



**Universidade do Estado do Rio de Janeiro**

Centro de Educação e Humanidades

Faculdade de Comunicação Social

Leonardo Mancini de Almeida

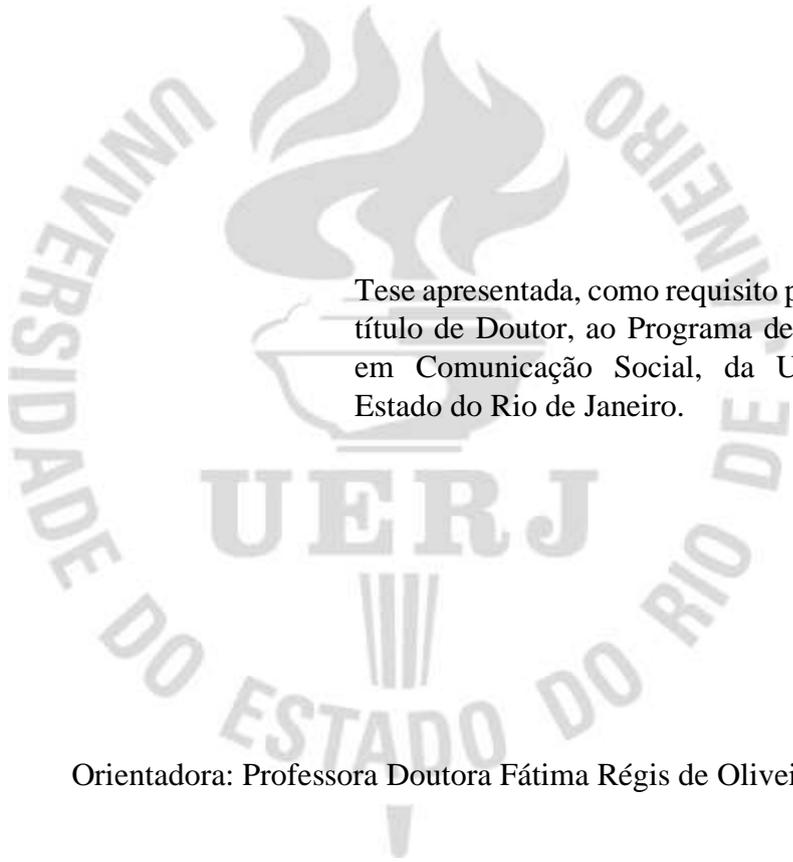
***Crowdfunding no Brasil: Determinantes do apoio em projetos do Catarse***

Rio de Janeiro

2020

Leonardo Mancini de Almeida

***Crowdfunding* no Brasil: Determinantes do apoio em projetos do Catarse**



Tese apresentada, como requisito para obtenção do título de Doutor, ao Programa de Pós-Graduação em Comunicação Social, da Universidade do Estado do Rio de Janeiro.

Orientadora: Professora Doutora Fátima Régis de Oliveira

Rio de Janeiro

2020

CATALOGAÇÃO NA FONTE  
UERJ / REDE SIRIUS / BIBLIOTECA CEH/A

A447 Almeida, Leonardo Mancini de.  
Crowdfunding no Brasil: Determinantes do apoio em projetos do Catarse /  
Leonardo Mancini de Almeida. – 2020.  
127 f.

Orientador: Fátima Régis de Oliveira.  
Tese (Doutorado) – Universidade do Estado do Rio de Janeiro. Faculdade de  
Comunicação Social.

1. Comunicação – Teses. 2. Financiamento coletivo– Teses. 3. Internet –  
Teses. I. Oliveira, Fátima Régis de. II. Universidade do Estado do Rio de Janeiro.  
Faculdade de Comunicação Social. III. Título.

es CDU 316.77

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese,  
desde que citada a fonte.

\_\_\_\_\_  
Assinatura

\_\_\_\_\_  
Data

Leonardo Mancini de Almeida

***Crowdfunding* no Brasil: Determinantes do apoio em projetos do Catarse**

Tese apresentada, como requisito para obtenção do título de Doutor, ao Programa de Pós-Graduação em Comunicação Social, da Universidade do Estado do Rio de Janeiro.

Aprovada em 18 de dezembro de 2020.

Banca Examinadora:

---

Professora Doutora Fátima Régis de Oliveira (Orientadora)  
Universidade do Estado do Rio de Janeiro - UERJ

---

Professor Doutor Carlos Alberto Messeder Pereira  
Universidade Federal do Estado do Rio de Janeiro - UFRJ

---

Professor Doutor Diogo Tavares Robaina  
Escola Superior de Propaganda e Marketing - ESPM Rio

---

Professor Doutor Thiago Moreira da Silva  
Universidade do Estado do Rio de Janeiro - UERJ

---

Professora Doutora Veranise Jacubowski Correia Dubeux  
Escola Superior de Propaganda e Marketing - ESPM Rio

Rio de Janeiro

2020

## DEDICATÓRIA

Dedico esse trabalho a: Letícia, Rafael, Beatriz, Gina, João, Camila, Thays, Cândida, Ângela, Arnaldo (*in memorian*), Zilá (*in memorian*), Joaõ (*in memorian*), Ieda (*in memorian*) todos coautores, de alguma forma.

Amo vocês.

## AGRADECIMENTOS

Realizar pesquisa acadêmica é um desafio imenso no Brasil de hoje. Mais ainda tendo dois filhos (que nasceram ao longo do desenvolvimento deste projeto), trabalhando em tempo integral, atravessando a maior pandemia do século e um momento político tenebroso.

Desenvolver esta pesquisa não seria possível sem:

o apoio da minha esposa, não só minha companheira de vida, mas minha maior incentivadora e maior crítica;

a orientação generosa da Professora Fátima Régis, que não só me mostrou o caminho, mas soube puxar as orelhas e aliviar o chicote com precisão;

meus filhos, Bia e Rafa, que tornam o mundo mais agitado e complexo, mas também um lugar mais bonito e feliz, e nos dão a certeza de que o amanhã será muito melhor;

o incentivo do Professor Carlos Alberto Messeder Pereira, grande responsável por me trazer para a academia e mentor por tanto tempo;

as orientações do Professor Thiago Moreira, que facilitou imensamente essa caminhada maluca de um comunicador pela estatística;

os amigos André Korenblum, Bruno Seixas, Fernando Hermont, Pedro Curi, Pedro Drevon, e tantos outros que trazem um pouco de equilíbrio ao sanatório geral;

os professores, funcionários e colegas de turma do PPGCOM–Uerj e do Cibercog, que me acompanharam e apoiaram ao longo do programa.

A todos o meu mais sincero agradecimento.

There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened

— *Douglas Adams, The Hitchhiker's Guide to the Galaxy*

## RESUMO

MANCINI DE ALMEIDA, L. M. A. *Crowdfunding no Brasil: Determinantes do apoio em projetos do Catarse*. 2020. 127 f. Tese (Doutorado em Comunicação Social) – Faculdade de Comunicação Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

O financiamento coletivo, ou *crowdfunding*, é uma designação nova para uma prática ancestral, a cotização coletiva para a viabilização de um projeto. Como atividade contemporânea, passa a ser organizada por plataformas digitais e permite que desconhecidos se organizem em torno de projetos que não teriam incentivos para ser realizados em outras circunstâncias. Buscamos identificar quais são os determinantes dos apoios em projetos publicados no Catarse, principal plataforma do tipo no Brasil. Para realizar esse estudo, analisamos a literatura existente a respeito do contexto tecnológico e social no qual surgem iniciativas no campo da Economia do Compartilhamento e da *Gig economy*, e dos estudos anteriores sobre *crowdfunding* para delimitar o escopo teórico. Posteriormente, coletamos informações a respeito de aproximadamente 16 mil projetos e 950 mil usuários cadastrados através de *web scraping* e analisamos os dados utilizando regressão binomial negativa em multinível. Como conclusão, percebemos que realizadores que planejam melhor suas estratégias de comunicação atraem mais apoios. Se mostraram variáveis significativas o número de atualizações, a utilização de outras plataformas de comunicação e o tamanho das descrições. Algumas características dos projetos também se mostraram significativas, como o número de recompensas, a duração da campanha e a modalidade no qual estão inseridas. Da mesma forma, algumas características dos realizadores se mostraram relevantes, como uma rede pessoal de contatos e experiências anteriores. Por fim, percebemos efeitos de geografia no número de apoios, como a densidade populacional e a concentração de renda dos municípios de onde se originam os projetos. Como contribuição ao campo da comunicação, este trabalho ambiciona avançar a compreensão a respeito das novas configurações sociais surgidas com a internet e propor uma reflexão a respeito de como as pesquisas em comunicação podem se apropriar de ferramentas de outras áreas para aprimorar seus próprios métodos.

Palavras-chave: Financiamento coletivo. Internet. Ciência de dados. Comunicação.

## ABSTRACT

MANCINI DE ALMEIDA, L. M. A. *Crowdfunding in Brazil: the determinants of backer's funding in Catarse projects*. 2020. 127 f. Tese (Doutorado em Comunicação Social) – Faculdade de Comunicação Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

Crowdfunding is a new designation for an ancestral practice, the collective raising of small amounts of money in order to make a project viable. Today, it is organized by digital platforms, which permit complete strangers to participate in projects that would otherwise have few if any chances of success. The purpose of this study is to identify the determinants of support for projects published in Catarse, the main crowdfunding platform in Brazil. In addition to analyzing the extant social and technological literature on the so-called Sharing and Gig Economy, this study analyzes approximately 16 thousand projects and 950 thousand registered users through web scraping. Employing a negative binomial multilevel regression, we find that entrepreneurs who plan their communication strategies better attract more support, with the number of updates, the use of alternative communication platforms and the size of project descriptions all being statistically significant predictors of backing. Several characteristics of the projects are also significant, such as the number of rewards, the duration of the campaign, and the modality in which they are registered. Likewise, several characteristics of the entrepreneurs are relevant, including personal networks and previous experiences with crowdfunding. Finally, we found that geography plays an important role, as population density and income concentration in cities where projects originate have influence on the number of backers. As a contribution to the field of communications, this work aims to advance the understanding of new social configurations that have emerged in the wake of the internet and to reflect on how research in communications can use tools from other fields to improve their own methods.

Keywords: Crowdfunding. Internet. Data science. Communications.

## LISTA DE FIGURAS

Figura 1 - Distribuição normal e de potência.....	23
Figura 2 - A Cauda Longa.....	30
Figura 3 - <i>Hotsite</i> de campanha de financiamento do Catarse.....	46
Figura 4 - Interação através dos comentários dos projetos de financiamento coletivo.....	47
Figura 5 - Tela de captura de informações do import.io.....	79
Figura 6 - Dados de projetos do Catarse, estruturados pelo import.io.....	79
Figura 7 - Interface RStudio durante o processo de raspagem.....	80
Figura 8 - Projetos publicados por usuários.....	87
Figura 9 - Contribuições realizadas por usuários.....	87
Figura 10 - Distribuição da variável apoiadores.....	88
Figura 11 - Distribuição da variável atualizações.....	89
Figura 12 - Distribuição do número de perfis em sites de redes sociais.....	90
Figura 13 - Distribuição da variável tamanho das descrições.....	90
Figura 14 - Distribuição da variável número de recompensas, considerando cada categoria.....	92
Figura 15 - Localização e concentração geográfica dos projetos.....	94
Figura 16 - Principais categorias por estado.....	96
Figura 17 - Preditores do número de apoiadores.....	98
Figura 18 - Preditores do número de apoiadores - efeitos aleatórios.....	98
Figura 19 - Efeitos preditos da variação do número de recompensas.....	99
Figura 20 - Efeitos preditos da variação do número de atualizações.....	100
Figura 21 - Efeitos preditos da variação da duração dos projetos.....	101
Figura 22 - Efeitos preditos do número de palavras na descrição.....	102
Figura 23 - Efeitos preditos do número de seguidores.....	103
Figura 24 - Efeitos preditos do número de usuários que os realizadores seguem.....	104
Figura 25 - Efeitos preditos do número de contribuições feitas por realizadores.....	104
Figura 26 - Efeitos preditos do número de projetos já publicados por realizadores.....	105
Figura 27 - Efeitos preditos da concentração de renda.....	105
Figura 28 - Efeitos preditos da densidade populacional.....	106
Figura 29 - Normalidade dos resíduos.....	127
Figura 30 - QQ plot.....	127
Figura 31 - Normalidade dos efeitos aleatórios.....	127

## LISTA DE TABELAS

Tabela 1 -	Número de plataformas de <i>crowdfunding</i> por país .....	38
Tabela 2 -	Número de projetos por categoria .....	39
Tabela 3 -	Número de projetos bem-sucedidos por categoria .....	40
Tabela 4 -	Categorias com mais sucesso .....	41
Tabela 5 -	Graus de liberdade e índices AIC dos modelos de contagem .....	81
Tabela 6 -	Diagnóstico de ajuste dos modelos multinível .....	83
Tabela 7 -	Relação de dados faltantes .....	83
Tabela 8 -	Resultados da imputação múltipla.....	83
Tabela 9 -	Relação de apoios por categoria .....	88
Tabela 10 -	Características dos projetos .....	91
Tabela 11 -	Modalidade dos projetos por categoria .....	92
Tabela 12 -	Características dos realizadores.....	93
Tabela 13 -	Projetos por cidade .....	94
Tabela 14 -	Projetos por Estado.....	95
Tabela 15 -	Sumário Estatístico.....	122
Tabela 16 -	Matriz de correlações - Coeficiente de Pearson .....	124
Tabela 17 -	Matriz de correlações - Coeficiente de Spearman.....	124
Tabela 19 -	Variância dos efeitos aleatórios - Modelo 1 .....	126
Tabela 20 -	Variância dos efeitos aleatórios - Modelo 2 (ATUALIZAR.....	126
Tabela 21 -	Variância dos efeitos aleatórios - Modelo 3 (ATUALIZAR.....	126

## SUMÁRIO

	<b>INTRODUÇÃO</b> .....	12
1	<b>INTERNET E ECONOMIA</b> .....	17
1.1	<b>Internet e sociedade</b> .....	17
1.1	<b>As redes</b> .....	22
1.2	<b>Internet e economia</b> .....	24
2	<b>O FINANCIAMENTO COLETIVO</b> .....	34
2.1	<b>Definições e histórico</b> .....	34
2.2	<b>A relevância do financiamento coletivo</b> .....	36
2.3	<b>O mercado de <i>crowdfunding</i> hoje</b> .....	37
2.4	<b>Modalidades de financiamento coletivo</b> .....	41
2.5	<b>Processo de criação de um projeto e seus elementos</b> .....	43
2.6	<b>Formas de comunicação entre empreendedores e investidores</b> .....	45
2.7	<b>Incentivos e oportunidades do financiamento coletivo</b> .....	48
2.8	<b>Riscos do financiamento coletivo</b> .....	50
2.9	<b>Sinais de qualidade e confiança</b> .....	54
2.10	<b>Sinais de qualidade e financiamento coletivo</b> .....	57
3	<b>A PESQUISA EM COMUNICAÇÃO NA ERA DO BIG DATA</b> .....	62
3.1	<b>A pesquisa na Era do <i>Big Data</i></b> .....	62
3.2	<b>As Ciências Sociais Computacionais (CSC)</b> .....	64
3.3	<b>CSC e pesquisa quantitativa</b> .....	66
3.4	<b>Pesquisas Quantitativas em Comunicação</b> .....	68
3.5	<b>Extração automatizada de dados e implicações éticas</b> .....	70
4	<b>METODOLOGIA E ANÁLISE DE DADOS</b> .....	78
4.1	<b>Problema de pesquisa e escopo do estudo</b> .....	78
4.2	<b>Coleta de dados</b> .....	78
4.3	<b>Modelos binomiais negativos e modelos multinível</b> .....	81
4.4	<b>Tratamento dos dados e descrição das variáveis</b> .....	83
4.5	<b>Definição de variáveis e desenvolvimento das hipóteses</b> .....	84
4.6	<b>Descrição das variáveis</b> .....	86
4.6.1	<u>Características dos usuários da plataforma</u> .....	86
4.6.2	<u>Estratégias de divulgação</u> .....	89

4.6.3	<u>Características dos projetos</u> .....	91
4.6.4	<u>Características dos realizadores</u> .....	93
4.6.5	<u>Efeitos de geografia</u> .....	93
4.7	<b>Determinantes do apoio em projetos do Catarse</b> .....	96
	<b>CONCLUSÃO</b> .....	107
	<b>REFERÊNCIAS</b> .....	110
	<b>GLOSSÁRIO</b> .....	118
	<b>APÊNDICE A</b> – Diagnósticos da Imputação múltipla .....	120
	<b>APÊNDICE B</b> – Diagnósticos dos Modelos de Contagem.....	121
	<b>APÊNDICE C</b> – Sumário estatístico.....	122
	<b>APÊNDICE D</b> – Matrizes de correlação.....	124
	<b>APÊNDICE E</b> – Coeficientes dos modelos estimados .....	125
	<b>APÊNDICE F</b> – Coeficientes dos modelos estimados - efeitos aleatórios.....	126
	<b>APÊNDICE G</b> – Diagnósticos dos modelos estimados .....	127

## INTRODUÇÃO

Em 2016, Maycom Passos de Melo tornou-se uma pequena celebridade no mundo dos cafés. Junto com Juca Esmanhoto, dono da Rause Café, um estabelecimento famoso de Curitiba, Maycom lançou o projeto de desenvolvimento da Cafeteira Aram, uma máquina de espresso que não utiliza energia elétrica, é portátil, aparenta ser um objeto de decoração e promete extrair “o melhor do café”. Sua preocupação era produzir um aparelho que conseguisse fazer um café de qualidade, mas que fosse mais sustentável e que demandasse pouca manutenção (CATARSE, 2017a).

Depois de elaborar um protótipo, Maycom e Juca recorreram ao Catarse, o principal site de *crowdfunding* no Brasil, para financiar sua produção. Nesta plataforma, qualquer um pode contribuir financeiramente para que um projeto de seu interesse se concretize. Em apenas dois dias, a dupla conseguiu os R\$ 35 mil que imaginavam ser necessários para construir a cafeteira. Ao final, o projeto recebeu R\$ 253,3 mil, mais de 7 vezes a sua meta inicial, e foi apoiado por 334 pessoas que, em sua maioria, nunca encontraram a dupla ou viram a cafeteira de perto.

No mesmo ano, Caroline Pasternak e Mayara Trevisan se propuseram a expandir a pesquisa que realizavam como conclusão de curso de graduação em jornalismo na FIAM-FAAM, em São Paulo, a respeito das ocupações nas escolas municipais e estaduais, que estavam ocorrendo em todo o estado naquele ano. Seu projeto, chamado “Quem mexeu na minha merenda?”, buscava recursos para a produção de um *webdocumentário*, “investigativo e independente”, a respeito das histórias dos estudantes que participavam das ocupações, “com o objetivo de mostrar ao grande público que lutas como estas são imensamente necessárias e acontecerão cada vez mais” (CATARSE, 2017b).

O projeto foi inscrito na plataforma em 2 de agosto de 2017 e pôde receber doações até o dia 31 de outubro. Neste período, somente uma pessoa o apoiou, com R\$ 10, um valor muito abaixo dos R\$ 2 mil desejados pelas estudantes, o que inviabilizou sua execução.

