unlock crowdfunding success for small businesses. arXiv.org. <https://arxiv.org/abs/2407.09480>
- Yu, X., & Zhao, L. (2023). Textual disclosure complexity and analysts' weighting of information. *Journal of Contemporary Accounting & Economics*, 20(1), 100395. <https://doi.org/10.1016/j.jcae.2023.100395>
- Zarya, D. (2023, August 17). IPO vs Equity Crowdfunding. CrowdSpace. <https://thecrowdspace.com/blog/ipo-vs-equity-crowdfunding/>
- Zavertiaeva, M., Parshakov, P., & Сепгеевна, Ф. B. (Éds.). (2026). IDLab Seminar : The Impact of AI-Generated Project Descriptions on Crowdfunding Campaign Outcomes.

9. Appendix

9.1. Formula Explanations

Formula (1) and (2)

$$(1) \quad Success_i = \beta_0 + \beta_1 AIUse_i + \gamma X_i + \varepsilon_i$$

$$(2) \quad \log Pledged_i = \beta_0 + \beta_1 AIUse_i + \gamma X_i + \varepsilon_i$$

$Success_i$ is the dependent variable of (1). It equals 1 if campaign i successfully reached its funding goal and 0 if it did not. i refers to any specific campaign.

$\log Pledged_i$ is the dependent variable of (2). It is the natural logarithm of the total amount pledged in USD by backers for campaign i .

β_0 represents the base probability of success when all other variables are equal to zero. On its own it has little impact, but it is a necessary mathematical component of the model.

β_1 is the coefficient of interest. It captures the effect of AI usage on the probability of success. A negative β_1 means that more AI content is associated with a lower probability of success. A positive β_1 would mean the opposite.

$AIUse_i$ is the main independent variable. It measures the proportion of campaign i 's text that was identified as generated by AI by GPTZero, expressed as a value between 0 and 1. This is the variable effect we are trying to estimate.

γ is the vector of coefficients associated with the control variables. Rather than a single number, it represents all coefficients, one for each control variable included in the model.

X_i is the vector of control variables for campaign i .

ε_i is the residual error term. It captures everything that influences the success of campaign i but is not included in the model.

Formula (3) to (11)

$$(3) \text{ to } (11) \quad \text{TextVariable}_i = \beta_0 + \beta_1 \text{AIUse}_i + \gamma X_i + \varepsilon_i$$

TextVariable_i is the dependent variable. This formula is applied nine times, once for each textual variable tested. Each regression isolates the effect of AI usage on one specific textual dimension of campaign i .

β_0 represents the basevalue of the textual variable when all other variables are equal to zero. It has limited practical interpretation on its own but is a necessary part of the model.

β_1 is the coefficient of interest. It captures the effect of AI use on the specific textual variable being tested. A positive β_1 means that more AI content is associated with a higher value of the textual variable, and a negative β_1 means the opposite.

AIUse_i is the main independent variable. It measures the proportion of campaign i 's text identified as AI-generated by GPTZero, expressed as a value between 0 and 1.

γ is the vector of coefficients associated with the control variables. Rather than a single number, it represents all coefficients, one for each control variable included in the model.

X_i is the vector of control variables for campaign i .

ε_i is the residual error term. It captures everything that influences the success of campaign i but is not included in the model.

Formula (12) and (13)

$$(12) \quad Success_i = \beta_0 + \sum_{k=1}^9 \beta_k TextVariable_{k,i} + \gamma X_i + \varepsilon_i$$

$$(13) \quad \log Pledged_i = \beta_0 + \sum_{k=1}^9 \beta_k TextVariable_{k,i} + \gamma X_i + \varepsilon_i$$

$Success_i$ is the dependent variable of (1). It equals 1 if campaign i successfully reached its funding goal and 0 if it did not. i refers to any specific campaign.

$\log Pledged_i$ is the dependent variable of (2). It is the natural logarithm of the total amount pledged in USD by backers for campaign i.

$\sum_{k=1}^9 \beta_k TextVariable_{k,i}$ is the key element of this formula. Instead of including a single independent variable of interest, this formula includes all nine textual variables simultaneously.

Each β_k is the coefficient associated with the k textual variable. This captures its independent effect on the probability of success while holding all other textual variables and control variables constant.

$TextVariable_{k,i}$ refers to the value of the k textual variable for campaign i.

γ is the vector of coefficients associated with the control variables. Rather than a single number, it represents all coefficients, one for each control variable included in the model.

X_i is the vector of control variables for campaign i.

ε_i is the residual error term. It captures everything that influences the success of campaign i but is not included in the model.

9.2. Raw Results

Hypothesis 1

```
# Régressions
> formule_1a <- as.formula(paste("Success ~ AI_use_per +", formule_controles))
> reg_1a <- glm(formule_1a, data=data_crowd_funding, family=binomial(link="logit"))
>
> formule_2a <- as.formula(paste("log_pledged ~ AI_use_per +", formule_controles))
> reg_2a <- lm(formule_2a, data=data_crowd_funding)
>
> summary(reg_1a)
```

Call:

```
glm(formula = formule_1a, family = binomial(link = "logit"),
    data = data_crowd_funding)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.987e+00	7.803e-01	2.546	0.01090 *
AI_use_per	-7.336e-01	5.922e-02	-12.388	< 2e-16 ***
log_goal	-4.288e-01	5.873e-03	-73.008	< 2e-16 ***
duration	-2.186e-02	9.868e-04	-22.154	< 2e-16 ***
US_Project1	4.691e-03	1.799e-02	0.261	0.79426
staff_pick	2.513e+00	3.535e-02	71.079	< 2e-16 ***
experience	2.892e-01	6.917e-03	41.806	< 2e-16 ***
competing_campaigns	-1.407e-04	5.304e-05	-2.652	0.00801 **
google_trend	-9.397e-03	1.834e-03	-5.123	3e-07 ***
categorydesign	1.395e+00	4.094e-02	34.086	< 2e-16 ***
categorygames	1.280e+00	4.493e-02	28.495	< 2e-16 ***
categorytechnology	7.337e-01	4.208e-02	17.437	< 2e-16 ***
year2011	2.158e+00	7.889e-01	2.735	0.00623 **
year2012	1.784e+00	7.826e-01	2.280	0.02260 *
year2013	1.251e+00	7.846e-01	1.595	0.11082
year2014	6.732e-01	7.865e-01	0.856	0.39207
year2015	9.030e-01	7.850e-01	1.150	0.24999
year2016	8.241e-01	7.808e-01	1.056	0.29118
year2017	9.641e-01	7.802e-01	1.236	0.21656
year2018	1.070e+00	7.791e-01	1.374	0.16952
year2019	1.286e+00	7.789e-01	1.651	0.09873 .
year2020	1.252e+00	7.786e-01	1.609	0.10771
year2021	1.719e+00	7.786e-01	2.207	0.02728 *
year2022	1.692e+00	7.785e-01	2.174	0.02971 *
year2023	1.753e+00	7.777e-01	2.254	0.02417 *
year2024	1.578e+00	7.781e-01	2.028	0.04258 *
year2025	1.239e+00	7.790e-01	1.590	0.11175

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 113596 on 81941 degrees of freedom
 Residual deviance: 79818 on 81915 degrees of freedom
 (3810 observations deleted due to missingness)
 AIC: 79872

Number of Fisher Scoring iterations: 6

> summary(reg_2a)

Call:

lm(formula = formule_2a, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-10.8972	-1.5379	0.3684	1.8956	8.3728

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.924e+00	9.262e-01	2.077	0.037803 *
AI_use_per	-8.930e-01	6.666e-02	-13.397	< 2e-16 ***
log_goal	2.145e-01	5.802e-03	36.962	< 2e-16 ***
duration	-1.968e-02	1.079e-03	-18.245	< 2e-16 ***
US_Project1	1.056e-02	2.007e-02	0.526	0.598648
staff_pick	3.236e+00	3.198e-02	101.212	< 2e-16 ***
experience	6.816e-02	2.162e-03	31.534	< 2e-16 ***
competing_campaigns	-1.977e-04	5.701e-05	-3.469	0.000523 ***
google_trend	-2.126e-02	1.947e-03	-10.918	< 2e-16 ***
categorydesign	2.386e+00	4.567e-02	52.251	< 2e-16 ***
categorygames	2.078e+00	5.001e-02	41.555	< 2e-16 ***
categorytechnology	1.031e+00	4.709e-02	21.898	< 2e-16 ***
year2011	2.494e+00	9.349e-01	2.668	0.007626 **
year2012	3.250e+00	9.291e-01	3.499	0.000468 ***
year2013	3.453e+00	9.305e-01	3.711	0.000206 ***
year2014	2.290e+00	9.322e-01	2.457	0.014017 *
year2015	2.071e+00	9.306e-01	2.226	0.026043 *
year2016	2.115e+00	9.266e-01	2.282	0.022467 *
year2017	2.409e+00	9.261e-01	2.601	0.009286 **
year2018	2.623e+00	9.252e-01	2.835	0.004577 **
year2019	2.795e+00	9.250e-01	3.021	0.002517 **
year2020	2.790e+00	9.247e-01	3.017	0.002552 **
year2021	3.082e+00	9.247e-01	3.333	0.000858 ***
year2022	2.882e+00	9.246e-01	3.117	0.001829 **
year2023	2.950e+00	9.237e-01	3.193	0.001406 **
year2024	2.690e+00	9.240e-01	2.911	0.003600 **
year2025	2.443e+00	9.248e-01	2.642	0.008246 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.768 on 81915 degrees of freedom

(3810 observations deleted due to missingness)
 Multiple R-squared: 0.2576, Adjusted R-squared: 0.2573
 F-statistic: 1093 on 26 and 81915 DF, p-value: < 2.2e-16

```
# Effets marginaux (Succes)
> marg_1a <- margins(reg_1a, variables="AI_use_per")
> cat("AI_use_per → P(Success) :", round(summary(marg_1a)$AME, 4),
+     "(p =", round(summary(marg_1a)$p, 3), ")\n")
AI_use_per → P(Success) : -0.1186 (p = 0 )
>
> # Effets économiques (Amount Raised)
> coef_2a <- coef(reg_2a)["AI_use_per"]
> p_2a <- summary(reg_2a)$coefficients["AI_use_per", "Pr(>|t|)"]
> cat("AI_use_per : coef =", round(coef_2a, 4), "| p =", round(p_2a, 3), "\n")
AI_use_per : coef = -0.893 | p = 0
> cat(" → +10 pp d'IA :", round((exp(coef_2a * 0.1) - 1) * 100, 2), "% sur usd_pledged_w\n\n")
→ +10 pp d'IA : -8.54 % sur usd_pledged_w
```

Hypothesis 2

```
# Une régression OLS par variable textuelle (variable dépendante)
>
> reg_h2_list <- lapply(vars_textuelles, function(var) {
+   formule <- as.formula(paste(var, "~ AI_use_per +", formule_controls))
+   lm(formule, data=data_crowd_funding)
+ })
> names(reg_h2_list) <- vars_textuelles
>
> # Vérifier les résultats
> lapply(reg_h2_list, summary)
$AboutWords
```

Call:
 lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-1564.5	-413.7	-143.3	249.2	5158.0

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-448.75438	215.50499	-2.082	0.03731 *
AI_use_per	-31.07564	15.51055	-2.004	0.04513 *
log_goal	72.55483	1.35004	53.743	< 2e-16 ***
duration	-0.96539	0.25098	-3.847	0.00012 ***
US_Project1	-73.04672	4.66957	-15.643	< 2e-16 ***
staff_pick	321.43647	7.44032	43.202	< 2e-16 ***
experience	7.18675	0.50297	14.289	< 2e-16 ***

competing_campaigns	-0.08813	0.01326	-6.644	3.07e-11	***
google_trend	-1.57006	0.45309	-3.465	0.00053	***
categorydesign	180.11898	10.62672	16.950	< 2e-16	***
categorygames	350.37346	11.63619	30.111	< 2e-16	***
categorytechnology	150.89074	10.95731	13.771	< 2e-16	***
year2011	214.71541	217.52498	0.987	0.32360	
year2012	536.46531	216.17463	2.482	0.01308	*
year2013	664.20016	216.51380	3.068	0.00216	**
year2014	484.81832	216.90245	2.235	0.02541	*
year2015	404.34743	216.52787	1.867	0.06185	.
year2016	384.62544	215.60452	1.784	0.07444	.
year2017	389.70151	215.47913	1.809	0.07053	.
year2018	407.72291	215.26631	1.894	0.05822	.
year2019	528.36146	215.22646	2.455	0.01409	*
year2020	508.62988	215.15919	2.364	0.01808	*
year2021	557.25137	215.15494	2.590	0.00960	**
year2022	506.24871	215.14227	2.353	0.01862	*
year2023	538.68035	214.92708	2.506	0.01220	*
year2024	518.20575	215.00504	2.410	0.01595	*
year2025	523.36436	215.19376	2.432	0.01502	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 644 on 81915 degrees of freedom
(3810 observations deleted due to missingness)

Multiple R-squared: 0.1068, Adjusted R-squared: 0.1065

F-statistic: 376.7 on 26 and 81915 DF, p-value: < 2.2e-16

\$RiskWords

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-286.0	-59.0	-23.0	29.7	4019.8

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.158e+02	3.484e+01	-3.325	0.000885 ***
AI_use_per	1.215e+02	2.507e+00	48.469	< 2e-16 ***
log_goal	1.378e+01	2.183e-01	63.132	< 2e-16 ***
duration	-3.657e-01	4.057e-02	-9.014	< 2e-16 ***
US_Project1	-2.965e+00	7.549e-01	-3.927	8.61e-05 ***
staff_pick	2.944e+01	1.203e+00	24.476	< 2e-16 ***
experience	9.542e-01	8.131e-02	11.735	< 2e-16 ***
competing_campaigns	-4.003e-03	2.144e-03	-1.867	0.061943 .
google_trend	-2.580e-01	7.325e-02	-3.523	0.000427 ***
categorydesign	1.611e+01	1.718e+00	9.380	< 2e-16 ***