O financiamento coletivo, nome contemporâneo para a tradicional vaquinha, surge, em sua forma atualizada, em meio a um conjunto de novas práticas típicas da transição, que vivemos, para um sistema econômico baseado no uso de computadores e de tecnologias digitais. Alguém que considere ter uma boa ideia, mas não tem acesso a recursos financeiros para sua execução, pode, ao invés de pedir aos amigos ou vender seu carro, divulgar o projeto em um site na internet e, caso convença um número grande de desconhecidos da importância, da inovação ou da viabilidade de sua criação, começa a receber doações em troca de acesso facilitado, ou antecipado, a produtos e serviços, ou somente de um agradecimento.

A promessa maior do financiamento coletivo contemporâneo é servir como alternativa em situações nas quais os incentivos econômicos tradicionais não estão presentes e, portanto, não interessaria a investidores institucionais, como bancos, *venture capitalists* e fundos de investimento. Nenhuma dessas organizações financiaria uma biblioteca comunitária, por exemplo, com alguma expectativa de retorno financeiro, independentemente da necessidade social e da importância dada por uma determinada comunidade à tal biblioteca.

Nosso objetivo, desta forma, é identificar quais são os determinantes dos apoios nos projetos de financiamento coletivo do Catarse. Nossa hipótese é que elementos constitutivos dos projetos e das campanhas, além de características dos realizadores, sinalizam aos potenciais apoiadores as qualidades, a relevância ou o nível de preparo dos proponentes, o que acaba por atrair os membros da comunidade.

A definição do Catarse como objeto de estudo tem como justificativa o fato de ser a primeira plataforma brasileira do tipo e a que, hoje, atrai mais atenção e mais projetos. Além disso, é um site desenvolvido com código aberto, compartilhado no GitHub, e com uma estrutura com poucas restrições, o que permite a automatização da extração e da manipulação de seus dados.

Apesar das plataformas de *crowdfunding* estarem crescendo em importância, os estudos a respeito no Brasil ainda são poucos, quase todos preliminares, descritivos e incompletos, o que não é surpreendente, considerando este ser um tema relativamente recente. Após uma revisão bibliográfica, que será detalhada ao longo dos capítulos 2 e 4, consideramos que a principal lacuna existente refere-se justamente às motivações dos apoiadores. Há autores que se concentraram na dinâmica das campanhas com foco em seu sucesso ou seu fracasso, mas não nos pareciam claros que fatores mobilizam os apoios, ou seja, que elementos intrínsecos aos projetos podem ser decisivos no momento em que alguém escolhe ajudar, ou não, um projeto desenvolvido, na maior parte das vezes, por um desconhecido.

Mesmo fora do Brasil, Mollick (2014), que realizou um dos trabalhos mais relevantes sobre o tema, defende que ainda não está claro como os realizadores se preparam para apresentar suas ideias, como desenham suas estratégias, ou mesmo se há uma busca, por parte dos apoiadores, de qualidades intrínsecas aos projetos de *crowdfunding*, algo que o autor sugere merecer mais atenção por parte da academia.

Desta forma, o estudo conduzido tem caráter exploratório e empírico, considerando que a literatura a respeito da prática no Brasil ainda é incipiente e que o fenômeno ainda está em curso.

Do ponto de vista teórico, essa pesquisa tem a ambição de auxiliar na compreensão a respeito das novas configurações sociais surgidas com a internet. Se, como defende Sodré (2015), a comunicação é um processo de vinculação social através dos quais os sujeitos são construídos e os modos de vida são organizados, plataformas como o Catarse nos parecem de especial interesse já que servem como exemplo de como as relações entre indivíduos, especialmente os desconhecidos, se reorganizaram com a mediação da tecnologia.

Do ponto de vista metodológico, essa pesquisa pretende contribuir com a reflexão a respeito de como o campo da comunicação pode se apropriar de procedimentos e ferramentas de outras áreas do conhecimento, como as ciências sociais computacionais e a estatística. O estudo que empreendemos demandou a utilização de ferramentas e conceitos que, até recentemente, eram vistos como pertencentes exclusivamente a outros campos de conhecimento. Para Barabási e Frangos (2002, n.p.), a ciência do século XX se esforçou em observar a natureza em seus menores detalhes, quebrando-as em pedaços tão pequenos que se tornou quase impossível, hoje, juntar novamente o quebra-cabeças. A percepção de que o mundo, a natureza, as relações sociais, são parte de sistemas complexos, “cujos componentes podem se encaixar de tantas maneiras diferentes que nos tomaria bilhões de anos para tentá-las todas” (BARABÁSI; FRANGOS, 2002), nos obriga a olhar para esse objeto de estudo de forma interdisciplinar, aberta. Como defende Regis (2006), é preciso, para dar conta da complexidade do real, que as pesquisas contemporâneas rompam com as barreiras epistemológicas modernas e coloquem “em comunicação ciências que se distinguem pelo método, mas têm em comum a investigação da complexidade do mundo”(p. 152).

No primeiro capítulo, antes de apresentar propriamente o financiamento coletivos, desenhamos o atual contexto sócio-econômico-tecnológico, que permite o surgimento de negócios dessa natureza. Estamos vivendo um momento de transição, em direção a uma economia e a uma sociedade nas quais os relacionamentos são cada vez mais imbricados com as tecnologias digitais e, com isso, experimentamos novas formas de sociabilidade e de compartilhamento da vida social, das quais o financiamento coletivo faz parte. Nesta análise, partiremos dos estudos iniciais de Lévy (1999) e Castells (2003), alguns dos primeiros a analisar as dimensões sociais, culturais e econômicas das tecnologias digitais, e, em seguida, aprofundaremos essa reflexão com Simondon (2007), Latour (2008) e Callon (2008), colocando em perspectiva a relação entre cultura, seres humanos e objetos técnicos, além de analisar as teorias de redes com Castells (2004) e Barabási e Frangos (2002). Ao final, tentaremos situar com Anderson (2006 e 2012), Costa (2015 e 2017),

Shirky (2008 e 2010), Tapscott e Williams (2007), Fisk (2020) e Acevedo (2020) as novas práticas no campo da economia e dos negócios, mostrando as potencialidades e os limites das empresas da *Gig Economy*.

No segundo capítulo, situamos teoricamente o financiamento coletivo, analisando sua história, sua importância social e econômica e seu potencial em viabilizar empreendimentos que não seriam possíveis de outra forma. Para a economia, o valor do *crowdfunding* tem a ver com seu papel enquanto alternativa de financiamento para projetos que não oferecem os atrativos comuns a financiadores tradicionais. Para tanto, analisaremos autores como Agrawal et al (2011 e 2013), Brúntje e Gajda (2016), Carr (2013), Dresner (2014), Mollick (2014 e 2015), Gierczak et al. (2016) e Kuppaswamy e Bayus (2015), que se dedicaram a aprofundar o conhecimento a respeito das principais plataformas de *crowdfunding*, sobretudo nos Estados Unidos e Europa, que são as regiões onde a prática se desenvolveu com mais força. Como fonte no Brasil, analisaremos os estudos de Felipe (2017) e Martins e Martins (2016), principalmente.

No terceiro capítulo, apresentamos uma breve discussão a respeito dos efeitos da chamada “Era do Big Data” nas pesquisas em comunicação. Considerando que as relações sociais cada vez mais acontecem em plataformas digitais, há novas possibilidades e desafios, inclusive éticos, para as pesquisas em comunicação. Há, no entanto, certa resistência no campo a respeito das pesquisas de natureza quantitativa, o que pode ser um dificultador para que se adotem novos procedimentos típicos deste ambiente. Neste sentido, analisaremos os estudos, principalmente, de Cioffi-Revilla (2014), a respeito das Ciências Sociais Computacionais, de Pilny e Poole (2017) a respeito da “dataficação” das relações sociais, de Flick (2016) e de Alim (2014) sobre as limitações éticas de pesquisas que envolvam dados extraídos de sites de redes sociais.

No último capítulo, apresentamos as técnicas de coleta e de análise de dados utilizadas, além dos resultados de nossa pesquisa. Há um desafio metodológico singular a este tipo de pesquisa: o volume de informações. Na base que estruturamos, com dados coletados diretamente da plataforma, dispomos de informações a respeito de cerca de 16 mil projetos publicados no Catarse, entre 2011 e 2018, e outra, com informações a respeito de mais de 950 mil usuários. Ambas foram criadas utilizando-se uma técnica chamada *web scraping*, ou raspagem de dados, que permite extração e formatação de informações a partir de *websites*, de forma automatizada. Consideramos que essas bases representam a totalidade dos projetos e dos usuários cadastrados na plataforma de sua abertura até julho de 2018.

Após a base estruturada, utilizamos o algoritmo de processamento de linguagem natural do Google para analisar os textos das descrições e, assim, extrair duas variáveis de sentimento

textual. Todas as demais variáveis foram determinadas após a leitura da bibliografia correspondente.

Posteriormente, utilizamos inferência estatística para identificar os determinantes dos apoios, especificamente uma técnica de regressão chamada binomial negativa em

multinível, também conhecido por modelo hierárquico, considerando que nossa variável dependente, o número de apoiadores de um projeto, tem uma distribuição não linear, não negativa, e que as demais variáveis sofrem influência das categorias nas quais os projetos de financiamento estão cadastrados, ou seja, são dependentes dos contextos em que existem. Como resultado, conseguimos perceber os efeitos de cada categoria no comportamento das variáveis independentes, o que nos permite realizar previsões em um grau muito sofisticado.

Acreditamos que a escolha por esse desenho de pesquisa é inovadora não só pelo fato de serem raros os trabalhos eminentemente quantitativos no campo a comunicação, mas, sobretudo, porque em nenhum dos estudos analisados, mesmo de outras áreas, encontramos a aplicação de modelos mistos para explicar esse fenômeno, o que, nos parece, são os que melhor se adéquam aos dados existentes.

# 1 INTERNET E ECONOMIA

## 1.1 Internet e sociedade

Imagine que o projeto da cafeteira de Maycom de Melo e Juca Esmanhoto tivesse sido desenvolvido há 30 ou 40 anos, provavelmente, a dupla desenharia sua cafeteira em uma prancheta, com réguas e lápis. Se tivessem uma oficina à disposição, criariam um protótipo e, aí, teriam duas opções: ou venderiam sua invenção por um preço relativamente baixo a uma empresa já estabelecida, ou tentariam buscar financiamento em bancos ou com investidores profissionais para construir uma fábrica. Depois disso, buscariam jornais ou rádios para divulgar sua cafeteira entre entusiastas, e precisariam montar uma rede de distribuidores e revendedores para fazer chegar seus produtos a lojas físicas.

O processo de criação da Aram, no entanto, foi bem diferente. A dupla entrou em contato com especialistas em fóruns dedicados a café na internet e desenvolveu grande parte de sua criação de forma colaborativa. Para financiar o projeto, recorreram ao Catarse, um site de *crowdfunding* e, assim, conseguiram os recursos necessários para terminar o projeto e iniciar a produção (em poucos dias, receberam R\$ 253,3 mil, de 334 pessoas que nunca viram nem tiveram contato com seu protótipo). Para produção das peças, escolheram ateliês e fábricas brasileiras, mas poderiam ter encomendado todas as peças para fábricas asiáticas, o que teria barateado e acelerado a produção e a montagem.

O processo que descrevemos é típico do que vários autores estão chamando de nova revolução industrial, Era da Informação, Economia do Compartilhamento, *Gig Economy*, dentre várias designações, uma reorientação do sistema econômico possibilitada pela migração de uma economia baseada em tecnologias analógicas para outra, com foco no uso de computadores. Essa nova dinâmica permite não só a circulação facilitada de informação ou a criação de novos produtos e serviços, dentre elas as plataformas de financiamento coletivo, mas abre também espaço para novas formas de relacionamento entre indivíduos.

Em uma retrospectiva histórica, um dos primeiros autores a se dedicar ao estudo da internet foi Pierre Lévy (1999). Naquela que é, talvez, sua obra mais conhecida, o autor evidencia uma série de mudanças que vinham ocorrendo no ambiente político, econômico e cultural relacionadas à expansão da internet.

Em sua análise, Lévy (1999) defende que a metáfora do impacto das tecnologias é inadequada. Para ele, não há técnica externa à sociedade e à cultura. A tecnologia não é, em sua visão, uma bala que possa ser projetada na direção do ambiente, tornado alvo, e,

assim, transformá-lo. Ainda que possamos considerar que existam essas três categorias (tecnologia, sociedade e cultura) de maneira independente umas das outras, o que Lévy (1999) não acredita, a tecnologia será sempre produto de uma sociedade e de sua cultura. Do domínio do fogo ao lançamento de satélites, um dos elementos que constituem a humanidade é o desenvolvimento de ferramentas, elementos absolutamente indissociáveis da vida do homem.

Para Lévy (1999), melhor do que pensar em impactos é lembrar que toda técnica carrega consigo um projeto, uma nova correlação de forças com implicações sociais e culturais que são, em muitos casos, múltiplas e dúbias, resultantes dos esquemas imaginados pelos homens que a desenvolveram, com seus interesses específicos.

No caso da internet, Lévy (1999) considera que não é possível reconhecer um único projeto, já que sua expansão foi encorajada por estados, por grandes empresas de eletrônica e de software, por desenvolvedores e por usuários, cada um interessado em um aspecto: supremacia militar, autonomia, expansão de suas capacidades cognitivas, colaboração ou artes. É, portanto, o resultado de ambições heterogêneas, muitas vezes conflitantes, com implicações sociais e econômicas relevantes.

Para ele, não se pode falar, ainda, na determinação de uma cultura ou de uma sociedade por uma tecnologia, ou seja, que uma tecnologia seja a causa identificável de uma mudança cultural ou social. Para Lévy (1999), tecnologias condicionam, abrem possibilidades que não existiriam se elas não existissem. Nesse sentido, não são boas ou más; integram-se à vida social de formas distintas, de acordo com as pessoas e o contexto que a utilizam. Não significa que sejam neutras, no entanto; agem abrindo ou fechando portas, criando possibilidades, que são aproveitadas pelos atores sociais, que exploram seus limites e acabam criando novas técnicas e novos comportamentos, às vezes imprevisíveis e improváveis.

No caso da internet, é compreensível para Lévy (1999) que seja vista por muitos como um outro ameaçador dadas as mudanças acarretadas pelo seu uso. Quando se refere à existência de uma cibercultura, Lévy (1999) está apontando o surgimento de “práticas, de atitudes, de modos de pensamento e de valores que se desenvolvem juntamente com o crescimento do ciberespaço” (p. 13), que ele define, por sua vez, como o “meio de comunicação que surge da interconexão mundial de computadores” (p. 13).

Ainda em retrospectiva histórica, outro autor que se dedicou a analisar a internet foi Manuel Castells (2003), que considera seu surgimento um evento comparado com a invenção do motor elétrico na era Industrial por sua capacidade de fazer circular a força da informação por “todo o domínio da atividade humana” (p. 3). Da mesma forma que a energia tornou possível a criação de fábricas, a internet tornou-se o suporte, tecnológico e cultural, sobre o qual a rede, a forma de organização típica de nosso tempo, se estrutura.

Castells (2003) considera a internet o primeiro meio de comunicação “de muitos com muitos, num momento escolhido, em escala global” (p. 8). O autor lembra, da mesma forma que Lévy (1999), que todo sistema tecnológico é socialmente construído e, conseqüentemente, em sua raiz está um conjunto de valores, crenças e costumes compartilhados por seus criadores, que acaba moldando os sistemas e influenciando nas maneiras como esses sistemas podem ser apropriados pelos seus usuários.

Para Castells (2003), a cultura da internet é, pelo menos em sua origem, impregnada pelos valores daqueles que foram seus pioneiros: as forças armadas, as empresas de tecnologia e os jovens programadores e hackers americanos dos anos 60 a 80. Essa cultura, para ele, é estruturada em quatro camadas: a tecnomeritocrática, a *hacker*, a comunitária virtual e a empresarial que, juntas, formam um conjunto ideológico que fomenta a cooperação, o comunitarismo, a valorização da ciência e da tecnologia, o empreendedorismo e a liberdade, de criação e de ação, valores estes que não só estruturam a internet, mas condicionam seus usuários. Elementos desta cultura acabam por ter reflexos em diversas dimensões da vida cotidiana, em especial naquelas mais intensamente envolvidas com suas práticas.

Apesar de sua inegável contribuição, tanto Lévy (1999) quanto Castells (2003) são autores que parecem separar, em suas análises, o substrato técnico da inteligência humana, ou seja, as ferramentas digitais da ação e do pensamento do homem, algo que parece estranho nos estudos mais contemporâneos.

Lévy (1999), por exemplo, ainda que afirme vivermos em um sistema sociotécnico, no qual o humano é indissociável de seu ambiente material e dos artefatos que desenvolve, separa o que chama de inteligência coletiva dos dispositivos de comunicação digital. No esquema que desenhou, a internet é um imenso dispositivo de comunicação interativo, global e comunitário. Por sobre esse substrato técnico se desenvolve a inteligência coletiva, a “conjugação eficaz das inteligências e imaginações humanas” de forma “variada, distribuída por todos os lugares, constantemente valorizada, colocada em sinergia em tempo real, que engendra mobilização otimizada das competências” (p. 199). Lévy (1999) considera que o projeto da internet é o de emancipação dessas imaginações e inteligências para colocá-las a serviço do desenvolvimento das pessoas e dos coletivos, “uma exaltação do humano que se apoia nas disponibilidades técnicas de nossos dias” (p.208). O que é transformador, para Lévy (1999), não é a atuação humana coletiva em um ambiente sociotécnico, essa simbiose homem-máquina em rede, mas a inteligência coletiva que parece existir livre e independente das amarras tecnológicas.

Essa aparente oposição entre homem e máquina, ou cultura e técnica, é fortemente rejeitada por autores como Simondon (2007), por exemplo. Para ele, a cultura vem se constituindo como uma defesa do homem contra as técnicas, uma forma de se retirar dos objetos

técnicos sua humanidade. Há, para Simondon (2007), duas atitudes contraditórias em relação aos objetos técnicos: por um lado, esses objetos são vistos como um acúmulo de matéria, desprovidos de significados, importantes somente por sua utilidade; de outro, se são percebidos como animados, esses objetos carregam intenções hostis em relação ao homem, sendo, assim, um perigo constante de agressão e insurreição.

No entanto, em sua natureza e essência, objetos técnicos são parte intrínseca da cultura humana, “gestos humanos fixos e cristalizados em estruturas que funcionam” (p. 34). Nesse sentido, o que “reside nas máquinas é a realidade humana”(p. 12), com todas as funcionalidades e incertezas que derivam dessa natureza. O homem, para Simondon (2007), ao invés de ser um senhor de escravos, eternamente vigiando as intenções hostis das máquinas, é um organizador permanente de uma sociedade de “seres técnicos”(p. 13); ao mesmo tempo coordena e inventa esses objetos, é sociólogo e psicólogo das máquinas. Simondon (2007) considera, inclusive, que a relação homem-máquina resulta em um indivíduo técnico complexo, uma fusão entre objeto e homem que o permite agir no mundo natural.