```

categorygames      2.033e+01 1.881e+00 10.805 < 2e-16 ***
categorytechnology 6.141e+00 1.771e+00 3.467 0.000527 ***
year2011           -3.100e+00 3.517e+01 -0.088 0.929754
year2012           7.814e+01 3.495e+01 2.236 0.025363 *
year2013           1.451e+02 3.500e+01 4.146 3.38e-05 ***
year2014           1.337e+02 3.507e+01 3.814 0.000137 ***
year2015           1.216e+02 3.500e+01 3.473 0.000515 ***
year2016           1.210e+02 3.486e+01 3.472 0.000518 ***
year2017           1.265e+02 3.484e+01 3.630 0.000283 ***
year2018           1.229e+02 3.480e+01 3.530 0.000415 ***
year2019           1.134e+02 3.479e+01 3.260 0.001115 **
year2020           1.145e+02 3.478e+01 3.293 0.000993 ***
year2021           1.138e+02 3.478e+01 3.272 0.001070 **
year2022           1.079e+02 3.478e+01 3.102 0.001926 **
year2023           1.053e+02 3.475e+01 3.029 0.002453 **
year2024           1.047e+02 3.476e+01 3.012 0.002592 **
year2025           9.952e+01 3.479e+01 2.861 0.004230 **

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 104.1 on 81915 degrees of freedom
(3810 observations deleted due to missingness)

Multiple R-squared: 0.1045, Adjusted R-squared: 0.1043

F-statistic: 367.8 on 26 and 81915 DF, p-value: < 2.2e-16

\$Flesch_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-2096.71	-7.92	0.63	9.00	87.66

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.597e+01	6.035e+00	10.932	< 2e-16 ***
AI_use_per	-2.957e+01	4.353e-01	-67.936	< 2e-16 ***
log_goal	-1.362e+00	3.783e-02	-36.008	< 2e-16 ***
duration	-2.217e-03	7.032e-03	-0.315	0.752622
US_Project1	1.696e+00	1.308e-01	12.962	< 2e-16 ***
staff_pick	-1.461e+00	2.085e-01	-7.010	2.4e-12 ***
experience	-1.242e-02	1.409e-02	-0.882	0.377945
competing_campaigns	1.221e-03	3.718e-04	3.284	0.001024 **
google_trend	1.273e-02	1.270e-02	1.002	0.316108
categorydesign	-4.860e+00	2.978e-01	-16.320	< 2e-16 ***
categorygames	-1.151e+00	3.261e-01	-3.529	0.000417 ***
categorytechnology	-6.835e+00	3.070e-01	-22.265	< 2e-16 ***
year2011	-1.111e+01	6.091e+00	-1.824	0.068128 .

year2012	-1.180e+00	6.054e+00	-0.195	0.845478
year2013	3.678e+00	6.063e+00	0.607	0.544148
year2014	3.724e+00	6.074e+00	0.613	0.539823
year2015	4.065e+00	6.063e+00	0.670	0.502638
year2016	3.805e+00	6.038e+00	0.630	0.528572
year2017	3.843e+00	6.034e+00	0.637	0.524212
year2018	3.381e+00	6.028e+00	0.561	0.574864
year2019	3.266e+00	6.027e+00	0.542	0.587909
year2020	2.310e+00	6.025e+00	0.383	0.701413
year2021	2.186e+00	6.025e+00	0.363	0.716706
year2022	2.588e+00	6.025e+00	0.430	0.667545
year2023	2.035e+00	6.019e+00	0.338	0.735273
year2024	4.102e-01	6.021e+00	0.068	0.945677
year2025	1.247e+00	6.026e+00	0.207	0.836007

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 18.03 on 81810 degrees of freedom
(3915 observations deleted due to missingness)

Multiple R-squared: 0.1282, Adjusted R-squared: 0.1279

F-statistic: 462.6 on 26 and 81810 DF, p-value: < 2.2e-16

\$SentimentLM_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-0.96475	-0.03079	0.00419	0.03720	1.03946

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-6.468e-02	2.124e-02	-3.044	0.00233 **
AI_use_per	2.012e-02	1.532e-03	13.127	< 2e-16 ***
log_goal	1.155e-03	1.332e-04	8.670	< 2e-16 ***
duration	1.228e-05	2.476e-05	0.496	0.61996
US_Project1	-2.351e-03	4.607e-04	-5.102	3.36e-07 ***
staff_pick	7.839e-03	7.339e-04	10.681	< 2e-16 ***
experience	6.371e-04	4.960e-05	12.845	< 2e-16 ***
competing_campaigns	-3.008e-06	1.309e-06	-2.298	0.02157 *
google_trend	-1.812e-04	4.471e-05	-4.054	5.04e-05 ***
categorydesign	1.030e-02	1.049e-03	9.826	< 2e-16 ***
categorygames	6.548e-03	1.148e-03	5.703	1.18e-08 ***
categorytechnology	1.042e-02	1.081e-03	9.641	< 2e-16 ***
year2011	3.712e-02	2.144e-02	1.731	0.08347 .
year2012	3.605e-02	2.131e-02	1.691	0.09076 .
year2013	1.871e-02	2.134e-02	0.877	0.38063
year2014	1.553e-02	2.138e-02	0.726	0.46774

year2015	1.567e-02	2.135e-02	0.734	0.46295
year2016	1.768e-02	2.125e-02	0.832	0.40542
year2017	1.965e-02	2.124e-02	0.925	0.35488
year2018	1.985e-02	2.122e-02	0.935	0.34957
year2019	1.532e-02	2.122e-02	0.722	0.47023
year2020	1.005e-02	2.121e-02	0.474	0.63576
year2021	1.014e-02	2.121e-02	0.478	0.63271
year2022	1.140e-02	2.121e-02	0.537	0.59104
year2023	2.892e-02	2.119e-02	1.365	0.17230
year2024	1.408e-02	2.120e-02	0.664	0.50638
year2025	1.228e-02	2.121e-02	0.579	0.56280

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06349 on 81775 degrees of freedom
(3950 observations deleted due to missingness)