Latour (2008), por sua vez, argumenta que o sistema sociotécnico é o resultado de uma negociação entre pessoas, instituições, organizações e máquinas. Segundo afirma, sociólogos, ao tentar analisar as ligações sociais, só observam “humanos frágeis e moralidades fracas” (p. 152), o que torna suas descrições insuficientes. Para ele, para que se construa uma descrição mais equilibrada das forças sociais, é necessário incluir elementos socialmente fortes e moralmente elevados, as máquinas. Uma porta eletrônica, que se abre na presença de um humano, é, para o autor, um personagem não-humano a quem delegamos uma função, mas também valores, deveres e ética. Ele decide, de acordo com o que é prescrito por quem a criou, quem pode ou não passar, geralmente um humano razoavelmente habilidoso, rápido o suficiente para passar antes que a porta se feche. Esse ator não-humano, também seguindo a prescrição de quem o criou, discrimina pessoas muito baixas ou muito idosas, além de quem carrega móveis e pacotes.

Considerar humanos e não-humanos como parte de um mesmo sistema é relevante, para Callon (2008), porque, desde que os estudos a respeito da cognição distribuída ganharam importância, considera-se que o conhecimento e as maneiras de se agir não se encontram encerrados no cérebro humano, mas estão distribuídos também em dispositivos materiais. Para Callon (2008, n.p.), é importante que se percebam os objetos e os não-humanos também como parte do social, em especial ao se discutir o que chama de “agenciamento sociotécnico”, ou seja, os efeitos de agenciamento que não se resumem aos corpos humanos e aos objetos que o prolongam, mas também aos “textos, materialidades, técnicas, conhecimentos abstratos e formais, etc.”, e aos conjuntos de configuração de arranjos, como mercados econômicos, por

exemplo. A maior vantagem, para Callon (2008), de se abandonar a oposição humano não-humano é a possibilidade de se poder estudar as diferentes formas de agência e ação, de forma empírica, sem termos que nos preocupar com essas duas categorias.

Uma constatação que, para nós, é interessante ao analisarmos as novas formas de sociabilidade que são percebidas com o surgimento da internet. O mundo social, como afirmam Barabási e Frangos (2002), se constitui sob a forma de uma rede complexa, densamente inter-relacionada, que engloba todos os aspectos da vida. Nesse mundo, ponderam os autores, há, certamente, um caminho que liga duas pessoas, ainda que desconhecidas, da mesma forma em que há “um caminho entre dois neurônios, em nosso cérebro, entre duas empresas no mundo, entre dois compostos químicos e nossos corpos” (BARABÁSI; FRANGOS, 2002).

Uma outra crítica necessária a Lévy (1999) e a Castells (2003) relaciona-se à tentativa de ambos de atribuir à existência da internet uma espécie de elevação inexorável da humanidade, como se as sociedades fossem se tornar mais justas, coesas e democráticas por conta das potencialidades da rede mundial de computadores. Lévy (1999), por exemplo, considera que o surgimento do ciberespaço “acompanha, traduz e favorece uma evolução geral da civilização”(p. 25). Ainda que considere a possibilidade de que as transformações operadas pelo ciberespaço possam ser tão bruscas a ponto de excluir determinados grupos, ou de favorecer o surgimento de um conjunto de processos “retroativos”, tais como novas formas de dominação, de exploração, de controle e de “bobagem coletiva”(p. 30), abraçar a inteligência coletiva, ainda assim, seria o melhor caminho, já que “seu aspecto socializante, descompartmentalizante e emancipador”[...]”constitui um dos melhores remédios para o ritmo desestabilizante, por vezes excludente, da mutação técnica”(p. 25)<sup>1</sup>.

Da mesma forma, Castells (2003) considera que:

A história da criação e do desenvolvimento da Internet é a história de uma aventura humana extraordinária. Ela põe em relevo a capacidade que têm as pessoas de transcender metas institucionais, superar barreiras burocráticas e subverter valores estabelecidos no processo de inaugurar um mundo novo. Reforça também a ideia de que a cooperação e a liberdade de informação podem ser mais propícias à inovação do que a competição e os direitos de propriedade. (p. 4)

Posteriormente, CASTELLS (2013), ao discutir os movimentos sociais que surgiram a partir de 2009, vai sugerir uma espécie de democratização inevitável do mundo ao afirmar que os novos movimentos sociais vão se amplificar por conta da liberdade que lhes garante a internet. Para o autor, o projeto desses movimentos é o de reinventar a própria democracia, reforçando os valores das revoluções libertárias iluministas. Segundo afirma,

---

<sup>1</sup> Para Lévy (1999) esses processos sequer fazem parte da inteligência coletiva. Para ele, os processos “retroativos” são parte do ciberespaço, mas à inteligência coletiva atribui somente características transformadoras, inovadoras e positivas.

Esses movimentos sociais em rede são novos tipos de movimento democrático - de movimentos que estão reconstruindo a esfera pública no espaço de autonomia constituído em torno da interação entre localidades e redes da internet, fazendo experiências com as tomadas de decisão com base em assembleias e reconstituindo a confiança como alicerce da interação humana (CASTELLS, 2013, p. 142).

No entanto, essa visão pode ser contestada de diversas maneiras. Levitsky e Ziblatt (2018), por exemplo, ao discutir as fragilidades dos guardiões tradicionais do sistema democrático norte-americano, mostram que o enfraquecimento da mídia tradicional, especialmente pelo surgimento de uma mídia alternativa, que inclui TVs a cabo e sites de redes sociais, abriu caminho para o aparecimento de *outsiders*. Isso se dá porque, segundo os autores, o processo de reconhecimento nacional não está mais concentrado em um pequeno número de veículos mais ligados ao *establishment*. Nesse novo ambiente, surgem celebridades com reconhecimento público, e apoio popular, a toda hora, em plataformas que favorecem um discurso próximo ao do entretenimento, com mensagens mais radicais e que, assim, acabaram por radicalizar o eleitor conservador. Mesmo Castells (2018, n.p.) reconhece que a democracia liberal está em uma crise, cuja a raiz, acredita, é um enfraquecimento do Estado-nação decorrente da globalização da economia e da comunicação, que limita as capacidades de um estado em fazer frente a questões que se tornaram globais. A internet seria, assim, o lugar de construção de um universo que conduziria à personalização da política e campo de embates em que se busca a “destruição moral e da imagem de quem se postula como líder”.

Nesse sentido, para se compreender melhor as potencialidades sociais e econômicas da internet, nos parece necessário analisar aquela que é sua estrutura fundamental, a rede.

### 1.1 As redes

Um dos efeitos dos estudos a respeito da internet foi ter colocado em evidência a própria noção de rede. Castells (2003) considera a rede como sendo a forma mais básica de organização, tão antiga quanto o próprio homem. Citando Fritjof Capra, Castells (2004) afirma que não é uma estrutura típica do século XXI, mas um padrão comum a toda a vida: “onde quer que vejamos vida, vemos redes” (p. 4).

Para Castells (2004), é uma forma de organização mais eficiente do que as estruturas rigidamente hierarquizadas por conta de sua flexibilidade, escalabilidade e capacidade de sobrevivência. Flexibilidade, para Castells (2004), é a capacidade das redes de se reconfigurar de acordo com o ambiente e de contornar eventuais bloqueios para encontrar novas conexões; já a escalabilidade refere-se à capacidade das redes aumentarem e diminuírem seu tamanho sem

que se quebrem; e capacidade de sobrevivência refere-se à impossibilidade de se destruir uma rede atacando somente um nó, já que as redes, para ele, não têm centro.

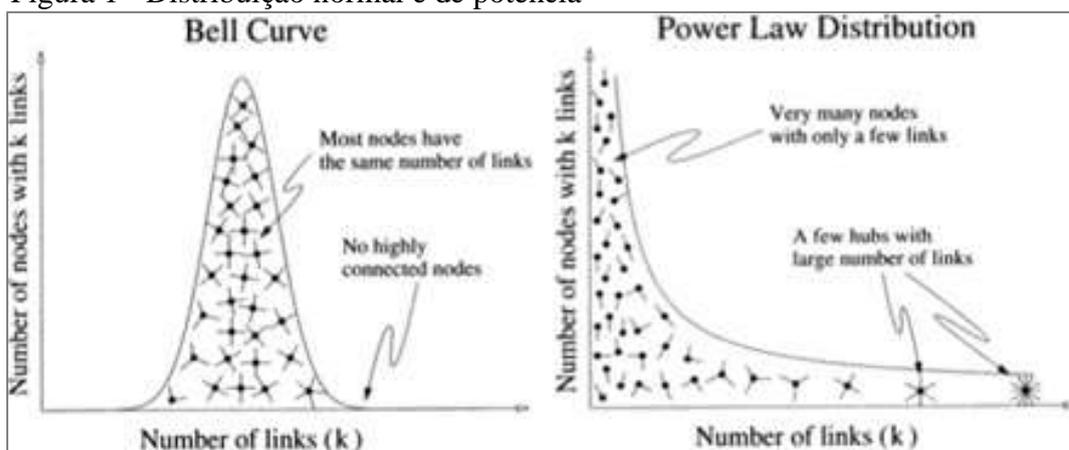
O trabalho de Barabási e Frangos (2002), no entanto, nos permite aprofundar a compreensão a respeito da topologia das redes e de suas regras básicas de constituição. Redes como a internet, para o autor, não são aleatórias, mas seguem uma lógica de auto-organização, própria de sistemas complexos. Seu crescimento, contínuo, é resultado das decisões individuais de milhares de usuários e segue dois princípios básicos: tempo de existência e adequação.

Os elementos básicos de qualquer rede são conexões e nós. À medida em que uma rede cresce, mais nós vão se conectando uns aos outros. Essa distribuição de conexões não é aleatória: cada novo nó tende a se conectar a nós com mais visibilidade e importância. Por isso, quanto mais antigo, mais provável que um nó tenha mais conexões. Daí a expressão de Barabási e Frangos (2002), *rich-get-richer*, ou seja, ricos ficam ainda mais ricos, ou nós que concentram muitos nós atrairão cada vez mais.

Essa primeira tendência explica a existência dos *hubs*, nós altamente conectados, e o fato de que a distribuição de nós em uma rede aberta, sem escalas e não aleatória, raramente segue uma curva normal, mas uma curva de potência.

Na natureza, assume-se que a distribuição mais comum segue uma forma de sino, também chamada de distribuição normal (figura 1). Como exemplo prático, se contarmos a altura de todos os indivíduos de dada comunidade, perceberemos que a maior parte deles terá uma altura média, e que alguns poucos indivíduos serão ou muito baixos, ou muito altos.

Figura 1 - Distribuição normal e de potência



Legenda: Exemplos de distribuição normal e distribuição de potência Fonte: Barabási e Frangos (2002)

Em redes como a internet, no entanto, existe um pequeno número de nós com um número muito grande de conexões, enquanto a maior parte deles são muito pouco conectados. A esse tipo de distribuição damos o nome de curva de potência. Sites como Facebook e Google, por exemplo, concentram milhares de vezes mais conexões do que um *website* particular,

conectado a uma ou duas páginas. Segundo Barabási e Frangos (2002), para efeitos práticos, esse *website* com poucas conexões é praticamente invisível e quase impossível de ser encontrado, já que os mecanismos de busca são enviesados contra ele<sup>2</sup>.

Outro efeito prático dessa característica das redes é que elas não são democráticas, justas ou igualitárias. Suas conexões e visibilidade não são igualmente distribuídas entre todos os nós. Redes como a internet são, de fato, dominadas por um pequeno número de conectores, visíveis de praticamente todos os pontos, que representam, para Barabási e Frangos (2002), o verdadeiro centro da rede.

Essa posição de viabilidade e centralidade, no entanto, não garante que esses conectores centrais ocuparão esse lugar para sempre. Google e Facebook substituíram, em espaço relativamente curto de tempo, a posição de outros atores mais relevantes, como o Yahoo! e o Orkut.

É por isso que Barabási e Frangos (2002) apontam para outra característica importante das redes: alguns nós parecem ter a capacidade de atrair conexões em velocidade maior do que outros. A esse comportamento Barabási e Frangos (2002) deu o nome de adequação. Toda rede tem um tipo específico de adequação. Em redes nos quais os nós têm níveis semelhante de adequação, a distribuição é normal. Naquelas em que alguns nós têm mais adequação que outros, a distribuição se torna uma curva de potência. Google, para usar o exemplo do autor, é centenas de milhares de vezes mais interessante do que um site individual. Naturalmente, acaba por atrair muito mais usuários.

## 1.2 Internet e economia

Quando analisa a internet a partir de uma ótica econômica, Castells (2003) afirma que apesar da expansão da rede ter sido estimulada por grandes empresas, que ajudaram na construção da infraestrutura necessária, a base dos sistemas foi inventada por pioneiros hackers e pelas elites tecnológicas, militares e acadêmicas, de forma comunitária e colaborativa. Como resultado, a internet não é, para o autor, dominada por comércio ou negócios, apesar desta ser uma parte muito significativa, mas traz em sua lógica estruturante os elementos que permitiram uma mudança profunda em todas as dimensões de um negócio.

---

<sup>2</sup> Veremos mais à frente, mas essa lógica se estende também aos projetos de financiamento coletivo. Enquanto alguns poucos projetos atraem milhares de apoiadores, e, à medida em que vão atraindo mais apoiadores, se tornam mais atraentes, a imensa maioria deles não recebe um número significativo de apoios. Essa dinâmica e os motivos serão discutidos nos capítulos 3 e 4.

Internet é, para Castells (2003), a espinha dorsal de uma nova economia, “erigida em torno de normas e processos novos de produção, administração e cálculo econômico” (p.49). Sua existência dá às empresas a possibilidade de circundar as regras e as limitações definidas pelas instituições dominantes, o que muda potencialmente todos os seus aspectos: “a relação com fornecedores e compradores, sua administração, seu processo de produção, sua cooperação com outras firmas, seu financiamento e a avaliação de ações em mercados financeiros” (p. 56).

Na base desses novos processos está a rede como forma de organização, algo que altera os componentes-chave dos processos de criação de valor capitalista: capital e trabalho.

Redes, segundo Castells (2003), ficaram historicamente relegadas aos ambientes privados e familiares quando houve o surgimento de organizações capazes de atingir metas centralmente definidas, caso de Estados e corporações. O que aconteceu ao surgir a internet foi que as redes se proliferaram em diversos domínios da economia e da sociedade, superando corporações e governos em capacidade de organização e desempenho. Em uma economia baseada em computador, essas redes extrapolam os ambientes locais e se tornam capazes de flexibilidade, de desempenho de tarefas, de comunicação global e horizontal, e de tomada de decisão coordenada e descentralizada, o que torna a rede, em sua visão, a melhor forma organizacional inventada pelo homem.

Uma empresa em rede não significa, para Castells (2003), que ela faz parte de uma rede de empresas ou que possui uma rede interna de computadores, mas sim que possui suas atividades econômicas organizadas em torno de projetos empresariais específicos. Ainda que exista uma empresa, com uma marca identificável, suas etapas produtivas, o conjunto administração-produção-distribuição, ou seja, o trabalho, é fragmentado e realizado em rede, sob forma cooperativa, compartilhada e deslocalizada.

Junto com a organização em rede, na fonte dessa nova economia estão as novas formas de circulação de capital. Desde a origem, *venture capitalists*, aqueles que assumem mais risco em busca de retornos superiores, financiaram a criação das empresas de internet, mesmo depois da crise da Nasdaq, ao final dos anos 1990. Em paralelo, os mercados financeiros aproveitaram a expansão da infraestrutura de telecomunicações para criar uma economia global interconectada, que os permite migrar investimentos de um canto a outro do globo quase instantaneamente, de maneira quase independente das regulações nacionais.

Esse mercado interconectado deu agilidade tanto a investidores institucionais, como bancos ou fundos de investimento, quanto a indivíduos, que conseguem fazer transações financeiras eletrônicas sem grande esforço, usando um *smartphone*, um serviço privado, como o PayPal, e um número de cartão de crédito. Com isso, a quantidade de investidores aumentou

exponencialmente, aumentando também a volatilidade dos mercados e o volume de transações. Contudo, criou-se, com isso, um incentivo para a oferta de novas formas de investimento: títulos de governo, moedas, participação em *startups*, *crowdfunding*, etc.

A conjunção desses dois elementos (infraestrutura digital e novas fontes de investimento) em um ambiente cultural propício deu força para uma nova dinâmica de inovação e empreendedorismo, que utiliza como combustível o conhecimento e a cooperação. Segundo Castells (2003), a inovação depende da geração de conhecimento, acessível em abundância na internet. E a cooperação é típica de um sistema de fonte aberta, como a internet. O resultado é que a economia digital tem como características, para o autor, efeitos de rede (quanto mais nós tem uma rede, mais benefício para cada nó), dependência de caminho (a cada inovação e em seus desdobramentos, seus descobridores têm uma vantagem competitiva decisiva) e retornos crescentes (apesar do custo inicial de uma inovação ser alto, os custos marginais decrescem rapidamente ao logo da sua adoção em produtos e serviços).

Esse cenário que descrevemos (ambiente cultural, infraestrutura de comunicação, organização em rede e liberalização dos fluxos de capital) tem efeitos significativos. Para Shirky (2008), um desses efeitos está na redução drástica dos custos de transação. Segundo Shirky (2008), Ronald Coase (1937), em seu trabalho “A Natureza da Firma”, respondeu a uma das questões mais importantes da economia: por que as organizações existem? Por que criar empresas, governos, ou qualquer outro tipo de instituição, se os indivíduos podem ir ao mercado e vender e comprar serviços?

A resposta está na lógica dos incentivos e dos custos de transação. Vendedores e compradores precisam se encontrar, firmar contratos, produzir, etc, atividades que consomem recursos (tempo, capital, mão-de-obra, material, etc.). Logo, um mercado completamente aberto sempre teria uma performance do trabalho reduzida, já que parte do trabalho estaria desviada da produção. Daí surge a necessidade dos processos de gestão e das tarefas compartilhadas.

Estruturas hierarquizadas têm como atributos, para Shirky (2008), a racionalização de recursos e a simplificação dos esforços de comunicação. Elas existem porque conseguem tirar vantagem do esforço conjunto. O problema é que a existência de gestores é, em si, um custo de transação, que, por sua vez, é tanto maior quanto for o tamanho desta organização. Portanto, para continuar crescendo, organizações precisam encontrar maneiras de superar esses custos: as empresas precisam aumentar faturamento, o estado, os impostos, e a igreja, as doações.

Daí surgem as duas condições para a existência da firma: a organização só é justificada se a riqueza produzida por ela for superior aos custos de transação (também chamado de Piso Coaseano) e inferior aos custos de gestão (ou Teto Coaseano). É por isso que projetos ou

atividades que não têm um grande potencial econômico, que são exóticos, que dariam pouco ou nenhum retorno financeiro, como é o caso da maior parte dos projetos apresentados em plataformas de financiamento coletivo, raramente seriam executados em um ambiente tradicional porque os investidores não teriam os incentivos necessários.

A questão, para Shirky (2008), é que o mercado atual e as ferramentas digitais fizeram que os custos de transação – tais como compartilhamento, cooperação, coordenação

– desabassem, o que permite não só a criação de novas estruturas organizacionais, mas também que atividades que anteriormente seriam economicamente inviáveis, ou sequer imaginadas como possibilidade, hoje existam com relativa facilidade.

Amadores que se reúnem para criar uma enciclopédia, fotógrafos não profissionais que juntam para produzir imagens de atentados terroristas, organizações políticas extremistas globais, grupos de ódio, analistas econômicos não profissionais, todo tipo de grupo hoje é capaz de se juntar, de forma distribuída, e produzir algo com custos de transação quase inexistentes. Para Shirky (2008), isso é resultado do surgimento de ferramentas de comunicação e de padrões sociais novos, que não criam, por si, ação coletiva, mas removem os obstáculos para que ela aconteça.

Outra característica deste novo momento, segundo Shirky (2010), tem a ver com o que chamou de excedente cognitivo. Como consequência da industrialização e das legislações trabalhistas, nunca tivemos tanto tempo livre. Uma das palavras de ordem utilizadas nos protestos de trabalhadores no século XIX, lembra Shirky (2010, n.p.), dizia “oito horas de trabalho, oito horas para dormir, e oito horas para fazermos o que quisermos”. Daí a origem das 40 horas de trabalho semanais.

Esse tempo conquistado acabou não sendo todo investido, ao longo do século XX, em atividades comunitárias, em leitura de livros ou desenvolvimento de atividades artísticas. Para Shirky (2008), grande parte deste tempo foi destinado ao consumo de mídia audiovisual, especialmente a televisiva. Segundo afirma, o cidadão médio americano assiste a milhares de horas de televisão todos os anos, uma atividade que tem um custo de acesso especialmente baixo.

Com o surgimento das tecnologias digitais, esse cenário mudou ligeiramente. A população mundial conectada, segundo Shirky (2008), continua assistindo a trilhões de horas de TV todos os anos. No entanto, parte significativa desse excedente cognitivo migrou para atividades que permitem mais participação e colaboração, também com custo de aquisição baixo. Quando se possui um *smartphone*, não só temos acesso a uma plataforma de consumo, como é a TV, mas também a uma máquina que permite produzir e fazer circular essa produção.

E as evidências apontam, para o autor, que quando se oferece a possibilidade de produzir e compartilhar, parte das pessoas, ainda que nunca tenham feito isso antes, nem que consigam produzir conteúdo profissional, aproveitam essa oportunidade.

Para Shirky (2008), ter acesso a esse excedente, ainda que a pequena parte dele, já permite a organização de enormes projetos colaborativos; 1% que seja da atenção do mundo conectado corresponde a centenas de iniciativas do tamanho da Wikipédia todos os anos.

Atividades que demandam participação, para Shirky (2008), têm uma natureza diferente daquelas meramente contemplativas. A participação demanda presença, uma resposta para dado estímulo ou evento, e esse desejo de participação, de comunhão e compartilhamento é algo intrínseco à humanidade. A questão, para Shirky (2008) é que os custos de descobrimento eram altos demais para que atividades antes vistas como hobby tivessem visibilidade ou fossem facilmente acessíveis.

Usando o exemplo do início do capítulo, imagine que alguém seja um aficionado por café em um mundo analógico. Teria que participar presencialmente de feiras e eventos temáticos (a maior feira de café do Brasil acontece anualmente em Belo Horizonte); se corresponder por carta com seus pares, se conseguisse estabelecer esses laços nos encontros; assinar revistas especializadas ou *newsletters* impressas, que chegariam em sua casa provavelmente uma vez por mês; e, para encomendar grãos raros, seria necessário telefonar ou enviar cartas para esses fornecedores (e talvez não encontrá-los no momento da ligação).

Em um mundo de comunicação digital, esses custos de descobrimento foram praticamente eliminados. Em poucos minutos, pode-se encontrar grupos que compartilham seus gostos e seus interesses, por mais exóticos. Com isso, é relativamente simples para um amador se engajar em projetos colaborativos, ainda que não tenha retorno financeiro por isso. Essas atividades, que hoje competem com profissões estabelecidas, são, muitas vezes, remuneradas com reconhecimento social e com satisfação pessoal. Para Shirky (2008), essa produção social não demanda supervisão gerencial ou incentivos salariais – basta que um grupo a veja como interessante, importante ou urgente o suficiente para que se justifique.

Tapscott e Williams (2007) identificaram tendência similar ao analisar os projetos de colaboração existentes na internet. Para eles, é inédita a capacidade atual de se canalizar o conhecimento, os recursos e o esforço intelectual de um volume enorme de pessoas. O termo que cunharam, *wikinomics*, é inspirado nos softwares colaborativos de edição de texto, e designa um momento de colaboração massiva, tanto entre indivíduos como com corporações e governos, de diferentes formas, que vai ter como resultado a dispersão do poder, do conhecimento e da capacidade produtiva.

Na base dessa dinâmica economia, para os autores, está o acesso à tecnologia, sobretudo à internet. Qualquer pessoa com um computador, uma conexão de rede, iniciativa e criatividade pode participar de processos colaborativos globais e ajudar a desenvolver produtos como Linux, MySQL e a Wikipédia. Através do *peer production*, ou *peering*, a produção por pares, por iguais, é possível desenvolver um projeto, resolver questões intelectuais, ou criar bens e serviços, com ou sem interesse financeiro.

Já Anderson (2006) notou, como resultado deste novo ambiente socioeconômico, o surgimento de novos modelos de negócio com base em comportamentos de nicho. A sua teoria, chamada Cauda Longa, parte do pressuposto de que em uma economia de abundância e de eficiência em distribuição, fabricação e comunicação, os critérios de viabilidade econômica mudam.

Em uma curva de oferta e demanda, a demanda quase nunca chega a zero. Isso significa que sempre há algum interessado, em algum lugar, por alguma oferta. Ao mesmo tempo, vivemos uma situação na qual a oferta, de qualquer coisa, entretenimento, produtos, serviços, é abundante. A questão é que em uma economia de altos custos e de ineficiências de estocagem, produção, distribuição e venda, nem toda oferta faz sentido econômico porque não há retorno financeiro significativo em um volume pequeno de vendas. Logo, quem oferece, reduz a quantidade de itens comercializados àqueles que são mais vendidos ou que possuem uma margem maior (figura 2).

Uma loja de departamento, por exemplo, mesmo possuindo uma área muito grande, seu estoque é sempre limitado. Por isso, a oferta dessa loja estará sempre condicionada pela preferência dos públicos que a acessam fisicamente. Tradicionalmente, considera-se que 20% da oferta pode trazer 80% do faturamento. Logo, essa loja tentará sempre adivinhar qual o *mix* de ofertas tem uma relação ótima entre oferta e demanda.

Se essa mesma loja tiver uma versão na internet, a sua lógica de funcionamento muda completamente. Seu público consumidor não está necessariamente localizado geograficamente, logo não está restrita aos gostos dos seus vizinhos. Além disso, pode fazer um acordo com seus fornecedores e não ter custo de estoque: a cada pedido de um cliente, o fornecedor envia diretamente o produto, sem passar pela loja. Com isso, ao invés de oferecer somente aqueles produtos que mais vendem, pode oferecer praticamente todos os produtos. Obviamente, cada um desses produtos menos buscados será vendido com menos frequência, mas o conjunto dessas ofertas de nicho é potencialmente tão grande que isso se torna uma parte muito relevante do faturamento.

Para Anderson (2006), a Cauda Longa não é exatamente um efeito da internet, mas o resultado de três forças econômicas baseadas nas inovações técnicas desenvolvidas nos últimos anos: a democratização das ferramentas de produção, a democratização das ferramentas de distribuição e a ligação entre oferta e procura.

A democratização das ferramentas de distribuição resulta da expansão da internet em conjunto com novas técnicas de logística e de cadeia de suprimentos. A partir dos anos 70, grandes varejistas, aproveitando-se, por sua vez, da evolução das empresas de transporte, de todos os modais, desenvolveram estruturas supereficientes de distribuição. Essas estruturas se tornaram ainda mais baratas com a possibilidade de comunicação de baixo custo e instantânea da internet.

A ligação entre oferta e procura, por sua vez, tem raízes nos algoritmos de recomendação, como o que existe na Amazon, por exemplo, nos agregadores de publicidade, e na possibilidade de se utilizar as opiniões de usuários de internet para direcionar quem busca por informação, como faz o Google.

A democratização das ferramentas de produção, por fim, é resultado dos avanços da eletrônica e da computação. Para ANDERSON, hoje indivíduos conseguem produzir conteúdo com qualidade que era reservada a profissionais há alguns anos, num custo muito baixo. Com isso, a oferta de obras audiovisuais, textos, etc, é a maior que se tem registro.

Figura 2 - A Cauda Longa



Legenda: Relação entre negócios de *cauda longa* e de *cabeça* da curva de oferta e demanda. Fonte: Anderson (2006)

Anderson (2012) percebe, ainda, que a democratização das ferramentas de produção não aconteceu somente do lado do conteúdo. Para ele, estamos vivendo um novo momento

econômico, que ele chamou de *Maker Movement*, no qual é possível utilizar computadores para desenvolver projetos digitalmente e, ao invés de vender suas ideias para uma indústria estabelecida, esse inventor pode se tornar empreendedor e passar a produzir e comercializar, por si mesmo, suas invenções.

Para ele, estamos em uma nova revolução industrial, que combina manufatura digital e pessoal, uma espécie de industrialização do espírito “faça-você-mesmo” com base na digitalização dos projetos, na automação e no barateamento das ferramentas de corte e impressoras 3D. Nesse ambiente, um inventor pode produzir suas criações por si mesmo ou mandar seu projeto digitalmente para fábricas em qualquer lugar do mundo onde encontre os recursos que necessite. Essa fábrica não só poderia produzir dezenas de milhares de unidades de sua invenção, mas ainda participar do processo de distribuição para os clientes.

Como esses realizadores são da “geração *web*”, acabam por compartilhar suas criações online, alimentando a cultura de colaboração. Ao invés de criarem a partir do zero, esses novos inventores participam de comunidades de aficionados, dividindo avanços e dilemas, apoiam-se nos outros membros de sua comunidade para avançar seus projetos, e utilizam-se dessa força de trabalho em rede para aprimorar o design – o que, para ANDERSON, é um processo de criação de valor único.

Costa (2017) descreve um outro aspecto desse momento ao analisar o que chamou de economia da confiança. Neste ambiente que descrevemos, algumas práticas que existiam organizadas de forma privada ou reguladas por grandes instituições mediadoras, tais como “o hábito de pegar algo emprestado com um vizinho ou parente, de pedir caronas, de organizar a chamada ‘vaquinha’, dividir o espaço de trabalho com um colega, jantar com alguns amigos e hospedar outros, usar bibliotecas e o transporte coletivo” (p.23), tornaram-se viáveis como modelo de negócio.

Alguns destes novos negócios tornaram-se extremamente relevantes e parte do nosso cotidiano. Uber (monetizar a carona), Airbnb (alugueis de temporada), Couchsurfing (hospedagem na casa de amigos), Udemy (aulas particulares), eBay (brechó), são todos serviços bem-sucedidos que se apropriaram de comportamentos já existentes, mas que não teriam escala para serem viáveis financeiramente em outro momento.

COSTA (2015) faz uma distinção entre economia colaborativa, criativa e solidária. Para ele, a economia criativa é um ramo da economia que teria uma relação maior com as artes e com o design, e cujo insumo principal é a criatividade. Já a economia solidária carregaria uma crítica ao sistema capitalista; seria uma forma de produção que se propõe mais ética e igualitária, na qual o objetivo não seria o lucro, mas a divisão do excedente produtivo.

Apesar de encontrar similaridades com essas ideias, COSTA (2015) defende que a economia colaborativa é uma aproximação entre relações comerciais e comunitárias facilitada pelos suportes tecnológicos disponíveis, e não uma atividade econômica com insumo específico, a criatividade, ou que se apresente como contraponto ao capitalismo, já que não carrega as críticas habituais ou propõe alternativas ao sistema. Pelo contrário:

é um modelo de negócios bem estabelecido, com um volume de mercado importante, cujas características nos fazem refletir a respeito das noções tradicionais de consumo e de propriedade<sup>3</sup>.

O que diferencia as empresas da economia colaborativa, para COSTA (2015), é a existência de plataformas tecnológicas que organizam atividades que já aconteciam informalmente, como empréstimos de roupas, livros ou automóveis, compartilhamento de escritório, caronas, hospedagem, etc., em troca de comissão. Seus modelos de negócio tendem a ser fortemente ancorados em dois eixos: 1) tecnologia, mais especificamente na internet e em aplicativos mobile, o que aponta para um papel central das tecnologias na criação de redes e comunidades; e 2) a confiança entre desconhecidos<sup>4</sup>, garantida parcialmente pela empresa que oferece a plataforma de colaboração.

A crítica mais frequentemente em relação a negócios desta natureza relaciona-se ao esforço que fazem para se sobrepôr à regulação de mediadores institucionais, seja para se estabelecer como negócio e concorrer com serviços altamente regulados, como por uma busca por garantir que seus custos fixos mantenham-se baixos.

Fisk (2020), por exemplo, defende que a *Gig Economy*<sup>5</sup> é nova no que se refere à sua capacidade de coordenar uma força de trabalho imensa enquanto garante um bom nível de serviço, o que se tornou possível graças à tecnologia digital. No entanto, herda os velhos problemas do capitalismo moderno quando busca reduzir as proteções garantidas pelas leis trabalhistas. Para a autora, o seu comportamento é o mesmo daquelas que cortaram os vínculos trabalhistas dos funcionários, passando a considerá-los prestadores de serviços independentes, quando as leis passaram a garantir mais direitos a eles, o que significou mais custos para a

---

<sup>3</sup> Acevedo (2020), inclusive, critica o uso da expressão economia do compartilhamento por dar um ar altruísta à prática e sugerir que as empresas típicas dessa economia não teriam como objetivo a comercialização de bens ou serviços, o que é falso.

<sup>4</sup> Para COSTA (2015) esse ponto é importante porque não se pode considerar toda empresa que se utilize fortemente da internet, como o Netflix, por exemplo, como sendo parte da economia colaborativa. Tampouco esforços colaborativos como a Wikipedia, que se apresentam mais como ambientes de troca de conteúdo do que como plataformas comerciais, o deveriam.

<sup>5</sup> Gig economy é o termo utilizado para caracterizar um conjunto de empresas que proveem serviço a consumidores com base no trabalho de indivíduos que não têm contrato de trabalho formalizado, ou que possuem contratos considerados precários. É o caso, por exemplo, do Uber, que considera seus motoristas como sendo seus parceiros e não seus funcionários.

empresas. Para Fisk (2020), os grupos de trabalhadores que se saem melhor numa economia assim são os que conseguem negociar coletivamente, e não individualmente, pagamentos e benefícios.

Para Acevedo (2020), o que caracteriza essas empresas é o fato delas participarem das transações que mediam ao mesmo tempo que consideram que as salvaguardas regulatórias que deveriam orientar essas relações não se aplicam a elas<sup>6</sup>. Para Acevedo (2020), esse comportamento acontece em meio a uma tendência de enfraquecimento das condições trabalhistas, refletida no processo de dessindicalização da força de trabalho, enquanto há um aumento no poder dessas empresas e uma percepção da importância de continuação desses serviços. A pandemia da COVID-19 mostrou o quão relevantes são os serviços de entrega e de distribuição de refeições, por exemplo, realizados, em sua maior parte, por esses trabalhadores independentes. Em muitos países, esses serviços foram considerados essenciais e, com isso, nunca foram submetidos às regras de quarentena ou de restrição de movimento. Acevedo (2020) acredita que a crise sanitária pode até avançar a discussão a respeito da necessidade de maior proteção a esse grupo de trabalhadores, sobretudo considerando que em vários lugares eles foram incluídos, como categoria, nos esforços de proteção dos trabalhadores. No entanto, considera que não há consenso a respeito de como o que chama de *gig work* deve ser ordenado, uma discussão, considera, que deve ainda se alongar.

Naquilo que cabe a este projeto, essa reflexão serve como mais um alerta contra os determinismos e nos ajudam a colocar em perspectiva uma visão original, talvez excessivamente otimista, a respeito das maneiras como as relações mediadas por tecnologias se estabelecem.

---

<sup>6</sup> Acevedo (2020) alerta que nem toda empresa de tecnologia se encaixa nessa definição, que é fruto de bastante controvérsia. Couchsurfer, uma empresa típica da economia do compartilhamento, não interfere, como faz Airbnb e Uber, na relação entre quem precisa de um lugar para dormir e quem pode oferecer. Tanto Airbnb como Uber construíram sistemas de controle de qualidade, através de ranqueamento e monitoramento em tempo real, que obrigam seus “parceiros” a se comportar de determinada maneira, ao mesmo tempo em que se eximem da responsabilidade de se adotar padrões e exigências legais, como higienização dos locais e dos veículos, pagamento de salário mínimo e recolhimento de tributos, o que aconteceria com empresas tradicionais.

## 2 O FINANCIAMENTO COLETIVO

### 2.1 Definições e histórico

O *crowdfunding* é a formalização, através de um intermediário, da tradicional “vaquinha”, a cotização entre membros de um grupo para viabilizar um projeto. Segundo Dresner (2014), é geralmente definido como sendo uma cooperação coletiva de pessoas que concentram recursos financeiros para apoiar iniciativas de outras pessoas ou organizações. Em sua versão contemporânea, esses intermediários são plataformas digitais, sob forma de sites ou aplicativos mobile, que servem de ponte entre os empreendedores e os financiadores, em troca de comissão sobre os valores arrecadados.

É uma prática construída em torno do compartilhamento social (CARR, 2013; MOLLICK, 2014) que permite que grupos ajam coletivamente para apoiar iniciativas de diversas naturezas, como: produção de filmes, música, apoio a filantropia, pesquisa científica, reportagens especiais, revistas em quadrinhos, desenvolvimento de novos produtos, etc.

É também, do ponto de vista da economia, uma maneira de se financiar projetos em estágio embrionário (os chamados *early-stage projects*), que podem, para Agrawal, Catalini e Goldfarb (2013), ter consequências econômicas significativas, tanto positivas como negativas.

Para Dresner (2014), é uma maneira nova de descrever uma prática milenar. Segundo o autor, há 5 mil anos, antes de existirem bancos ou instituições financeiras, famílias ricas e governantes se juntavam para viabilizar os empreendimentos necessários às suas comunidades, desde novos negócios a infraestrutura. Mais recentemente, depois da Grande Depressão norte-americana, em 1930, fazendeiros passaram a se juntar em cooperativas de crédito para escapar dos altos custos de empréstimos financeiros bancários. Esse modelo foi adaptado, nos Estados Unidos, por outros grupos para financiar, por exemplo, a compra de casas para a classe média.

Dresner (2014) aponta, ainda, outros exemplos de financiamento coletivo antes dos sites de *crowdfunding*. A Estátua da Liberdade, ícone de Nova Iorque e dos Estados Unidos, foi construída em 1876 com doações de cidadãos franceses, que cotizaram para a construção da estátua, e de americanos, que pagaram pelo pedestal depois de uma campanha liderada por Joseph Pulitzer em seu jornal, o *The World*. Brüntje e Gajda (2016) lembram que, em 2008, a campanha de Barack Obama para a presidência dos Estados Unidos arrecadou US\$ 250 milhões de dólares em contribuições de pessoas físicas, que doaram cada um, em média, US\$ 200.