Multiple R-squared: 0.02687, Adjusted R-squared: 0.02656

F-statistic: 86.84 on 26 and 81775 DF, p-value: < 2.2e-16

\$SentimentHE_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-1.00580	-0.01663	0.00599	0.02362	0.50999

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.080e-02	1.965e-02	-0.550	0.58258
AI_use_per	3.855e-03	1.418e-03	2.719	0.00654 **
log_goal	2.480e-03	1.232e-04	20.122	< 2e-16 ***
duration	-9.913e-05	2.291e-05	-4.327	1.51e-05 ***
US_Project1	-1.209e-03	4.262e-04	-2.836	0.00457 **
staff_pick	9.727e-03	6.789e-04	14.326	< 2e-16 ***
experience	6.141e-04	4.588e-05	13.385	< 2e-16 ***
competing_campaigns	1.816e-06	1.211e-06	1.499	0.13376
google_trend	-2.169e-04	4.136e-05	-5.244	1.57e-07 ***
categorydesign	6.689e-03	9.701e-04	6.896	5.41e-12 ***
categorygames	2.806e-03	1.062e-03	2.642	0.00825 **
categorytechnology	-2.958e-03	1.000e-03	-2.957	0.00310 **
year2011	2.398e-03	1.984e-02	0.121	0.90377
year2012	2.028e-03	1.972e-02	0.103	0.91806
year2013	-2.137e-03	1.975e-02	-0.108	0.91382
year2014	-5.034e-03	1.978e-02	-0.254	0.79911
year2015	-6.931e-03	1.975e-02	-0.351	0.72560
year2016	-8.620e-03	1.966e-02	-0.438	0.66111
year2017	-7.801e-03	1.965e-02	-0.397	0.69140

year2018	-4.583e-03	1.963e-02	-0.233	0.81541
year2019	-8.258e-03	1.963e-02	-0.421	0.67397
year2020	-1.238e-02	1.962e-02	-0.631	0.52812
year2021	-1.072e-02	1.962e-02	-0.547	0.58471
year2022	-9.615e-03	1.962e-02	-0.490	0.62410
year2023	-6.006e-03	1.960e-02	-0.306	0.75931
year2024	-1.763e-02	1.961e-02	-0.899	0.36862
year2025	-1.569e-02	1.963e-02	-0.799	0.42416

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05873 on 81775 degrees of freedom
(3950 observations deleted due to missingness)

Multiple R-squared: 0.02024, Adjusted R-squared: 0.01993

F-statistic: 64.98 on 26 and 81775 DF, p-value: < 2.2e-16

\$LC_PC1_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

	Min	1Q	Median	3Q	Max
	-13.125	-1.449	0.054	1.519	45.464

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.516e-01	8.554e-01	0.528	0.597556
AI_use_per	3.744e+00	6.157e-02	60.815	< 2e-16 ***
log_goal	3.705e-01	5.359e-03	69.124	< 2e-16 ***
duration	-8.466e-03	9.962e-04	-8.498	< 2e-16 ***
US_Project1	-7.335e-02	1.854e-02	-3.958	7.58e-05 ***
staff_pick	1.046e+00	2.953e-02	35.432	< 2e-16 ***
experience	1.445e-02	1.997e-03	7.238	4.60e-13 ***
competing_campaigns	-9.792e-05	5.265e-05	-1.860	0.062919 .
google_trend	-1.457e-02	1.798e-03	-8.103	5.45e-16 ***
categorydesign	7.291e-01	4.218e-02	17.285	< 2e-16 ***
categorygames	6.266e-01	4.619e-02	13.567	< 2e-16 ***
categorytechnology	4.183e-01	4.349e-02	9.618	< 2e-16 ***
year2011	2.035e+00	8.634e-01	2.357	0.018419 *
year2012	6.612e-01	8.581e-01	0.771	0.440991
year2013	-3.027e+00	8.594e-01	-3.522	0.000428 ***
year2014	-3.436e+00	8.609e-01	-3.990	6.60e-05 ***
year2015	-3.732e+00	8.595e-01	-4.342	1.41e-05 ***
year2016	-3.894e+00	8.558e-01	-4.550	5.38e-06 ***
year2017	-3.746e+00	8.553e-01	-4.379	1.19e-05 ***
year2018	-3.875e+00	8.545e-01	-4.535	5.77e-06 ***
year2019	-4.302e+00	8.543e-01	-5.035	4.78e-07 ***
year2020	-4.405e+00	8.540e-01	-5.158	2.50e-07 ***

year2021	-4.418e+00	8.540e-01	-5.173	2.30e-07	***
year2022	-3.796e+00	8.540e-01	-4.445	8.78e-06	***
year2023	-1.028e+00	8.531e-01	-1.206	0.227982	
year2024	-4.783e+00	8.534e-01	-5.605	2.09e-08	***
year2025	-4.783e+00	8.542e-01	-5.599	2.16e-08	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.556 on 81911 degrees of freedom
(3814 observations deleted due to missingness)

Multiple R-squared: 0.3178, Adjusted R-squared: 0.3176

F-statistic: 1468 on 26 and 81911 DF, p-value: < 2.2e-16

\$Soph_PC1_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-6.071	-1.040	-0.145	0.836	52.470

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.041e-01	5.413e-01	0.747	0.455360
AI_use_per	2.930e+00	3.896e-02	75.206	< 2e-16 ***
log_goal	1.037e-01	3.391e-03	30.564	< 2e-16 ***
duration	-4.234e-03	6.304e-04	-6.716	1.88e-11 ***
US_Project1	-6.904e-02	1.173e-02	-5.886	3.97e-09 ***
staff_pick	4.225e-01	1.869e-02	22.608	< 2e-16 ***
experience	3.666e-03	1.263e-03	2.902	0.003710 **
competing_campaigns	1.418e-04	3.332e-05	4.255	2.09e-05 ***
google_trend	-4.092e-03	1.138e-03	-3.596	0.000324 ***
categorydesign	2.458e-01	2.669e-02	9.210	< 2e-16 ***
categorygames	5.513e-02	2.923e-02	1.886	0.059273 .
categorytechnology	3.106e-01	2.752e-02	11.285	< 2e-16 ***
year2011	1.195e+00	5.464e-01	2.186	0.028801 *
year2012	2.886e-01	5.430e-01	0.531	0.595097
year2013	-1.292e+00	5.438e-01	-2.376	0.017504 *
year2014	-1.430e+00	5.448e-01	-2.625	0.008666 **
year2015	-1.465e+00	5.439e-01	-2.694	0.007057 **
year2016	-1.497e+00	5.416e-01	-2.764	0.005719 **
year2017	-1.455e+00	5.412e-01	-2.689	0.007164 **
year2018	-1.524e+00	5.407e-01	-2.819	0.004813 **
year2019	-1.738e+00	5.406e-01	-3.215	0.001306 **
year2020	-1.729e+00	5.404e-01	-3.199	0.001378 **
year2021	-1.706e+00	5.404e-01	-3.157	0.001596 **
year2022	-1.456e+00	5.404e-01	-2.694	0.007069 **
year2023	-1.940e-01	5.398e-01	-0.359	0.719395

```

year2024      -1.731e+00  5.400e-01 -3.206 0.001346 **
year2025      -1.721e+00  5.405e-01 -3.183 0.001457 **

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.618 on 81911 degrees of freedom
(3814 observations deleted due to missingness)

Multiple R-squared: 0.2087, Adjusted R-squared: 0.2085

F-statistic: 831.1 on 26 and 81911 DF, p-value: < 2.2e-16

\$Var_PC1_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