Um exemplo mais próximo é a construção do Estádio Mario Filho, o Maracanã. Palco principal da Copa do Mundo de 1950, com capacidade de 200 mil pessoas, o que representa

10% da população do Rio de Janeiro naquele momento, o projeto só foi viável com a venda de 30 mil títulos de cadeiras cativas (GLOBO, 2013).

Com a internet, surgiram o ambiente cultural e os comportamentos necessários para que o financiamento coletivo se tornasse uma prática mais significativa e relevante do ponto de vista econômico e social. Para Agrawal, Catalini e Goldfarb (2013), isso tem a ver com três características do contexto atual:

- a) o custo de fazer com que empreendedores e potenciais financiadores se encontrem caiu drasticamente;
- b) a exposição ao risco é imensamente reduzida porque esse tipo de financiamento é realizado de forma incremental e fragmentada, e não de uma só vez, como no caso de investidores de risco; e
- c) a redução dos custos de comunicação facilitam uma melhor aquisição de informação e monitoramento do andamento dos projetos, além de permitir que financiadores participem do desenvolvimento das ideias, o que reduz, mas não elimina, as assimetrias de informação entre criadores e financiadores, em especial no que diz respeito à habilidade do criador em materializar seu projeto e entregar os produtos.

As primeiras plataformas similares ao que conhecemos hoje como *crowdfunding* apareceram em meados dos anos 2000. Em 2005, por exemplo, foi lançada a Kiva, uma empresa cujo propósito era oferecer micro-empréstimos de cidadãos norte-americanos para empreendedores de países em desenvolvimento, com um propósito claramente social. Segundo Dresner (2014), era possível pegar empréstimos de US\$ 100 a US\$ 1 mil, que esses empreendedores utilizavam para comprar motocicletas, vacas, fornos – insumos necessários para o estabelecimento de pequenas iniciativas comerciais.

A primeira iniciativa de financiamento coletivo tal como estamos analisando neste estudo foi a plataforma alemã Sellaband. Criada em 2006, funcionava como um espaço “to unite artists and fans in an independent movement that aims to level the playing field in the global music industry”<sup>7</sup>. O Sellaband, que pediu concordata em 2010 e fechou definitivamente em 2015, auxiliava a financiar eventos musicais, produções e novas bandas e podia ser utilizado tanto por artistas independentes como por produtoras, gravadoras ou empresas de mídia, para financiar seus projetos musicais e para criar uma base de fãs para novos artistas (COCATE; JÚNIOR, 2016).

---

<sup>7</sup> Tradução nossa: “para unir artistas e fãs em um movimento independente que almeja nivelar o terreno da indústria global de música”

A crise de 2008 talvez tenha sido, no entanto, o maior impulsionador para a criação de plataformas de crowdfunding. Iniciada após o estouro de uma bolha no mercado imobiliário norte-americano, a crise atingiu a economia de quase todos os países em diversos setores, mas, muito mais fortemente, os bancos. Se em condições normais essas instituições já impunham muitas barreiras para emprestar capital, principalmente para a criação de novos negócios, logo depois da quebra do Lehman Brothers, uma das cinco grandes instituições do tipo nos Estados Unidos naquele momento, esse capital praticamente desapareceu (DRESNER, 2014).

É nesse contexto que aumentam as iniciativas desta natureza em diversas partes do mundo e não é surpreendente que a maior plataforma hoje, o Kickstarter, surja logo depois, em 2009.

## 2.2 A relevância do financiamento coletivo

O *crowdfunding* como prática é relevante porque se apresenta como forma não tradicional de financiamento para novos empreendimentos, principalmente para aqueles que não reúnem as características típicas do mundo dos negócios.

Em resumo, a teoria econômica considera que investidores privados apoiam a criação de novos empreendimentos em troca da oportunidade de se apropriar de parte dos retornos financeiros futuros (CARR, 2013). No entanto, há diversos riscos intrínsecos a esse tipo de investimento, como margens menores, problemas operacionais, entraves regulatórios, pouco talento de gestão, etc., riscos estes que são pouco previsíveis antes de um negócio funcionar de fato. O trabalho de um investidor, portanto, consiste em equilibrar as possibilidades de retornos futuros com os riscos percebidos em um determinado investimento (MOLLICK, 2014). Em outras palavras, o investidor vai tentar buscar o maior retorno possível dentro de uma faixa que considera adequada de risco.

O problema é que mesmo depois de análises econômicas refinadas, as informações disponíveis antes de um empreendimento existir de fato continuam sendo, quase sempre, parciais e de confiabilidade limitada. Como consequência, investidores tradicionais são naturalmente resistentes a colocar seus recursos em projetos cujas expectativas de retorno são baixas e nos quais esses sinais de qualidade não são facilmente percebidos, o que pode tornar projetos menores, pouco estruturados ou que tenham aparência amadora potencialmente inviáveis (CARR, 2013).

E justamente era essa a preocupação de Perry Chen ao criar o Kickstarter. Como integrante de uma banda, Chen percebia a dificuldade que existia para se levantar o di-

necessário para fazer um tour ou gravar músicas. Inspirado no Sellaband, imaginou conectar artistas e seus fãs para que, juntos, conseguissem viabilizar novas produções e shows. Para cada doação, o fã receberia, em troca, um exemplar do que fosse produzido ou um presente, que consistia, geralmente, segundo Dresner (2014), em um CD, um DVD ou uma camiseta. Com isso, criou um dos principais princípios de funcionamento dos sites deste tipo hoje: o financiamento em troca de recompensas.

Guardadas as devidas proporções, essa é também uma das preocupações do Banco Mundial. Em um relatório chamado *Crowdfunding's Potential for the Developing World* (BEST et al., 2013), a instituição afirma acreditar que o financiamento coletivo tem o potencial de fomentar o desenvolvimento de novos negócios e de criar inovação tecnológica, ao estilo *Silicon Valley*, em várias partes do mundo. Há, para o relatório, um potencial não explorado de desenvolvimento econômico e social em países em desenvolvimento por falta de incentivos a investidores tradicionais. Nesse contexto, o financiamento coletivo pode, se criadas as condições para isso, auxiliar no desenvolvimento de uma cultura e um ecossistema de empreendedorismo nesses locais.

### 2.3 O mercado de *crowdfunding* hoje

Desde seu surgimento, o mercado de financiamento coletivo se desenvolveu bastante. Se no início era caracterizado pelos pedidos de pequenas doações e adiantamentos financeiros, hoje tornou-se economicamente relevante, com uma diversidade grande de modalidades e de aplicações.

Best et al. (2013) afirma que há projeções de todos os tipos para o tamanho deste mercado globalmente, de US\$ 3,5 bi até US\$ 300 bilhões anuais. No entanto, projetando o potencial de mercado a partir de uma perspectiva de renda familiar, ou seja, analisando

o quanto se pode utilizar da renda familiar média mundial para gastos desta natureza, o relatório acredita em um mercado potencial de US\$ 96 bilhões/ano em 2025, sendo que na América Latina esse mercado poderia chegar a quase US\$ 11 bilhões/ano.

Mesmo sendo muito otimista com relação à potencial utilização do financiamento coletivo em países em desenvolvimento, o relatório reconhece que esta ainda é uma prática típica de países com economia desenvolvida.

Nos Estados Unidos, por exemplo, Kuppuswamy e Bayus (2015) afirmam que, em 2014, os sites de *crowdfunding* movimentaram cerca de US\$ 16 bilhões, sendo que somente o Kickstarter foi responsável por faturar US\$ 2 bi. Best et al. (2013) considera que há um

crescimento de 60% anual nessas plataformas na região, com uma aceleração forte percebida a partir de 2012.

Tabela 1 - Número de plataformas de *crowdfunding* por país

País	Plataformas de CF	País	Plataformas de CF
EUA	344	Austrália	12
Reino Unido	87	Índia	10
França	53	Rússia	4
Holanda	34	África do Sul	4
Canadá	34	Hong Kong	1
Espanha	27	Estônia	1
Alemanha	26	Bélgica	1
Brasil	17	China	1
Itália	15	Emirados Árabes	1

Legenda: Relação de plataformas de financiamento coletivo identificadas até 2012 Fonte: Best et al. (2013)

Na Europa, os sites começaram a se desenvolver fortemente a partir de 2010, segundo Brüntje e Gajda (2016), sobretudo no Reino Unido, na Alemanha e na Holanda. O estudo que conduziram estima que foram arrecadados, em 2013, US\$ 5,1 bilhões.

Na África Subsaariana, Best et al. (2013) identificaram um crescimento sustentável no número de plataformas desde 2010, estimuladas pelo fortalecimento de uma classe média, pela rápida adoção de tecnologias *mobile* e uma necessidade real de mercado. Um dos exemplos apontados pelo relatório é o HomeStrings, que mobilizou US\$ 25 milhões para investir em projetos de infraestrutura em 13 países. Aceita inclusive projetos de governos, como o do Quênia, de Gana e da Nigéria.

Best et al. (2013) listaram iniciativas relevantes, ainda, na América do Sul, no Oriente Médio, no Norte da África e na Ásia, o que, segundo afirmam, indica o interesse de empreendedores e governos em utilizar e adaptar os modelos de *crowdfunding* existentes às suas regiões (tabela 1). Apesar de perceberem um aumento nos projetos dessas plataformas em países em desenvolvimento, Best et al. (2013) percebem, ainda, que esse crescimento depende de fatores que variam por região. No entanto, em termos gerais, nestes locais, dado o pouco desenvolvimento tecnológico, são poucos os projetos ligados a tecnologia. Na maior parte das vezes, os projetos tendem a ser mais voltados a iniciativas artísticas e causas sociais.

Tabela 2 - Número de projetos por categoria

Categoria	Total
Música	2.025 (13%)
Cinema e vídeo	1.627 (10%)
Literatura	1.402 (9%)
Pessoais	1.398 (9%)
Educação	1.375 (9%)
Socioambiental	1.373 (9%)
Teatro e dança	954 (6%)
Ciência e tecnologia	828 (5%)
Artes	798 (5%)
Eventos	765 (5%)
Quadrinhos	702 (4%)
Esporte	495 (3%)
Jogos	442 (3%)
Fotografia	395 (3%)
Design e moda	303 (2%)
Jornalismo	287 (2%)
Gastronomia	260 (2%)
Arquitetura e urbanismo	189 (1%)
<b>Total</b>	<b>15.618 (100%)</b>

Legenda: Número de projetos apresentados, separados por categoria, na plataforma Catarse, de 2011 a 2018

Fonte: Elaboração do autor

No Brasil, há diversas plataformas deste tipo<sup>8</sup>, sendo que a principal é o Catarse. Criado em 2011 e inspirado na experiência do Kickstarter, seu objetivo é ser “o ponto de encontro de uma crescente comunidade de pessoas interessadas em, juntas, contribuírem financeiramente para tirar do papel projetos de diversas categorias” (CATARSE, 2017c, n.p.). Oficialmente, o Catarse arrecadou, em 2019, R\$ 29,5 milhões e publicou 2.850 projetos receberam recursos (CATARSE, 2019).

Na base de dados com que estamos trabalhando, obtida através de *web scraping* do site do Catarse, foram realizados, até 25 de maio de 2018, 15.618 projetos, sendo que a maior parte deles nas categorias música, cinema e vídeo e literatura (tabela 2).

T

<sup>8</sup> Costa (2017) identificou como relevantes, naquele momento, pelo menos 7: [www.doare.org](http://www.doare.org), [www.juntos.com](http://www.juntos.com), [www.benfeitoria.com](http://www.benfeitoria.com), [bicharia.com.br](http://bicharia.com.br), [www.variavel5.com.br](http://www.variavel5.com.br); [www.cineasta.cc](http://www.cineasta.cc) e [www.freedomsporsors.org](http://www.freedomsporsors.org). No momento da elaboração deste texto, encontramos, além destas, [www.apoia.se](http://www.apoia.se), [sharity.com.br](http://sharity.com.br), [www.kickante.com.br](http://www.kickante.com.br), [evoe.cc](http://evoe.cc) e [eqseed.com](http://eqseed.com)

Tabela 3 - Número de projetos bem-sucedidos por categoria

Categoria	n	Percentual
Música	685	20%
Cinema e vídeo	490	14%
Quadrinhos	371	11%
Socioambiental	343	10%
Literatura	325	9%
Teatro e dança	283	8%
Educação	160	5%
Artes	159	5%
Eventos	122	4%
Jogos	116	3%
Fotografia	111	3%
Jornalismo	81	2%
Ciência e tecnologia	50	1%
Design e moda	47	1%
Pessoais	45	1%
Esporte	42	1%
Arquitetura e urbanismo	38	1%
Gastronomia	9	0%
<b>Total 3477</b>	<b>100%</b>	

Legenda: Número de projetos que alcançaram suas metas, por categoria, no Catarse entre 2011 e 2018

Fonte: Elaboração do autor

Nessas categorias estão também o maior número de projetos que tiveram sucesso, ou seja, aqueles que atingiram ao menos 100% da meta estabelecida, independentemente da modalidade na qual se insere (tabela 3).

Esses dados colocam o Catarse em um estágio similar ao início do Kickstarter, quando os projetos eram mais ligados às artes e à música. Segundo Agrawal, Catalini e Goldfarb (2013), somente mais tarde é que surgiram os projetos ligados a tecnologia.

Se observarmos quais categorias têm melhor taxa de sucesso (tabela 4), perceberemos que o Catarse é uma plataforma dedicada aos quadrinhos (53% de sucesso), à música (34%), ao Cinema e ao vídeo (30%), e ao Teatro e a Dança (30%). As categorias que não conseguem bons resultados são Pessoais (3%), Gastronomia (3%) e Ciência e tecnologia (6%).

Considerando os dados levantados em nosso banco, o valor total arrecadado pelo Catarse desde 2011 é de pouco mais de R\$ 55 milhões e, considerando a comissão cobrada (13%), é provável que eles tenham faturado cerca de R\$ 7 milhões e distribuído R\$ 47.940.697,5 aos empreendedores<sup>9</sup>.

<sup>9</sup> Na retrospectiva de 2019, o site afirma já ter faturado, desde a criação, pouco mais de R\$ 100 milhões, o que aponta um crescimento significativo nos últimos anos (CATARSE, 2019).

Tabela 4 - Categorias com mais sucesso

Categoria	Apresentados	Bem-sucedidos	% de sucesso
Quadrinhos	702	371	53%
Música	2025	685	34%
Cinema e vídeo	1627	490	30%
Teatro e dança	954	283	30%
Jornalismo	287	81	28%
Fotografia	395	111	28%
Jogos	442	116	26%
Socioambiental	1373	343	25%
Literatura	1402	325	23%
Arquitetura e urbanismo	189	38	20%
Artes	798	159	20%
Eventos	765	122	16%
Design e moda	303	47	16%
Educação	1375	160	12%
Esporte	495	42	8%
Ciência e tecnologia	828	50	6%
Gastronomia	260	9	3%
Pessoais	1398	45	3%
<b>Total</b>	<b>15618</b>	<b>3477</b>	<b>22%</b>

Legenda: Relação de projetos bem-sucedidos, separados por categoria, apresentados ao Catarse, de 2011 a 2018  
 Fonte: Elaboração do autor

## 2.4 Modalidades de financiamento coletivo

Diferente de outros intermediários financeiros tradicionais, sites de *crowdfunding* não emprestam, coletam ou pegam emprestado dinheiro de outras instituições ou atores financeiros. De acordo com Brüntje e Gajda (2016), seu foco é fazer com que empreendedores e financiadores se encontrem através do fornecimento de informações a respeito dos projetos, das funcionalidades, dos serviços ofertados e dos riscos envolvidos. Em outras palavras, os sites de financiamento coletivo são, ao mesmo tempo, plataformas nas quais as campanhas de funcionamento ocorrem e meios de transação entre empreendedores e investidores.

Existem diferentes modalidades de financiamento coletivo, que variam de acordo com o tipo de retorno oferecido aos financiadores. Segundo Best et al. (2013), os diferentes modelos de financiamento coletivo são:

- a) **Com base em doações (ou *donation-based crowdfunding*)**: de acordo com Best et al. (2013), são iniciativas que demandam geralmente valores relativamente pequenos, inferiores a US\$ 50 mil, e que são mais relacionados a conceitos e ideias ainda não testadas. São considerados neste grupo os projetos que envolvem filantropia, lançamento de novos produtos e serviços, ou viabilização de projetos artísticos e culturais. Não oferecem como contrapartida lucro ou participação no capital dos empreendimentos. Geralmente, o resultado é um protótipo, uma linha de produtos, um evento ou uma prova de conceito.
- b) **Com base em investimento (ou *equity-based crowdfunding*)**: referem-se ao financiamento de novos empreendimentos através da venda de instrumentos financeiros, tais como títulos

de dívida, propriedade intelectual ou ações. Em geral, participam empresas recém-inauguradas, que já estejam funcionando ou tenham ideias já testadas, mas que não são tão atraentes para os mecanismos tradicionais de financiamento dados os riscos. Essas empresas também não conseguem se financiar através de instituições e mecanismos tradicionais pelos altos custos envolvidos. Neste grupo, investidores fazem pequenas contribuições, ao contrário de modelos mais tradicionais como investidores-anjo ou *venture capitalists*.

#### Quadro 1 - Doação

<i>Doação simples:</i>	<ul style="list-style-type: none"> <li>- <b>Características:</b> Filantrópico: investidores doam sem expectativa de compensação monetária</li> <li>- <b>Prós:</b> Não há risco</li> <li>- <b>Contras:</b> Financiadores não adquirem títulos. Empreendedores têm dificuldade de obter maior volume de capital</li> </ul>
<i>Doação contra recompensa:</i>	<ul style="list-style-type: none"> <li>- <b>Características:</b> Investidores recebem um presente ou compram antecipadamente um produto ou serviço. Esse modelo está se desenvolvendo como um <i>marketplace</i>, com empresas arrecadando somas consideráveis em venda antecipada.</li> <li>- <b>Prós:</b> Baixo risco (principalmente, dificuldade em entregar produtos e fraudes). Não há potencial de retorno financeiro.</li> <li>- <b>Contras:</b> Retorno potencial é pequeno. Não há título adquirido e nem mecanismos de <i>accountability</i>. Muitos empreendedores podem ter dificuldade em levantar maior volume de capital sem um produto de apelo de massa.</li> </ul>

#### Quadro 2 - Investimento

<i>Baseado em equity (participação no capital):</i>	<ul style="list-style-type: none"> <li>- <b>Características:</b> Investidores recebem instrumentos de <i>equity</i> ou compartilhamento de lucros</li> <li>- <b>Pros:</b> Potencial participação nos lucros do empreendimento. Potencial ilimitado de ganhos financeiros. Pode atrair um número relativamente grande de investidores</li> <li>- <b>Contras:</b> Risco potencial de fracasso do empreendimento. Os detentores de participação estão subordinados aos credores em caso de falência. As leis que regulam títulos relacionados a financiamento coletivo podem ser complexas</li> </ul>
<i>Baseado em dívida</i>	<ul style="list-style-type: none"> <li>- <b>Características:</b> Investidores recebem um instrumento de débito que paga uma taxa fixa de juros, com retorno principal em data fixada</li> <li>- <b>Prós:</b> Taxa de retorno pré-estabelecida, acordada entre empreendedor e credor. Detentores de título de crédito são seniores em relação aos detentores de capital em caso de falência. A securitização pode tornar mais fácil para empreendedores levantar capital.</li> <li>- <b>Contras:</b> Pode ser subordinado a credores seniores. Tem uma taxa de risco de perda do investimento similar à a alta taxa de insucesso de <i>startups</i>. Requer um negócio que já esteja gerando um fluxo de caixa. Empresas que existem e que tenham fluxo de caixa positivo podem considerar esta opção por terem oportunidades de saída mais bem estruturadas do que as ofertas típicas.</li> </ul>
<i>Baseado em Royalties:</i>	<ul style="list-style-type: none"> <li>- <b>Características:</b> Menos comum que os outros modelos. Investidores recebem participação em um <i>trust</i>, que adquire os lucros dos royalties da propriedade intelectual da empresa financiada. Um percentual do lucro é pago a cada período de tempo. Esse pagamento depende do faturamento ao longo do tempo</li> <li>- <b>Prós:</b> Ganho potencial ilimitado, mas a taxa de ganho é pré-determinada pela taxa de lucro. Apresenta risco e potencial de retornos menores do que o investimento em <i>equity</i>, e é mais arriscado do que o baseado em dívida</li> <li>- <b>Contras:</b> Potencial perda do investimento. Risco de prejuízo comparável ao do baseado em <i>equity</i>, mas investimento oferece retorno potencial inferior. A empresa pode deixar de pagar royalties se escolher operar sem o capital intelectual em questão. Esses instrumentos tendem a atrair menos grupos de investidores do que os outros modelos, logo os empreendedores podem ter mais dificuldade de atrair capital desta forma.</li> </ul>

Neste trabalho, nos interessamos pelo modelo baseado em doação, o mais comum no Brasil, apesar de considerarmos que várias das nossas análises servem para compreender o fenômeno de forma mais ampla.

O Catarse, especificamente, não aceita projetos baseados em *equity*, dívida ou empréstimos, somente em doação simples ou contra recompensas, esta última a modalidade mais recomendada. A plataforma não permite, ainda, doações a campanhas ou partidos políticos, uma modalidade recentemente autorizada pela reforma eleitoral de 2017 (CATARSE, 2018a; CÂMARA., 2018).

## 2.5 Processo de criação de um projeto e seus elementos

O processo de elaboração e financiamento dos projetos baseados em doação acontece de modo muito similar em todas as plataformas de *crowdfunding*. Basicamente, projetos são chamadas públicas por recursos financeiros com vistas à realização de algo. Um empreendedor, que o Catarse chama de realizador e o Kickstarter chama de *creator*, define os tipos de contribuição financeira possíveis e os prêmios aos investidores, que variam de acordo com o valor doado. Analisando os sites europeus, Brüntje e Gajda (2016) afirmam que as contribuições podem variar de praticamente nada, comuns em projetos de caridade, a até €1 mil. Em nosso levantamento, encontramos pedidos no Catarse variando de R\$ 5 a R\$ 25 mil.

Um elemento central na definição das campanhas é a escolha de modalidades dentro das plataformas. Tanto o Kickstarter como o Catarse têm como modalidade principal o tudo-ou-nada, ou seja, caso o empreendedor consiga arrecadar o valor mínimo que definiu ao início do projeto, recebe todo o capital recebido, inclusive aquele que exceder à sua meta. Caso contrário, não recebe nada, nem paga a comissão que seria devida ao site. Os investidores, caso o projeto fracasse, recebem de volta suas doações.

O princípio básico, de acordo com Brüntje e Gajda (2016), é que determinados projetos possuem custos mínimos para sua realização. Assim, o tudo-ou-nada garante que um projeto arriscado não terá de ser realizado caso não tenha capital suficiente.

A segunda modalidade é o *keep-it-all*, que o Catarse chama de “Campanha Flexível”. Neste caso, o empreendedor recebe tudo o que for doado, por menor que seja o valor, ainda que o projeto não tenha atingido a meta estabelecida.

É uma forma adotada mais comumente, segundo Brüntje e Gajda (2016), em projetos de filantropia ou por aqueles empreendedores que utilizam o financiamento coletivo como forma de financiamento complementar. Foi a modalidade escolhida, por exemplo, pela

Deli&Co, uma empresa especializada em painéis de ferro fundido, ao lançar um produto em conjunto com a chef Roberta Sudbrack (CATARSE, 2017d).

A terceira forma de se arrecadar fundos é a doação recorrente. Ao invés de realizar uma chamada por recursos única, com tempo determinado, o empreendedor pode receber contribuições mensais fixas de seus investidores. O Kickstarter não trabalha com essa modalidade, mas o Catarse a lançou em dezembro de 2017, seguindo o exemplo de outras plataformas como o Benfeitoria e o Patreon (CATARSE, 2017e). Como essa modalidade é mais recente, não a analisaremos neste estudo.

Há vantagens e desvantagens em cada uma das modalidades. O tudo-ou-nada garante que um projeto arriscado não terá de ser realizado caso não consiga os recursos mínimos necessários para sua execução, o que pode acontecer em uma campanha flexível. Já a flexível, apesar de permitir que o empreendedor fique com todos os recursos, o que é positivo para quem já tem capacidade de produção e distribuição instalada e testada, não evita o dilema do *free rider*<sup>10</sup>, frequente em situações como o *crowdfunding*, quando os investidores têm incentivo a esperar até o último minuto para participar e, com isso, o projeto corre o risco de não ser financiado (AGRAWAL; CATALINI; GOLDFARB, 2013).

Nenhuma das modalidades, no entanto, impede que um projeto seja financiado muito além do projeto original, o que é muito bom para um realizador bem estruturado, mas pode causar problemas para alguém que tenha limitações de capacidade produtiva, o que frequentemente ocorre. Como veremos, quase todos os projetos bem-sucedidos descumprem seus cronogramas.

Além de definir o modelo de arrecadação, o empreendedor precisa escolher o que o investidor receberá em retorno por sua contribuição. Essas recompensas podem, em geral, ser, segundo Gierczak et al. (2016):

- a) *Nada*. O financiador faz uma doação sem nenhum tipo de retorno. É mais comum em projetos ligados a filantropia, mas pode ser encontrado também em empreendimentos com fins lucrativos. Nestes casos, a doação pode ter a ver com a noção de compartilhamento social.
- b) *Reconhecimento não material*. São os casos em que o empreendedor reconhece a contribuição, mas não entrega nenhuma recompensa física pela ajuda. É o caso de inserir o nome do financiador nos créditos de uma obra, enviar um e-mail de agradecimento ou um telefonema, um encontro com o empreendedor, etc.

<sup>10</sup> Essa questão será apresentada mais à frente, mas, em resumo, o dilema do free rider diz respeito à apropriação individual de esforços coletivos, quando um indivíduo, ou um grupo pequeno, faz um esforço consideravelmente menor que o resto da coletividade para ter acesso aos mesmos benefícios de todos. A expressão tem origem nos passageiros sem ticket de transportes coletivos.

- c) *Produtos*. Podem ser brinquedos, livros, aparatos eletrônicos, entrada para a uma atividade artística, etc.
- d) *Produtos financeiros*, como participação acionária <sup>11</sup>, títulos de dívida, *royalties*, etc. Como o Catarse não permite recompensas financeiras ou sob forma de participação acionária, não iremos estudar esse tipo de recompensa neste momento.

Outra definição, ainda na fase de planejamento, é a duração da campanha. No Catarse, o máximo que uma campanha pode ficar no ar são 60 dias corridos na modalidade Tudo ou Nada, e 365 dias corridos na Flexível. Essa definição é importante porque campanhas de financiamento coletivo tendem a receber mais recursos no início ou ao final do processo de captação (AGRAWAL; CATALINI; GOLDFARB, 2013; KUPPUSWAMY; BAYUS, 2015). Esse comportamento pode ter relação com dois fenômenos que serão abordados mais à frente, a influência de familiares e amigos no resultado das campanhas (estes tendem a contribuir mais no início) e o efeito manada (financiadores tendem a apoiar projetos que já se mostram viáveis). O próprio Catarse produziu um *e-book* (CATARSE, 2018b) com recomendações e exemplos de melhores práticas na definição do prazo de campanha. Segundo afirmam, as campanhas bem-sucedidas tendem a serem financiadas entre 30 e 45 dias, sendo que as contribuições se concentram nas primeira e última semanas.

## 2.6 Formas de comunicação entre empreendedores e investidores

A partir dessas definições, o empreendedor inscreve seu projeto na plataforma e, quando estiver pronto, o disponibiliza ao público através de um *hotsite* (figura 3), que contém todas as informações relevantes a seu respeito, tais como a modalidade, a categoria, os objetivos (inclusive financeiros) que deseja alcançar, um cronograma de execução e quais recompensas oferece.

É neste *hotsite* que acontece o contato mais importante entre empreendedores e financiadores e é nele que as contribuições são feitas. Os elementos utilizados para descrever os projetos são:

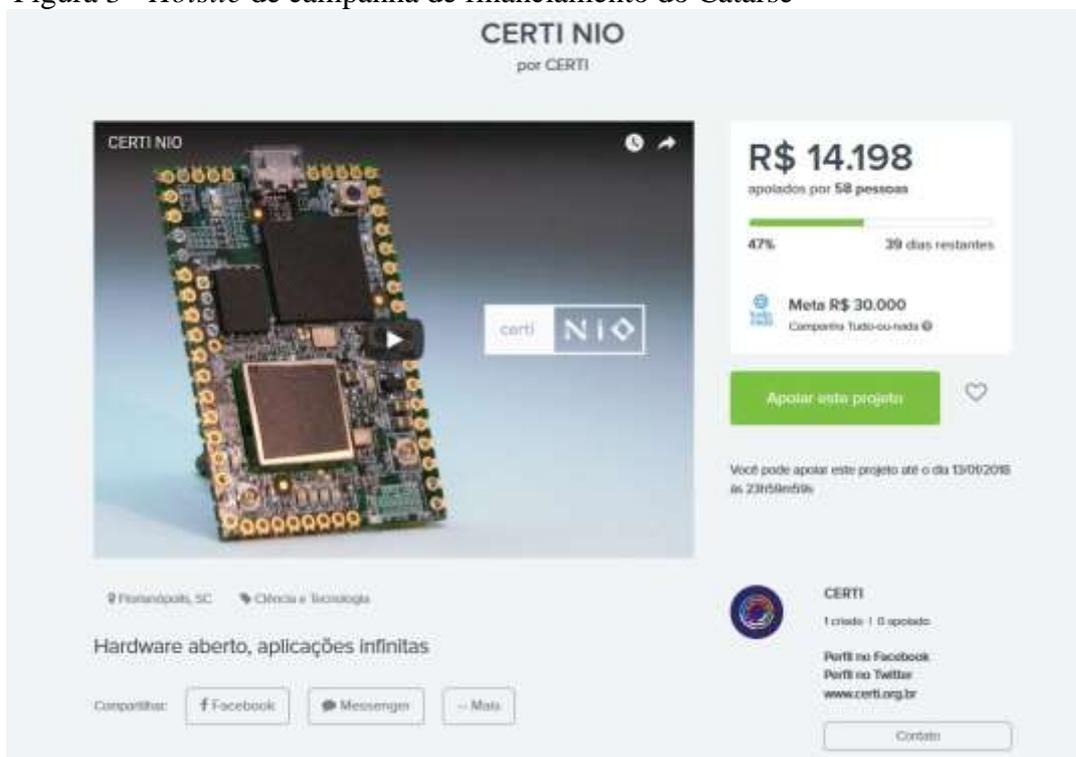
- a) Título da campanha;
- b) Categoria;

---

<sup>11</sup> Algumas plataformas que oferecem, em agosto de 2020, participação acionária em empresas, ou *equity*, como recompensa, são, nos Estados Unidos, o AngelList (<https://angel.co>) e o CircleUp (<https://circleup.com>). No Brasil, encontramos duas iniciativas desta natureza, o Startmeup (<https://www.startmeup.com.br>) e o eqseed (<https://eqseed.com/br>).

- c) Local do projeto;
- d) Tipo de financiamento;
- e) Meta de arrecadação;
- f) Tempo de arrecadação ;
- g) Descrição ;
- h) Vídeo ;
- i) Orçamento;

Figura 3 - *Hotsite* de campanha de financiamento do Catarse



Legenda: Exemplo de *hotsite* utilizado para reunir as informações a respeito das campanhas de financiamento do Catarse, disponível em <https://www.catarse.me/certinio>, acessado em 5 de dezembro de 2017 Fonte: Acervo do autor

- a) Recompensas;
- b) Perfil do realizador, incluindo seus sites e perfis em sites de redes sociais.

Além disso, ficam visíveis para quem acessar as páginas:

- a) O percentual de alcance da meta estabelecida;
- b) O número de apoiadores;
- c) Os comentários feitos por outros apoiadores.

Alguns desses elementos nos interessam em especial por indicarem o grau de preparo dos empreendedores, a qualidade dos projetos e fomentarem o envolvimento social em torno da campanha. Abordaremos esses estudos mais à frente.

Figura 4 - Interação através dos comentários dos projetos de financiamento coletivo



Legenda: Exemplo de interação através dos comentários na plataforma Catarse, disponível em <https://suporte.catarse.me/hc/pt-br/articles/115002214463-projeto-ONLINE>, acessado em 25 de julho de 2017

Fonte: Acervo do autor

Além dos elementos de descrição, os empreendedores podem, ainda dentro da plataforma, publicar novidades a respeito da campanha, enviar atualizações a respeito do projeto aos e-mails dos apoiadores, atualizar a descrição e responder aos comentários feitos (figura 4).

O Catarse utiliza, ainda, seus perfis nos sites de redes sociais (Facebook e Twitter), seu blog, a *home-page* do seu site institucional e uma *newsletter*, distribuída por e-mail, para dar visibilidade a projetos que sua equipe julga como sendo mais relevantes. Segundo afirmam, a divulgação por parte deles não é um serviço estendido a todos os empreendedores, mas uma escolha institucional e depende, segundo eles, “do desempenho da campanha e das escolhas editoriais da nossa equipe” (CATARSE, 2018c).

Além das ferramentas disponíveis na plataforma, alguns realizadores desenvolvem campanhas mais elaboradas de comunicação, envolvendo promoção em sites de redes sociais, eventos, links patrocinados e assessoria de imprensa (MARTINS; MARTINS, 2016). Esse tipo de esforço é recomendado pelo Catarse, já que aumenta a visibilidade do projeto e com isso, as chances de que ele seja financiado. Ao mesmo tempo, aumenta também a visibilidade da própria plataforma.

## 2.7 Incentivos e oportunidades do financiamento coletivo

Um dos questionamentos feitos por Agrawal, Catalini e Goldfarb (2013) tem a ver com a lógica dos incentivos. O que motiva criadores, financiadores e plataformas ao se envolverem com financiamento coletivo? Para os autores, empreendedores buscam o *crowdfunding* por dois motivos: redução do custo de capital e acesso a mais informação.

No que diz respeito ao custo de capital, tradicionalmente, empreendedores com projetos em estágio embrionário buscariam financiamento para suas ideias através de empréstimos pessoais, poupança individual, dívida em cartão de crédito, empréstimos de amigos e familiares, investidores-anjo ou *venture capitalists*<sup>12</sup>. Logo, em determinadas condições, o *crowdfunding* pode representar o acesso a capital de mais baixo custo<sup>13</sup>.

Outro incentivo para empreendedores é a facilidade em se acessar informações que não teria de outra forma, de duas maneiras. A primeira, tem a ver com a relação com potenciais financiadores, que podem sugerir modificações e melhorias nos projetos ao longo do processo de financiamento, o que Agrawal, Catalini e Goldfarb (2013) chamaram de *user-driven innovation*. A segunda, com a possibilidade de se testar protótipos e conceitos antes do lançamento de produtos e serviços no mercado.

Como contam Agrawal, Catalini e Goldfarb (2013), Eric Migicovsky, o empreendedor por detrás do Pebble, aproveitou-se dos dois mecanismos. O Pebble é um dos primeiros *smartwatches* e foi lançado em uma campanha no Kickstarter que conseguiu US\$ 10 milhões. Eric, antes de colocar seu projeto na plataforma, tinha apresentado sua ideia a investidores tradicionais, que negaram os recursos necessários por duvidarem do potencial de mercado de um produto como aquele. Mesmo com a negativa inicial, Eric resolveu testar o conceito usando a plataforma do Kickstarter, com sucesso. Ao mesmo tempo, o empreendedor recebeu, ao longo da campanha, uma série de sugestões da comunidade de financiadores, algumas incorporadas como funcionalidades, tais como incluir Bluetooth 4.0, uma tecnologia que usa menos bateria do que o Bluetooth 2.1, definida originalmente.

---

<sup>12</sup> *Venture Capitalists* são investidores de risco. Geralmente, são os primeiros a financiar um empreendimento buscando rendimentos superiores por conta do risco que correm.

<sup>13</sup> Essa é uma questão que nos parece necessitar uma investigação posterior. Agrawal, Catalini e Goldfarb (2013) afirmam que, apesar deste ser um argumento frequentemente utilizado, se somadas as comissões e os eventuais custos de *due diligence*, o custo de captar recursos no financiamento coletivo pode ser superior a outras formas de financiamento. Isso pode se tornar ainda mais relevante em países como o Brasil, conhecido por ter acesso restrito a financiamento de novos negócios. Além disso, frequentemente realizadores têm custos não previstos como a preparação e gestão da campanha, e custos extras de produção em caso de procura excessiva, o que reduz a lucratividade de um projeto. Nos parece, intuitivamente, que um dos principais incentivos, em países como o nosso, é a pouca formalização desta forma de financiamento, ou seja, o fato de existirem poucas barreiras formais para a proposição dos projetos.

Quem contribui com campanhas de financiamento coletivo, de maneira geral, parece motivado por questões sociais. São financiadores que apoiam projetos com os quais têm relação emocional, são realizados por empreendedores próximos, familiares ou amigos, ou com os quais conseguem criar identificação. Agrawal, Catalini e Goldfarb (2013), Carr (2013), Gierczak et al. (2016) acreditam que os apoiadores têm por incentivo:

- a) *Acesso antecipado a lançamentos de produtos.* Os apoiadores desconhecidos são, muitas vezes, entusiastas por acesso a novos produtos, ainda que venham de produtores sem reputação estabelecida ou que sejam protótipos.
- b) *Participação comunitária.* Para muitos, o *crowdfunding* é uma atividade social, que pode ter como resultado o acesso privilegiado ao criador, algum tipo de reconhecimento social pelo esforço conjunto ou o sentimento de participação. Por esse motivo, financiadores se sentem confortáveis em não só fornecer os recursos necessários para viabilizar o empreendimento, mas também em sugerir melhorias.
- c) *Filantropia.* Ainda que grande parte dos projetos de financiamento coletivo tenha como objetivo o lucro, diversas iniciativas não oferecem recompensas e ainda assim são viabilizadas. A justificativa é que existe uma parte dos apoiadores que se envolvem nesses projetos por filantropia e altruísmo.
- d) *Formalização de contratos.* Em geral, empreendedores com projetos embrionários buscam capital entre amigos e familiares. A utilização das plataformas de financiamento coletivo é uma maneira de formalizar acordos que seriam realizados em privado, o que muda a dinâmica social das pressões pela realização dos projetos.