```

      Min       1Q   Median       3Q      Max
-12.1314 -1.2947  0.1425  1.4952  21.7290

```

Coefficients:

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.351e-01  8.133e-01  0.412 0.680306
AI_use_per      2.650e+00  5.854e-02 45.277 < 2e-16 ***
log_goal        3.511e-01  5.096e-03 68.897 < 2e-16 ***
duration       -7.513e-03  9.472e-04 -7.931 2.19e-15 ***
US_Project1    -4.982e-02  1.762e-02 -2.827 0.004704 **
staff_pick      9.418e-01  2.808e-02 33.538 < 2e-16 ***
experience       1.412e-02  1.898e-03  7.439 1.02e-13 ***
competing_campaigns -1.620e-04  5.006e-05 -3.235 0.001217 **
google_trend    -1.394e-02  1.710e-03 -8.154 3.58e-16 ***
categorydesign    6.785e-01  4.011e-02 16.918 < 2e-16 ***
categorygames    6.414e-01  4.392e-02 14.606 < 2e-16 ***
categorytechnology 3.037e-01  4.135e-02  7.343 2.10e-13 ***
year2011         1.612e+00  8.209e-01  1.963 0.049639 *
year2012         5.653e-01  8.159e-01  0.693 0.488393
year2013        -2.642e+00  8.171e-01 -3.233 0.001226 **
year2014        -3.024e+00  8.186e-01 -3.694 0.000221 ***
year2015        -3.328e+00  8.172e-01 -4.073 4.64e-05 ***
year2016        -3.489e+00  8.137e-01 -4.288 1.80e-05 ***
year2017        -3.348e+00  8.132e-01 -4.117 3.85e-05 ***
year2018        -3.458e+00  8.124e-01 -4.256 2.08e-05 ***
year2019        -3.866e+00  8.123e-01 -4.760 1.94e-06 ***
year2020        -3.986e+00  8.120e-01 -4.909 9.19e-07 ***
year2021        -4.005e+00  8.120e-01 -4.932 8.17e-07 ***
year2022        -3.456e+00  8.120e-01 -4.257 2.07e-05 ***
year2023        -1.089e+00  8.111e-01 -1.342 0.179550
year2024        -4.406e+00  8.114e-01 -5.430 5.64e-08 ***
year2025        -4.421e+00  8.122e-01 -5.444 5.22e-08 ***

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.431 on 81911 degrees of freedom
(3814 observations deleted due to missingness)

Multiple R-squared: 0.2798, Adjusted R-squared: 0.2796

F-statistic: 1224 on 26 and 81911 DF, p-value: < 2.2e-16

\$Tone_withAI

Call:

lm(formula = formule, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-8.1494	-0.3531	0.0276	0.4092	5.2791

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.020e-01	2.381e-01	-1.268	0.20468
AI_use_per	3.759e-01	1.717e-02	21.884	< 2e-16 ***
log_goal	3.160e-02	1.493e-03	21.163	< 2e-16 ***
duration	8.327e-04	2.775e-04	3.000	0.00270 **
US_Project1	-2.302e-02	5.163e-03	-4.459	8.26e-06 ***
staff_pick	8.365e-02	8.225e-03	10.171	< 2e-16 ***
experience	5.563e-03	5.558e-04	10.008	< 2e-16 ***
competing_campaigns	-4.733e-05	1.467e-05	-3.226	0.00126 **
google_trend	-2.493e-03	5.010e-04	-4.975	6.53e-07 ***
categorydesign	1.367e-01	1.175e-02	11.635	< 2e-16 ***
categorygames	6.993e-02	1.287e-02	5.434	5.51e-08 ***
categorytechnology	3.809e-02	1.212e-02	3.144	0.00167 **
year2011	1.357e-01	2.403e-01	0.565	0.57229
year2012	1.746e-01	2.388e-01	0.731	0.46468
year2013	7.854e-02	2.392e-01	0.328	0.74266
year2014	6.331e-02	2.396e-01	0.264	0.79164
year2015	6.109e-02	2.392e-01	0.255	0.79843
year2016	3.460e-02	2.382e-01	0.145	0.88452
year2017	7.345e-02	2.381e-01	0.309	0.75769
year2018	8.849e-02	2.378e-01	0.372	0.70985
year2019	1.589e-02	2.378e-01	0.067	0.94673
year2020	-4.689e-02	2.377e-01	-0.197	0.84364
year2021	-6.127e-02	2.377e-01	-0.258	0.79660
year2022	-6.341e-02	2.377e-01	-0.267	0.78963
year2023	-4.161e-02	2.375e-01	-0.175	0.86090
year2024	-5.128e-02	2.375e-01	-0.216	0.82909
year2025	-5.889e-02	2.377e-01	-0.248	0.80436

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7115 on 81775 degrees of freedom

(3950 observations deleted due to missingness)
 Multiple R-squared: 0.02395, Adjusted R-squared: 0.02364
 F-statistic: 77.18 on 26 and 81775 DF, p-value: < 2.2e-16

```
# Résumé des coefficients et significativité
> for(var in vars_textuelles) {
+   coef_ai <- coef(reg_h2_list[[var]]["AI_use_per"])
+   p_ai <- summary(reg_h2_list[[var]])$coefficients["AI_use_per", "Pr(>|t|)"]
+   sig <- ifelse(p_ai < 0.01, "****", ifelse(p_ai < 0.05, "***", ifelse(p_ai < 0.1, "**", "ns")))
+   cat(var, ": coef =", round(coef_ai, 4), "| p =", round(p_ai, 3), sig, "\n")
+ }
AboutWords : coef = -31.0756 | p = 0.045 **
RiskWords : coef = 121.5356 | p = 0 ***
Flesch_withAI : coef = -29.5696 | p = 0 ***
SentimentLM_withAI : coef = 0.0201 | p = 0 ***
SentimentHE_withAI : coef = 0.0039 | p = 0.007 ***
LC_PC1_withAI : coef = 3.7442 | p = 0 ***
Soph_PC1_withAI : coef = 2.93 | p = 0 ***
Var_PC1_withAI : coef = 2.6504 | p = 0 ***
Tone_withAI : coef = 0.3759 | p = 0 ***
>
```

Hypothesis 3

```
# Régressions
> # Logit : variables textuelles → Success
> reg_h3_success <- glm(formule_h3_success, data=data_crowd_funding,
+   family=binomial(link="logit"))
>
> # OLS : variables textuelles → log(pledged)
> reg_h3_pledged <- lm(formule_h3_pledged, data=data_crowd_funding)
>
> summary(reg_h3_success)
```

Call:

```
glm(formula = formule_h3_success, family = binomial(link = "logit"),
    data = data_crowd_funding)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	2.572e+00	7.578e-01	3.395	0.000687 ***
AboutWords	6.959e-04	1.569e-05	44.365	< 2e-16 ***
RiskWords	-1.118e-03	1.272e-04	-8.784	< 2e-16 ***
Flesch_withAI	-1.869e-03	5.488e-04	-3.405	0.000661 ***
SentimentLM_withAI	2.093e+00	2.056e-01	10.176	< 2e-16 ***
SentimentHE_withAI	2.951e+00	2.443e-01	12.080	< 2e-16 ***
LC_PC1_withAI	-3.319e-01	7.067e-02	-4.696	2.65e-06 ***
Soph_PC1_withAI	1.566e-01	2.793e-02	5.606	2.07e-08 ***
Var_PC1_withAI	4.209e-01	6.552e-02	6.425	1.32e-10 ***

Tone_withAI	-1.183e-01	2.307e-02	-5.128	2.93e-07	***
log_goal	-5.520e-01	6.465e-03	-85.394	< 2e-16	***
duration	-2.243e-02	1.009e-03	-22.232	< 2e-16	***
US_Project1	7.797e-02	1.827e-02	4.268	1.97e-05	***
staff_pick	2.319e+00	3.513e-02	66.016	< 2e-16	***
experience	2.749e-01	6.736e-03	40.808	< 2e-16	***
competing_campaigns	-1.275e-05	5.384e-05	-0.237	0.812739	
google_trend	-7.035e-03	1.848e-03	-3.806	0.000141	***
categorydesign	1.264e+00	4.136e-02	30.551	< 2e-16	***
categorygames	1.042e+00	4.545e-02	22.917	< 2e-16	***
categorytechnology	5.845e-01	4.270e-02	13.690	< 2e-16	***
year2011	1.845e+00	7.649e-01	2.412	0.015871	*
year2012	1.594e+00	7.591e-01	2.100	0.035746	*
year2013	1.391e+00	7.619e-01	1.826	0.067798	.
year2014	9.209e-01	7.638e-01	1.206	0.227921	
year2015	1.239e+00	7.622e-01	1.625	0.104086	
year2016	1.191e+00	7.578e-01	1.571	0.116127	
year2017	1.332e+00	7.572e-01	1.759	0.078584	.
year2018	1.422e+00	7.561e-01	1.881	0.059970	.
year2019	1.625e+00	7.560e-01	2.150	0.031562	*
year2020	1.627e+00	7.557e-01	2.153	0.031296	*
year2021	2.104e+00	7.557e-01	2.784	0.005365	**
year2022	2.031e+00	7.554e-01	2.688	0.007183	**
year2023	1.665e+00	7.540e-01	2.207	0.027280	*
year2024	1.810e+00	7.550e-01	2.398	0.016499	*
year2025	1.390e+00	7.558e-01	1.839	0.065921	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 118668 on 85600 degrees of freedom
 Residual deviance: 78945 on 85566 degrees of freedom
 (151 observations deleted due to missingness)
 AIC: 79015

Number of Fisher Scoring iterations: 6

> summary(reg_h3_pledged)

Call:

lm(formula = formule_h3_pledged, data = data_crowd_funding)

Residuals:

Min	1Q	Median	3Q	Max
-14.3409	-1.4606	0.3042	1.7769	9.8631

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.799e+00	8.746e-01	3.200	0.001375 **

```

AboutWords      1.053e-03 1.480e-05 71.151 < 2e-16 ***
RiskWords       -3.696e-04 1.220e-04 -3.029 0.002454 **
Flesch_withAI   -3.452e-03 5.167e-04 -6.681 2.38e-11 ***
SentimentLM_withAI 1.625e+00 1.980e-01 8.206 2.32e-16 ***
SentimentHE_withAI 2.988e+00 2.171e-01 13.765 < 2e-16 ***
LC_PC1_withAI   -7.948e-01 6.790e-02 -11.706 < 2e-16 ***
Soph_PC1_withAI  3.091e-01 2.695e-02 11.471 < 2e-16 ***
Var_PC1_withAI   8.913e-01 6.286e-02 14.179 < 2e-16 ***
Tone_withAI     -2.667e-02 2.261e-02 -1.180 0.238019
log_goal        6.966e-02 5.586e-03 12.470 < 2e-16 ***
duration        -1.871e-02 9.945e-04 -18.818 < 2e-16 ***
US_Project1     1.073e-01 1.857e-02 5.779 7.55e-09 ***
staff_pick      2.746e+00 2.994e-02 91.731 < 2e-16 ***
experience       5.950e-02 2.001e-03 29.736 < 2e-16 ***
competing_campaigns -9.202e-06 5.254e-05 -0.175 0.860968
google_trend    -1.742e-02 1.790e-03 -9.728 < 2e-16 ***
categorydesign   2.050e+00 4.229e-02 48.484 < 2e-16 ***
categorygames   1.594e+00 4.635e-02 34.385 < 2e-16 ***
categorytechnology 7.739e-01 4.362e-02 17.741 < 2e-16 ***
year2011        1.956e+00 8.814e-01 2.219 0.026472 *
year2012        2.595e+00 8.763e-01 2.961 0.003065 **
year2013        3.149e+00 8.782e-01 3.586 0.000336 ***
year2014        2.245e+00 8.797e-01 2.552 0.010718 *
year2015        2.147e+00 8.782e-01 2.445 0.014486 *
year2016        2.237e+00 8.747e-01 2.557 0.010548 *
year2017        2.498e+00 8.742e-01 2.857 0.004275 **
year2018        2.684e+00 8.734e-01 3.074 0.002115 **
year2019        2.838e+00 8.733e-01 3.250 0.001155 **
year2020        2.888e+00 8.730e-01 3.308 0.000941 ***
year2021        3.117e+00 8.730e-01 3.571 0.000356 ***
year2022        2.889e+00 8.728e-01 3.310 0.000932 ***
year2023        2.476e+00 8.716e-01 2.841 0.004495 **
year2024        2.641e+00 8.724e-01 3.028 0.002463 **
year2025        2.320e+00 8.730e-01 2.658 0.007872 **

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.611 on 85566 degrees of freedom

(151 observations deleted due to missingness)

Multiple R-squared: 0.3394, Adjusted R-squared: 0.3391

F-statistic: 1293 on 34 and 85566 DF, p-value: < 2.2e-16

>

Effets marginaux (logit)