Já as plataformas, como o Kickstarter e o Catarse, são predominantemente negócios com fins lucrativos, cujo modelo básico é de cobrança de taxa por projeto bem-sucedido. Para que funcione, precisa atrair uma comunidade grande de apoiadores e de realizadores, principalmente daqueles que tenham projetos de qualidade, com potencial de sucesso, e que consigam atrair muita atenção midiática, já que esse tipo de projeto não somente expande a comunidade de financiadores, mas permite que a plataforma avance em novas categorias (GIERCZAK et al., 2016).

Por isso, tanto o Catarse como Kickstarter oferecem *e-books*, tutoriais e videoaulas sobre planejamento de campanha, orçamento básico, como definir prazos, como contar histórias cativantes, exemplos de boas práticas, dentre outros assuntos, para garantir que os projetos tenham um mínimo de qualidade.

Essas plataformas precisam facilitar, ainda, o encontro entre apoiadores e projetos de qualidade de seu interesse. Por isso existem algumas ferramentas de promoção de projetos, tais

como *newsletter*, blog, página inicial e assessoria de imprensa, utilizadas exclusivamente pelas plataformas, o que aumenta a visibilidade de projetos considerados estratégicos.

Outra preocupação, de ambas, é evitar a percepção de que suas plataformas são lojas virtuais, o que traria uma maior responsabilidade na curadoria dos projetos, no controle das entregas e na qualidade dos produtos, etc., o que nenhuma das duas garante. Segundo o KICKSTARTER (2017), a responsabilidade por completar um projeto é inteiramente de seu proponente e a plataforma não dá garantias, não gerencia os projetos, nem investiga a capacidade de um empreendedor de completar suas propostas. Da mesma forma, o Catarse, em sua seção de dúvidas mais comuns, afirma que é responsabilidade do apoiador “averiguar a execução e cobrar diretamente do realizador caso a sua expectativa não seja correspondida” CATARSE (2017f) e que a relação entre as partes é prevista no artigo 854, do Código Civil, que traz a seguinte redação:

Aquele que, por anúncios públicos, se comprometer a recompensar, ou gratificar, a quem preencha certa condição, ou desempenhe certo serviço, contrai obrigação de cumprir o prometido (BRASIL, 2002).

A respeito de curadoria, é interessante notar que, ainda que as plataformas tenham o incentivo de oferecer projetos de qualidade, nenhuma das duas escolhe, efetivamente, quais projetos são publicados. O Catarse limita-se a impedir que projetos com determinados temas sejam desenvolvidos, como aqueles que envolvem pornografia, pedofilia, violência explícita, atos ilegais etc. O Kickstarter também tem uma lista de 96 recompensas proibidas (KICKSTARTER, 2018a), que vão desde bebidas energéticas a armas, mas, em projetos que envolvem hardware e produção industrial, exigem que o empreendedor tenha um protótipo funcional e que seja apresentado um projeto de produção (KICKSTARTER, 2018b).

Outro interesse das plataformas é de garantir o menor número de fraudes possível, já que esse é um dos principais riscos para os financiadores. No entanto, ambas limitam-se a pedir que a comunidade aponte aqueles projetos ou empreendedores que não estejam agindo corretamente.

## 2.8 Riscos do financiamento coletivo

Em toda relação de financiamento há risco, devidos, em grande parte, à assimetria de informações entre financiadores e empreendedores.

Empreendedores quase sempre têm uma vantagem na quantidade e na qualidade das informações a respeito dos riscos de seu próprio empreendimento, principalmente em projetos

em estágio embrionário. Essa assimetria é maior em ambientes pouco regulados e em situações nas quais é caro demais fazer uma *due diligence*<sup>14</sup> aprofundada (AGRAWAL; CATALINI; GOLDFARB, 2013; BRÜNTJE; GAJDA, 2016).

E esse é exatamente o cenário do *crowdfunding*. É muito difícil que um financiador entre em contato pessoalmente com o empreendedor para que perceba o quão preparado ele está ou quais qualidades ocultas existem nos projetos. Do ponto de vista do empreendedor, esse é um risco porque pode fazer com que projetos de alta qualidade não sejam financiados. Para financiadores, pode esconder falhas graves de planejamento, falta de competência para a realização do que for proposto ou mesmo fraudes.

A principal maneira através da qual se tenta superar as assimetrias de informação no *crowdfunding*, para Agrawal, Catalini e Goldfarb (2013), é o que chamaram de *crowd due diligence*, ou seja, a verificação pela comunidade das credenciais dos criadores e da qualidade dos projetos. Os autores dão como exemplo um projeto de jogo eletrônico no Kickstarter que falhou após dois investidores potenciais o denunciarem como fraudulento. Esses investidores identificaram plágio e uso de imagens indevidas na campanha, e alertaram a comunidade, que parou de fazer contribuições. Em seguida, os administradores da plataforma retiraram o projeto do ar.

Vários sites e *marketplaces* digitais, como o eBay e a Wikipedia, dependem do monitoramento de seus usuários para detectar fraudes. A Wikipedia, por exemplo, não conseguiria eliminar sozinho fraudes e vandalismo por conta do tamanho que alcançou (AGRAWAL; CATALINI; GOLDFARB, 2013).

Outra incerteza, no caso do financiamento coletivo, tem a ver com o efeito manada. Como vimos, a comunidade tende a apoiar projetos que já tenham sucesso ou que já estejam próximos de atingir as suas metas, ou seja, investidores percebem a aprovação da comunidade, refletida no volume de contribuições, como um sinal de qualidade relevante (AGRAWAL; CATALINI; GOLDFARB, 2013; KUPPUSWAMY; BAYUS, 2015).

Esse tipo de comportamento pode ser eficiente em alguns casos, mas não em todos, sobretudo pelo risco de manipulação. Quando o criador tem uma rede social própria grande, esse investimento inicial pode não ter como base a racionalidade do investimento, ou seja, não se baseia na qualidade do projeto, mas nos laços sociais do criador. Da mesma forma, família

---

<sup>14</sup> *Due Diligence* é uma expressão que se refere a uma auditoria ou uma investigação que acumula a maior quantidade de informações possível a respeito dos aspectos pessoais, técnicos, financeiros e contextuais envolvendo os empreendimentos para que se avalie melhor riscos financeiros, legais operacionais, etc. antes da realização de um processo de financiamento.

e amigos tendem a ser os primeiros a investir, o que pode não ter relação, mais uma vez, com a qualidade ou viabilidade do projeto.

Outra questão relacionada à dinâmica do financiamento, segundo Agrawal, Catalini e Goldfarb (2013), tem a ver com o dilema do *free rider*, também conhecido como clandestinidade ou parasitismo. O *free rider* surge naquelas situações nas quais alguém pode se aproveitar de bens comuns sem assumir os custos da existência daqueles bens, ou seja, ele parasita os esforços comunitários.

É um comportamento pouco desejável porque limita os benefícios sociais das ações coletivas (FARIETA, 2015).

No caso do *crowdfunding*, o comportamento do *free rider* aparece porque os apoiadores têm um incentivo para esperar o que os outros vão antes de tomar a decisão de ajudar ou não o projeto e, assim, reduzir o seu risco individual, ou seja, o doador só coloca dinheiro nos projetos que sabem que vão dar certo. O que as plataformas fazem para reduzir esse problema, segundo Agrawal, Catalini e Goldfarb (2013), é somente disponibilizar os fundos de um projeto caso este alcance a meta demandada, o que elimina o risco, para os apoiadores, de investir em um projeto que não será realizado.

Por outro lado, há maneiras legítimas para que os criadores se aproveitem destes efeitos, como, por exemplo, oferecer recompensas mais generosas aos primeiros que investirem, o que no Kickstarter é chamado de *early bird perks*. Geralmente, essas ofertas precoces vêm com descontos melhores ou pacotes mais generosos de produtos.

O realizador pode também, à medida em que o projeto arrecada mais dinheiro, lançar novas metas com novas recompensas, o que cria um incentivo para que apoiadores iniciais contribuam cedo, o que lhes garante um custo menor por recompensa, e para que auxiliem no processo de divulgação, o que os fará receber outras recompensas não previstas. Esse comportamento aponta, por fim, para a importância de um planejamento de comunicação anterior ao lançamento do projeto, para que se crie uma comunidade de entusiastas dispostos a contribuir desde o início.

De modo geral, o principal risco para quem participa de financiamento coletivo é o do fracasso do empreendimento. Como as plataformas não fazem uma análise mais detalhada da capacidade gerencial de quem propõe os projetos, há sempre o risco de falência das empresas que estão promovendo os projetos, ou de subdimensionamento das capacidades produtivas ou dos desafios tecnológicos que deverão ser enfrentados para transformar um protótipo em um bem de consumo produzido em larga escala (BRÜNTJE; GAJDA, 2016; BEST et al., 2013).

O que acontece nesses casos é que os apoiadores não recebem os retornos como especificados inicialmente ou não os recebem inteiramente, apesar de Mollick (2015) afirmar que isso não acontece com muita frequência no Kickstarter. Em sua pesquisa, com uma amostra de pouco mais de 47 mil projetos concluídos, somente em 9% deles as recompensas não foram entregues.

A outra consequência, essa sim muito comum, é o atraso no desenvolvimento das recompensas e, conseqüentemente, de suas entregas. Mollick (2014) afirma, após analisar 48.500 projetos no Kickstarter, que somente 24,9% dos bem-sucedidos realizaram suas entregas nos prazos prometidos. Já Agrawal, Catalini e Goldfarb (2013) afirmam que 84% dos 50 maiores projetos do Kickstarter, no período em que realizaram seu estudo, atrasaram a entrega.

Para Brüntje e Gajda (2016), isso pode impactar negativamente a reputação dos empreendedores e reduzir suas margens, dados os custos não previstos. Para Agrawal, Catalini e Goldfarb (2013), financiadores têm sido muito otimistas com relação à habilidade dos criadores em entregar suas promessas. No entanto, isso pode mudar se cada vez mais projetos deixarem de entregar as recompensas definidas ou atrasaram demais.

O outro risco importante é a fraude, ou seja, quando o proponente de um projeto deliberadamente tenta enganar os investidores. Apesar de existirem alguns casos conhecidos (GIZMODO, 2015), Best et al. (2013) afirmam não ter encontrado nenhum bem-sucedido. De acordo com seu relatório, por conta da transparência inerente a essas plataformas, em geral, as tentativas de fraude são identificadas em menos de 24 horas.

Por conta desses riscos, órgãos regulatórios em todo o mundo estabeleceram regras para o funcionamento dessas plataformas. Segundo Dresner (2014), a *Securities and Exchange Commission* (SEC, ou Comissão de Títulos e Câmbio dos Estados Unidos) divulgou, em 2012, o segundo título do JOBS Act, que permite que empreendedores usem meios públicos, como a internet ou qualquer outro meio de comunicação, para solicitar financiamento para seus negócios desde que os investidores sejam acreditados. Ao mesmo tempo, o Reino Unido e a Itália estabeleceram regras para o funcionamento de financiamento coletivo, antes que a União Europeia se pronunciasse a respeito.

No Brasil, a Comissão de Valores Mobiliários (CVM) estabeleceu, em 13 de julho de 2017, a instrução 588 (CVM, 2017), que “dispõe sobre a oferta pública de distribuição de valores mobiliários de emissão de sociedades empresárias de pequeno porte realizada com dispensa de registro por meio de plataforma eletrônica de investimento participativo”, ou seja, autoriza o funcionamento de sites de *crowdfunding* sob condições específicas. Além da instrução da CVM, a resolução nº 23.553, de 2 de fevereiro de 2018, do Tribunal Superior

Eleitoral (TSE), regulamenta o uso de financiamento coletivo como modalidade de arrecadação de recursos para campanhas eleitorais, também sob condições específicas (TSE, 2018).

Agrawal, Catalini e Goldfarb (2013) analisaram, ainda, os desincentivos, ou seja, elementos que podem afastar os atores do processo de financiamento coletivo. Para os criadores, em projetos que não envolvam *equity*, o maior desafio é o relacionamento com financiadores. Diferentes de investidores tradicionais, os que utilizam essas plataformas são numerosos e demandam um nível de atenção alto, sobretudo quando os projetos frustram os prazos estabelecidos. Ainda que essa interação possa ser rica do ponto de vista do desenvolvimento dos projetos, alimentá-la significa desviar recursos (tempo, atenção, pessoal, etc.) da execução.

## 2.9 Sinais de qualidade e confiança

Por todas essas questões, Agrawal, Catalini e Goldfarb (2013) apontam para a importância da reputação como mecanismo para estabelecimento de confiança e para a redução de riscos. Para os autores, criadores oferecem informações a respeito do que pretendem fazer e pedem aos apoiadores que confiem neles seus recursos. Como não é possível realizar uma *due diligence* ou reunir presencialmente criadores e apoiadores para reuniões, plataformas de financiamento coletivo criaram outros mecanismos para que a confiança se estabeleça:

- a) *Sinais de qualidade*: plataformas demandam dos criadores que ofereçam evidências da qualidade de seus projetos. Isso pode acontecer através das suas marcas, da divulgação de dados pessoais, do estabelecimento de garantias, da apresentação de patentes, da exposição da experiência dos empreendedores ou de sua equipe, do nível educacional (se pesquisadores, doutores, ou se possuem MBAs ou especializações);
- b) *Sistemas de feedback*: muitos *marketplaces* online possuem sistemas de avaliação abertos a seus usuários, geralmente com o estabelecimento de notas ou de recomendações. Esse tipo de sinalização é importante para a construção de uma reputação online e facilita a relação comercial entre desconhecidos. Em plataformas de financiamento coletivo esses mecanismos são menos comuns porque a frequência na qual um criador busca financiamento é baixa, diminuindo as oportunidades de avaliação de um criador pela comunidade;
- c) *Intermediários confiáveis*: além dos sinais de qualidade fornecidos pelo criador, é possível utilizar terceiros para reforçar as características positivas de um projeto,

como selos de qualidade de agências certificadoras, recomendações de outras organizações ou mesmo indicação de *reviews* e matérias publicadas pela imprensa.

Como discutimos anteriormente, o que difere o financiamento coletivo de outras modalidades é justamente o fato de se apresentar como uma alternativa em situações nas quais os incentivos econômicos tradicionais não estão presentes ou não são facilmente percebidos. Em casos de projetos com baixo retorno financeiro, filantropia, eventos de curta duração, iniciativas culturais, projetos pessoais, por exemplo, investidores tradicionais tendem a não colocar recursos e, conseqüentemente, estes não seriam realizados, o que dá ao *crowdfunding* um papel relevante (CARR, 2013; MOLLICK, 2014; KUPPUSWAMY; BAYUS, 2015).

Ao mesmo tempo, há literatura que mostra que o comportamento dos apoiadores de campanhas de *crowdfunding* é, de várias maneiras, similar ao comportamento de investidores tradicionais no que diz respeito à maneira como decidem entre os projetos e como avaliam o risco de participar deles.

Em especial, há literatura que aponta dinâmicas similares às do *crowdfunding* na relação entre empreendedores e *venture capitalists*, também conhecidos como investidores de risco. Landström (2007) considera como sendo investimentos de risco aqueles realizados por instituições, firmas ou indivíduos com muita disponibilidade financeira em empreendimentos que não estão listados nas bolsas de valores e que tenham um grande potencial de crescimento. Esse tipo de investidor aceita assumir riscos maiores, como o lançamento de novos empreendimentos, por compreenderem que há um potencial de ganho ainda maior do que em investimentos de outras naturezas.

A existência desse tipo de investidor está associada tanto a exemplos de muito sucesso, como o Vale do Silício, como é conhecida a região no entorno de São Francisco, nos Estados Unidos, onde estão localizados os principais investidores em tecnologia atualmente, como a fracassos históricos, como a quebra da Nasdaq, a bolsa de ações de tecnologia, no final dos anos 1990.

Landström (2007) credita a esses fracassos, em parte, o surgimento de regulamentações em todo mundo a respeito dos limites do investimento de risco e, também, uma mudança nas estruturas e processos de gestão de empresas de *venture capital* a partir dos anos 1980, o que tornou essas empresas muito mais especializadas.

Por isso tudo, o processo de financiamento de novos negócios, sobretudo por investidores de risco, tende a ser longo, complexo e cercado de aspectos técnicos. Isso se dá porque esses investidores necessitam reduzir ao máximo o risco de perder seu capital por conta de um erro ou uma falha não prevista.

Essa investigação, no entanto, não parece suficiente para eliminar a ambiguidade e o risco desses investimentos. Por isso, nesses processos, é necessário que o empreendedor se apresente e defenda seus planos de negócio, dando informações complementares aos investidores.

Para Chen, Yao e Kotha (2009), a tomada de decisão de investimento, nestes casos, é resultante de um processo de persuasão, no qual os empreendedores tentam convencer os investidores da viabilidade de suas ideias. Há, portanto, nessa relação, um conjunto de elementos que não têm a ver necessariamente com as qualidades intrínsecas aos projetos, ou seja, como seus dados objetivos ou análises financeiras, mas com elementos subjetivos, mesmo que investidores afirmem buscar os mesmos indícios em todos os planos de negócio. O estudo de Chen, Yao e Kotha (2009) aponta que, em vários casos, a paixão demonstrada por um empreendedor ao defender o seu projeto resultava na percepção de que ele estaria mais motivado e preparado para executar seu plano do que outros, uma conclusão que corrobora descobertas feitas por outras pesquisas que indicam que, em larga medida, investidores de risco são influenciados pela personalidade e a história do empreendedor, as características da equipe de gestão, a “química pessoal” entre investidores e empreendedores e na própria intuição (o *gut feeling*, no inglês).

Um dos estudos mais conhecidos nesta mesma linha foi MacMillan, Siegel e Narasimha (1985). Nele, os autores perceberam que, mais do que o plano de negócios, no qual o produto, o mercado e os critérios financeiros são descritos, o que determinava a decisão de investimento por parte de um *venture capitalist* era a percepção de qualidade do empreendedor. Essa qualidade percebida apontaria para uma capacidade maior de gerenciar riscos, exercer poder, realizar um esforço maior, de ter mais conhecimento do mercado, ter histórico de desempenho em negócios, mais capacidade de articulação, dentre outros elementos. Em entrevista a Landström (2007), MacMillan afirmou que esses sinais podem funcionar tanto para auxiliar projetos ruins a serem financiados, o que comprovou ao estudar a performance de empreendimentos percebidos pelos investidores inicialmente como de alta qualidade, como para dificultar o lançamento de empreendimentos com alto potencial de ganhos.

Essas percepções a respeito da viabilidade e qualidade dos projetos e dos empreendedores não são técnicas, internas, mas elementos subjetivos, externos, que geram as ideias a respeito da qualidade, do potencial de sucesso ou da necessidade de uma iniciativa sobre as quais investidores se apoiam para ter mais confiança e segurança em suas decisões (MOLLICK, 2014; KUPPUSWAMY; BAYUS, 2015).