```
> marg_h3 <- margins(reg_h3_success, variables=vars_textuelles)
```

```
> print(summary(marg_h3)[, c("factor", "AME", "SE", "p")])
```

```

      factor  AME  SE  p
AboutWords 0.0001 0.0000 0.0000
Flesch_withAI -0.0003 0.0001 0.0007

```

```

LC_PC1_withAI -0.0503 0.0107 0.0000
RiskWords -0.0002 0.0000 0.0000
SentimentHE_withAI 0.4471 0.0369 0.0000
SentimentLM_withAI 0.3171 0.0311 0.0000
Soph_PC1_withAI 0.0237 0.0042 0.0000
Tone_withAI -0.0179 0.0035 0.0000
Var_PC1_withAI 0.0638 0.0099 0.0000
>
> # Effets économiques OLS
> for(var in vars_textuelles) {
+   coef_var <- coef(reg_h3_pledged)[var]
+   p_var <- summary(reg_h3_pledged)$coefficients[var, "Pr(>|t|)"]
+   sig <- ifelse(p_var < 0.01, "***", ifelse(p_var < 0.05, "***", ifelse(p_var < 0.1, "**", "ns")))
+   effet <- (exp(coef_var) - 1) * 100
+   cat(var, ": coef =", round(coef_var, 4), "| effet =", round(effet, 2), "% | p =", round(p_var, 3),
sig, "\n")
+ }
AboutWords : coef = 0.0011 | effet = 0.11 % | p = 0 ***
RiskWords : coef = -4e-04 | effet = -0.04 % | p = 0.002 ***
Flesch_withAI : coef = -0.0035 | effet = -0.34 % | p = 0 ***
SentimentLM_withAI : coef = 1.6251 | effet = 407.9 % | p = 0 ***
SentimentHE_withAI : coef = 2.9885 | effet = 1885.58 % | p = 0 ***
LC_PC1_withAI : coef = -0.7948 | effet = -54.83 % | p = 0 ***
Soph_PC1_withAI : coef = 0.3091 | effet = 36.22 % | p = 0 ***
Var_PC1_withAI : coef = 0.8913 | effet = 143.83 % | p = 0 ***
Tone_withAI : coef = -0.0267 | effet = -2.63 % | p = 0.238 ns
>

```

9.3. R Code

```

rm(list=ls())

library(stargazer)
library(ggplot2)
library(lmtest)
library(sandwich)
library(modelsummary)
library(margins)
install.packages("pandoc")
library(pandoc)
if(!dir.exists("graphiques")) dir.create("graphiques")

# Charger la base de données
setwd("/Users/lucasdaemen/Desktop/Mémoire/DATA")
data_crowd_funding <- readRDS(file="merged_data_all_5.rds")

# Voir à quoi ça ressemble
head(data_crowd_funding, 3)
options(max.print = 10000)

```

```

head(data_crowd_funding, 5)
length(unique(data_crowd_funding$country))
summary(data_crowd_funding$usd_goal)

# Hypothèse 1 - AI on Crowdfunding succes

# Variable dépendante 1 : Success (binaire)
data_crowd_funding$Success <- as.numeric(data_crowd_funding$Success)
table(data_crowd_funding$Success, useNA="always")

# Variable dépendante 2 : log(usd_pledged_w + 1)
summary(data_crowd_funding$usd_pledged_w)
sum(data_crowd_funding$usd_pledged_w <= 0, na.rm=TRUE)
data_crowd_funding$log_pledged <- log(data_crowd_funding$usd_pledged_w + 1)

# Variable IA — AVANT normalisation
summary(data_crowd_funding$AI_use_per)
# Normaliser (0-100 → 0-1)
data_crowd_funding$AI_use_per <- data_crowd_funding$AI_use_per / 100
# Vérifier APRÈS normalisation
summary(data_crowd_funding$AI_use_per)

# Contrôles
data_crowd_funding$log_goal <- log(data_crowd_funding$usd_goal + 1)
data_crowd_funding$category <- as.factor(data_crowd_funding$category)
data_crowd_funding$year <- as.factor(data_crowd_funding$year)
data_crowd_funding$US_Project <- as.factor(data_crowd_funding$US_Project)

contrôles <- c("log_goal", "duration", "US_Project", "staff_pick",
              "experience", "competing_campaigns", "google_trend",
              "category", "year")
formule_contrôles <- paste(contrôles, collapse=" + ")
na_controls <- sapply(contrôles, function(v) sum(is.na(data_crowd_funding[[v]])))
print(na_controls)

# Régressions
formule_1a <- as.formula(paste("Success ~ AI_use_per +", formule_contrôles))
reg_1a <- glm(formule_1a, data=data_crowd_funding, family=binomial(link="logit"))

formule_2a <- as.formula(paste("log_pledged ~ AI_use_per +", formule_contrôles))
reg_2a <- lm(formule_2a, data=data_crowd_funding)

summary(reg_1a)
summary(reg_2a)

# Tableaux

```

```

modelsummary(
  list("AI_use_per" = reg_1a),
  title = "Impact AI on success",
  coef_map = c("AI_use_per" = "Succes"),
  coef_omit =
"category|year|US_Project|log_goal|duration|staff_pick|experience|competing_campaigns|google_trend",
  stars = c('*' = 0.1, '**' = 0.05, '***' = 0.01),
  gof_map = c("nobs", "aic"),
  notes = "Control variables included."
)

```

```

modelsummary(
  list("AI_use_per" = reg_2a),
  title = "Impact AI on Amount Raised",
  coef_map = c("AI_use_per" = "Amount raised"),
  coef_omit =
"category|year|US_Project|log_goal|duration|staff_pick|experience|competing_campaigns|google_trend",
  stars = c('*' = 0.1, '**' = 0.05, '***' = 0.01),
  gof_map = c("nobs", "r.squared", "adj.r.squared"),
  notes = "Control variables included."
)

```

```

# Effets marginaux (Succes)
marg_1a <- margins(reg_1a, variables="AI_use_per")
cat("AI_use_per → P(Success) :", round(summary(marg_1a)$AME, 4),
    "(p =", round(summary(marg_1a)$p, 3), ")\n")

# Effets économiques (Amount Raised)
coef_2a <- coef(reg_2a)["AI_use_per"]
p_2a <- summary(reg_2a)$coefficients["AI_use_per", "Pr(>|t|)"]
cat("AI_use_per : coef =", round(coef_2a, 4), "| p =", round(p_2a, 3), "\n")
cat(" → +10 pp d'IA :", round((exp(coef_2a * 0.1) - 1) * 100, 2), "% sur usd_pledged_w\n\n")

```

Hypothèse 2

```

# Variables textuelles (withAI)
vars_textuelles <- c("AboutWords", "RiskWords",
  "Flesch_withAI",
  "SentimentLM_withAI", "SentimentHE_withAI",
  "LC_PC1_withAI", "Soph_PC1_withAI", "Var_PC1_withAI",
  "Tone_withAI")

# Vérifier NA
na_text <- sapply(vars_textuelles, function(v) sum(is.na(data_crowd_funding[[v]])))
print(na_text)

```

```

# Vérifier distributions
summary(data_crowd_funding[, vars_textuelles])

# Régressions
# Une régression OLS par variable textuelle (variable dépendante)

reg_h2_list <- lapply(vars_textuelles, function(var) {
  formule <- as.formula(paste(var, "~ AI_use_per +", formule_controles))
  lm(formule, data=data_crowd_funding)
})
names(reg_h2_list) <- vars_textuelles

# Vérifier les résultats
lapply(reg_h2_list, summary)

# Tableau
modelssummary(
  reg_h2_list,
  output = "graphiques/tableau_h2.html",
  title = "Impact AI on all Textual Variables",
  coef_map = c("AI_use_per" = "Utilisation IA (proportion)"),
  coef_omit =
"category|year|US_Project|log_goal|duration|staff_pick|experience|competing_campaigns|google_trend|Intercept",
  stars = c('*' = 0.1, '**' = 0.05, '***' = 0.01),
  gof_map = c("nobs", "r.squared"),
  notes = "Control variables included."
)

# Résumé des coefficients et significativité
for(var in vars_textuelles) {
  coef_ai <- coef(reg_h2_list[[var]])["AI_use_per"]
  p_ai <- summary(reg_h2_list[[var]])$coefficients["AI_use_per", "Pr(>|t|)"]
  sig <- ifelse(p_ai < 0.01, "***", ifelse(p_ai < 0.05, "**", ifelse(p_ai < 0.1, "*", "ns")))
  cat(var, ": coef =", round(coef_ai, 4), "| p =", round(p_ai, 3), sig, "\n")
}

# Hypothèse 3

# Formules
formule_texte <- paste(vars_textuelles, collapse=" + ")

formule_h3_success <- as.formula(paste("Success ~", formule_texte, "+", formule_controles))
formule_h3_pledged <- as.formula(paste("log_pledged ~", formule_texte, "+",
formule_controles))

```

```

# Régressions
# Logit : variables textuelles → Success
reg_h3_success <- glm(formule_h3_success, data=data_crowd_funding,
                      family=binomial(link="logit"))

# OLS : variables textuelles → log(pledged)
reg_h3_pledged <- lm(formule_h3_pledged, data=data_crowd_funding)

summary(reg_h3_success)
summary(reg_h3_pledged)

```

Tableaux

```

# Panel A : Success
modelsummary(
  output = "graphiques/tableau_h3a.html",
  list("Success (Logit)" = reg_h3_success),
  title = "Impact of textual variables on success",
  coef_map = c(
    "AboutWords"      = "AboutWords",
    "RiskWords"       = "RiskWords",
    "Flesch_withAI"   = "Flesch_withAI",
    "SentimentLM_withAI" = "SentimentLM_withAI",
    "SentimentHE_withAI" = "SentimentHE_withAI",
    "LC_PC1_withAI"   = "LC_PC1_withAI",
    "Soph_PC1_withAI" = "Soph_PC1_withAI",
    "Var_PC1_withAI"  = "Var_PC1_withAI",
    "Tone_withAI"     = "Tone_withAI"
  ),
  coef_omit =
"category|year|US_Project|log_goal|duration|staff_pick|experience|competing_campaigns|google_trend|Intercept",
  stars = c('*' = 0.1, '**' = 0.05, '***' = 0.01),
  gof_map = c("nobs", "aic"),
  notes = "Control variables included."
)

```

```

# Panel B : Montant levé
modelsummary(
  output = "graphiques/tableau_h3b.html",
  list("log(pledged) (OLS)" = reg_h3_pledged),
  title = "Impact of textual variables on amount raised",
  coef_map = c(
    "AboutWords"      = "AboutWords",
    "RiskWords"       = "RiskWords",
    "Flesch_withAI"   = "Flesch_withAI",
    "SentimentLM_withAI" = "SentimentLM_withAI",
    "SentimentHE_withAI" = "SentimentHE_withAI",

```

```

"LC_PC1_withAI"    = "LC_PC1_withAI",
"Soph_PC1_withAI"  = "Soph_PC1_withAI",
"Var_PC1_withAI"   = "Var_PC1_withAI",
"Tone_withAI"      = "Tone_withAI"
),
coef_omit
"category|year|US_Project|log_goal|duration|staff_pick|experience|competing_campaigns|google_trend|Intercept",
stars = c('*' = 0.1, '**' = 0.05, '***' = 0.01),
gof_map = c("nobs", "r.squared", "adj.r.squared"),
notes = "Control variables included."
)

# Effets marginaux (logit)
marg_h3 <- margins(reg_h3_success, variables=vars_textuelles)
print(summary(marg_h3)[, c("factor", "AME", "SE", "p")])

# Effets économiques OLS
for(var in vars_textuelles) {
  coef_var <- coef(reg_h3_pledged)[var]
  p_var <- summary(reg_h3_pledged)$coefficients[var, "Pr(>|t|)"]
  sig <- ifelse(p_var < 0.01, "***", ifelse(p_var < 0.05, "**", ifelse(p_var < 0.1, "*", "ns")))
  effet <- (exp(coef_var) - 1) * 100
  cat(var, ": coef =", round(coef_var, 4), "| effet =", round(effet, 2), "% | p =", round(p_var, 3),
  sig, "\n")
}

#####

```

Abstract :

This study aims to better understand the impact of artificial intelligence (AI) on crowdfunding campaign performance using a database of 85,752 Kickstarter campaigns across 123 countries. This paper tests if AI use affects campaign outcomes directly. It also tries to find out if AI impacts the textual characteristics of campaign descriptions and if those textual characteristics influence the probability of success and the total amount raised of the funding campaign.

Results show that higher AI use is significantly and negatively associated with both campaign success and amount raised. This study also finds, among other findings, that content generated by AI reduces readability by increasing lexical complexity and produces a more positive tone. The tested changes in textual characteristics themselves have a significant impact on campaign performance. Together, the findings suggest that using AI as an autonomous writer in crowdfunding descriptions has a bad impact on campaign performance. This effect operates through the ways AI changes the writing style of campaign descriptions.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management

Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique

www.uclouvain.be/lsm