## 2.10 Sinais de qualidade e financiamento coletivo

Em nosso levantamento bibliográfico, encontramos diversos trabalhos que analisaram os projetos de financiamento coletivo. A maior parte deles estava interessada em avaliar seu sucesso, ou seja, se as propostas alcançavam ou não as metas estabelecidas pelos realizadores.

Dentre esses estudos, destacamos os de Agrawal, Catalini e Goldfarb (2013), Carr (2013), Mollick (2014), Xu et al. (2014) e Kuppuswamy e Bayus (2015), que analisaram as plataformas Kickstarter, Sellaband e Symbid, e os de Martins e Martins (2016) e Felipe (2017), que se concentraram no Catarse.

Resumidamente, esses trabalhos indicam que os sinais de qualidade estão associados 1) às estratégias desenvolvidas pelos realizadores para o relacionamento com os apoiadores e com a comunidade, o que inclui as formas de divulgação dos projetos e a qualidade dos elementos de comunicação das campanhas; 2) a características dos realizadores, como sua reputação, o tamanho de suas redes pessoais e o engajamento com a comunidade de *crowdfunding*; 3) a algumas características específicas dos projetos; e 4) a efeitos de geografia.

De forma geral, as pesquisas indicam que a maior parte dos projetos não consegue atingir seus objetivos. Agrawal, Catalini e Goldfarb (2013) apontaram que no Sellaband, entre 2006 e 2009, 61% dos projetos não recebia nenhuma doação e 0,7% deles concentrava mais de 73% dos fundos recebidos. Perceberam, também, que, na amostra que analisaram do Kickstarter, 1% das campanhas recebe 36% dos fundos. Da mesma maneira, Mollick (2014) afirma que somente 48,1% dos 48.526 projetos publicados entre 2009 e 2012 atingiram as suas metas.

Ainda a respeito do sucesso, há uma tendência de que os projetos atinjam as metas por pequenas margens, e de que os fracassos sejam por grandes. No estudo de Mollick (2014), 25% dos que deram certo ficaram 3% acima da meta estabelecida e metade ficou cerca de 10% acima da meta. Somente 1 entre 9 conseguiu atingir 200% do pedido original. Por outro lado, a média de captação dos projetos malsucedidos é de 10,3% da meta original. Cerca de 10% captam cerca de 30% e somente 3% chegam perto dos 50%.

Mollick (2014) explica esse comportamento de duas formas. A primeira é que, como o sistema do Kickstarter é exclusivamente de tudo-ou-nada, os realizadores se sentem incentivados a completar do próprio bolso o que falta de um projeto que está quase todo financiado, ainda que o site estabeleça mecanismos para evitar que isso ocorra. O outro diz respeito à natureza dos projetos. Aqueles percebidos como sendo de mais qualidade, tendem a atrair mais financiadores. Nesses casos, os projetos acabam tendo também mais visibilidade

porque esses investidores iniciais se sentem compelidos a divulgar espontaneamente as iniciativas e, em paralelo, acabam atraindo mais mídia, o que lhes dá mais visibilidade.

Agrawal, Catalini e Goldfarb (2013) identificaram, também, uma espécie de efeito-manada. No Sellaband, apoiadores eram duas vezes mais propensos a investir em projetos que já haviam alcançado 80% da meta, em relação àqueles que tinham chegado a somente 20% dela. A aceleração é mais forte ao final da campanha, ainda que exista um efeito de desaceleração no meio, em alguns casos, por conta da percepção de que o projeto será bem-sucedido independentemente do esforço individual.

Agrawal, Catalini e Goldfarb (2013) perceberam, ainda, que criadores e financiadores tendem a ser excessivamente otimistas em relação aos resultados dos projetos, colocando prazos irrealistas ou subdimensionando os desafios operacionais de suas ideias. Como resultado, muitos projetos do Sellaband e do Kickstarter falham, ou seja, recebem os recursos e não conseguem realizar suas ideias, ou têm atrasos significativos.

Esses fracassos podem estar relacionados com uma característica percebida por Kuppuswamy e Bayus (2015), de que as campanhas são financiadas por múltiplas pequenas contribuições (cerca de US\$ 70 por pessoa, em média) e tendem a receber um volume pequeno de recursos no total, algo em torno de US\$ 2,1 mil dólares em média por projeto.

No que se refere às estratégias de relacionamento e divulgação, as pesquisas indicam que realizadores mais bem preparados, que investem mais tempo e esforços para garantir o sucesso das empreitadas têm, conseqüentemente, mais potencial de atrair apoiadores.

Mollick (2014), por exemplo, aponta que projetos com erros gramaticais têm menos 13% de chance de serem financiados; os que não possuem vídeos têm 25% menos chances de sucesso; e aqueles que não enviam atualizações no início têm, também, menos 13% de chance.

Da mesma forma, Xu et al. (2014) e Kuppuswamy e Bayus (2015) encontraram correlação positiva entre número de atualizações (a principal forma de contato entre realizadores e apoiadores) e o sucesso do financiamento.

Segundo Kuppuswamy e Bayus (2015), os *updates* ocorrem mais frequentemente bem no início e bem ao final dos projetos, quando estes estão próximos da meta, ainda que, na média, realizadores enviem poucas atualizações, cerca de duas por projeto.

Xu et al. (2014), por sua vez, notaram que as chances de sucesso de um projeto no Kickstarter sem nenhuma atualização eram de apenas 32%. Ao mesmo tempo, a chance de sucesso aumentava para 58,7% naqueles que tinham ao menos uma. Os autores analisaram, ainda, o conteúdo das atualizações e perceberam que projetos bem-sucedidos incentivavam os apoiadores a divulgar o projeto ainda no início, davam atualizações do progresso no meio e anunciavam novas recompensas ao final da campanha. Xu et al. (2014) concluem que a maneira

como os realizadores comunicavam durante a campanha é um preditor melhor de sucesso do que a forma como a página da campanha é estruturada.

Em um estudo a respeito do Catarse, Martins e Martins (2016) analisaram as estratégias de divulgação dos projetos a partir da ótica da comunicação organizacional em meios digitais e, para isso, acompanharam, entre agosto e outubro de 2014, 24 projetos de sucesso na plataforma. Segundo os autores, todos aqueles financiados apresentaram um histórico, descrevendo como cada um nasceu e pretende ser executado, e uma lista de perguntas mais frequentes. Apesar de não obrigatórios, os projetos bem-sucedidos traziam, em sua maioria, imagens, gráficos e desenhos para auxiliar na sua descrição e todos tinham, ainda, um vídeo explicativo.

Martins e Martins (2016) perceberam, também, que 13 dos 24 projetos fizeram uso do Facebook como plataforma auxiliar de divulgação de seus projetos, sendo que 12 desses projetos utilizaram mais de um perfil para realizar as postagens, que foram de 2 a 47, com média de 146 compartilhamentos cada, o que sugere que o uso de outros sites de redes sociais como forma de aumentar a visibilidade pode ser importante para atrair apoiadores.

A respeito das redes dos realizadores, os autores indicam que o financiamento coletivo é uma prática eminentemente comunitária. Isso significa que realizadores que tenham laços fortes com suas comunidades profissionais ou com os grupos que apoiam projetos desta natureza recorrentemente tendem a atrair mais apoios.

Carr (2013), depois de analisar 71.304 campanhas do Kickstarter entre 2009 e 2013 e entrevistar 15 realizadores da plataforma Symbid, baseada na Holanda, afirmou que, sob condições de incerteza, fatores não financeiros têm mais importância do que incentivos materiais. Para ele, perto de 60% do resultado pode ser explicado por fatores relacionais e sociais, como reputação, tamanho das redes sociais e outras formas de capital social. O autor notou, ainda, que o principal indicador para avaliar o sucesso de um projeto seria a resposta dos apoiadores às atualizações dos projetos, o que, para ele, é indício de que apoiadores se relacionam com as campanhas motivados sobretudo por relacionamentos sociais e pelo que chamou de status de insider que se obtém na participação.

Agrawal, Catalini e Goldfarb (2013), por sua vez, afirmam que famílias e amigos têm um papel fundamental no início das campanhas. Como pessoas próximas ao empreendedor tendem a ser os primeiros a participar, seu comportamento serve como sinal de qualidade precoce para outros investidores.

Já Kuppuswamy e Bayus (2015) perceberam que os apoiadores, em geral, tendem a contribuir para um projeto mais fortemente na primeira e na última semana do ciclo de captação, independentemente de serem projetos de sucesso ou não. Esse comportamento, segundo

afirmam, não está correlacionado com a maneira como o Kickstarter é construído<sup>15</sup>, mas sugerem que a participação dos membros das famílias e amigos dos realizadores tende a acontecer na primeira semana após o lançamento e na última semana de captação. Além disso, para os autores, a maioria dos apoiadores tende a vir do círculo social do empreendedor e costumam contribuir uma única vez somente para um projeto.

Mollick (2014), por sua vez, considera que o tamanho da rede dos realizadores é um elemento que influencia o sucesso de um projeto, considerando que as redes oferecem não somente conexões que podem auxiliar na divulgação, mas que também funcionam como endosso da competência dos empreendedores. Para medir o tamanho dessas redes, Mollick (2014) utilizou o número de conexões que os realizadores possuem em seus perfis no Facebook. Seu estudo mostrou, por exemplo, que, na categoria cinema, um empreendedor com 10 amigos no Facebook tem 9% de sucesso; um com 100 tem 20%; e um com 1000 teria 40% de chances de conseguir ser financiado.

Já sobre as características dos projetos, as pesquisas sugerem que algumas escolhas iniciais dos realizadores, como a meta e o número de recompensas oferecidas, influenciam de forma significativa o número de apoios recebidos.

Segundo Kuppuswamy e Bayus (2015), projetos malsucedidos tendem a ter, em média, metas mais ambiciosas (US\$14,7 mil, aproximadamente) do que projetos bem-sucedidos (cerca de US\$3,4 mil) e mais de 90% dos criadores recorre ao Kickstarter somente uma vez.

Felipe (2017)<sup>16</sup>, por sua vez, notou uma relação similar no Catarse. Segundo seu estudo, há uma correlação negativa entre a meta estabelecida e o sucesso do financiamento, ou seja, a cada 1% de incremento da meta, há uma redução de 0,59 ponto percentual na probabilidade de sucesso dos projetos. Para o autor, uma explicação pode vir da natureza arbitrária das metas, que podem não ter relação com a necessidade financeira do projeto.

A respeito das recompensas, Felipe (2017) não encontrou correlação significativa entre sua quantidade e o sucesso da campanha, ou seja, o número de recompensas não parece atrair mais contribuições, o que vai ao encontro de outras pesquisas similares, como a de Carr (2013), que afirma cerca de 15% do sucesso dos projetos pode ser explicado pela existência de recompensas, ainda que não seja seu preditor principal.

---

<sup>15</sup> No Kickstarter, é possível filtrar, na lista de projetos, quais são os mais recentes e quais estão para encerrar.

<sup>16</sup> Felipe (2017) analisou 4.262 projetos publicados no Catarse entre 2011 e 2016. As informações foram cedidas ao pesquisador diretamente pela empresa, e compreendem um total arrecadado de cerca de R\$ 38 milhões, ou seja, próximo de 70% de todo o recurso angariado pela plataforma até o momento da sua pesquisa. O autor utilizou como variáveis de análise valor arrecadado, meta, categoria, duração dos projetos, número de aportes e recompensas, além de questões ligadas à geografia das regiões de onde partiram os projetos, tais como a distribuição de renda, tendo como indicador o índice de Gini, e o percentual de idosos e de analfabetismo.

Outra contribuição de Felipe (2017) tem a ver com as categorias dos projetos. O pesquisador, ao analisar os projetos, criou uma variável *dummy* chamada artes. Em outras palavras, ele identificou no seu banco quais projetos têm ligação com atividades artísticas, independentemente da categoria na qual o projeto está inscrito, e a utilizou em sua regressão. Esses projetos, segundo afirma, têm uma chance 8,7% menor de sucesso do que os outros, o que, apesar de contrariar uma percepção nossa de que o Catarse ainda está em um estágio mais próximo do início do Kickstarter, quando os projetos eram mais ligados às artes, aponta para uma diferença no desenvolvimento de campanhas de determinadas naturezas.

Por fim, os estudos indicam que, mesmo sendo uma prática típica da internet, há efeitos de geografia importantes na dinâmica de financiamento coletivos. Não só apoiadores estão aparentemente mais propensos a participar de campanhas geograficamente mais próximas de sua residência, mas projetos originados em regiões com mais densidade populacional, mais desenvolvimento econômico ou que sejam *hubs* de tecnologia parecem ser mais atraentes.

Segundo Agrawal, Catalini e Goldfarb (2011), a maior parte dos apoiadores, 86%, do Sellaband vivem a menos de 100 km do realizador e a distância média entre realizadores e apoiadores era de pouco mais de 1.500 km, ou seja, havia uma tendência de concentração geográfica dos projetos e dos financiadores. Os autores perceberam, ainda, que, mesmo que não existam restrições geográficas no acesso às plataformas digitais, os apoios tendem a se concentrar em regiões onde existem fontes tradicionais de financiamento, talvez, acreditam os autores, pela concentração de capital humano e pela facilidade de acesso a capital.

Para Mollick (2014), a geografia afeta os projetos do Kickstarter de diferentes maneiras. Em primeiro lugar, a distribuição dos projetos é desigual, concentrando-se em lugares com maior densidade populacional. Além disso, a geografia afeta a categoria dos projetos. Em lugares onde há mais produção musical, há mais projetos ligados a música; em São Francisco, onde fica o Vale do Silício, há uma prevalência de projetos de tecnologia. O pesquisador notou, ainda, que projetos de lugares onde há mais profissionais das indústrias criativas, como artes, design, entretenimento e mídia, têm mais chances de sucesso do que outros.

Felipe (2017) também identificou efeitos de geografia em seu estudo. Há, em sua amostra, uma correlação positiva entre a concentração de renda domiciliar nas regiões de onde partem os projetos e o seu sucesso. Naquelas com maior coeficiente de Gini, os projetos têm 75% a mais de chances de serem bem-sucedidos.

### 3 A PESQUISA EM COMUNICAÇÃO NA ERA DO BIG DATA

#### 3.1 A pesquisa na Era do *Big Data*

Do ponto de vista da metodologia de pesquisa e das técnicas de análise, talvez o principal desafio enfrentado pelos estudos que envolvem comunicação e internet seja a quantidade de dados disponíveis a respeito dos fenômenos. Para Munzert et al. (2014), isso é inusitado justamente porque, há bem pouco tempo, a escassez e a falta de acesso a informações sobre o comportamento dos indivíduos eram alguns dos principais problemas das ciências humanas.

Essa virada se deu com o surgimento dos processos de comunicação digitais, que oferecem diversas oportunidades de coleta e análise de dados a respeito do comportamento e das interações entre os indivíduos. Há informações detalhadas a respeito dos desejos, das percepções e das relações sociais em fóruns, blogs, sites de redes sociais, jogos, dentre vários, que permitem que pesquisadores e empresas que tiverem acesso aos recursos e ao treinamento adequados “escutem” melhor esses públicos (CHEN; CHIANG; STOREY, 2012). Ao mesmo tempo, em quase todos os setores da economia existem processos de construção de grandes bases de dados – finanças, medicina, indústria farmacêutica, governos, educação, etc. – que também oferecem oportunidades de se compreender melhor como agem e o que pensam os indivíduos (SCHUTT; O’NEIL, 2014).

Esse momento, que Pilny e Poole (2017) vão apelidar de “Era do *Big Data*”, possui as seguintes características gerais:

- a) *volume*, nunca produzimos dados na magnitude atual;
- b) *variedade*, esses dados são de naturezas das mais diversas;
- c) *velocidade*, não só a quantidade é única, mas a velocidade na qual novos dados são criados também é inédita; e
- d) *complexidade*, a estrutura dos dados é complexa e heterogênea, o que exige integração e reorganização.

Para Schutt e O’Neil (2014), não somente há dados disponíveis em grande quantidade, mas os próprios dados vêm se tornando produtos e serviços, como os algoritmos<sup>17</sup> de

---

<sup>17</sup> Algoritmos são, para Neyland (2019), comumente definidos como um conjunto de regras que deve ser seguido para em operações com o objetivo de resolver problemas. Os algoritmos de recomendação se utilizam do comportamento de usuários anteriores e de informações a respeito do indivíduo acessando um determinado site para prever assuntos ou produtos de interesse e, em seguida, fazer uma recomendação.

recomendação de produtos da Amazon, de amigos do Facebook, de músicas e filmes do Spotify ou do Netflix. Em outras áreas, dados orientam a concessão de crédito, o desenvolvimento de percursos acadêmicos ou o desenvolvimento de políticas públicas. Por isso, na visão de Schutt e O’Neil (2014), estamos vivenciando o início de uma mudança cultural, “na qual nosso comportamento muda os produtos e os produtos mudam nosso comportamento” (p. 5).

É, para eles, um processo de “dataficação”<sup>18</sup> do mundo.

O termo *big data*, para Schutt e O’Neil (2014), precisa ser bem delimitado já que vem sendo utilizado em contextos diversos e pode se referir a ideias completamente diferentes. O que é grande hoje, em termos de dados, é muito diferente do que era grande nos anos 70. Por isso, os autores propõem que estamos trabalhando com *big data* quando as bases de dados utilizadas ultrapassam a capacidade de análise e tratamento das soluções que são atualmente estado-da-arte em termos de memória, armazenagem, complexidade e velocidade de processamento.

Uma outra abordagem, mais simples, tem a ver com os recursos disponíveis para indivíduos e empresas. Nela, o pesquisador lida com *big data* quando os dados são tão grandes que não cabem ou não conseguem ser processadas em um computador pessoal (PROVOST; FAWCETT, 2013; SCHUTT; O’NEIL, 2014).

Nesse contexto, Pilny e Poole (2017) afirmam que os recursos disponíveis e ferramentas utilizadas até recentemente para coletar informações e analisar determinados fenômenos precisam transformar-se. No lugar das pesquisas de opinião, por exemplo, é possível extrair as interações entre indivíduos pelos dados em seus celulares, e-mails, sites de mídias sociais. Por isso, defendem que os pesquisadores aprendam um conjunto novo de técnicas e métodos, como processamento em paralelo, armazenagem em nuvem, modelagem estatística, etc.

Em termos práticos, simplesmente não há recursos financeiros, tempo e recursos humanos suficientes para coletar, armazenar, organizar e analisar determinadas fontes de dados sem o auxílio de computadores. No caso específico desta tese, as bases que estão sendo utilizadas têm informações detalhadas de quase 16 mil projetos de financiamento coletivo e cerca 952 mil usuários da plataforma Catarse. Ainda que essa amostra não se encaixe precisamente na definição de *big data*, essa pesquisa não seria possível e os conhecimentos resultantes mais passíveis de erro se não estivesse sendo utilizado um conjunto de técnicas e métodos, chamado de Ciências Sociais Computacionais.

---

<sup>18</sup> Esse processo não é, no entanto, neutro e sem problemas. Noble (2018) nos mostra, por exemplo, que a quase onipresença dos algoritmos no dia-a-dia, sejam eles visíveis ou não, reforça as desigualdades raciais, de gênero e de renda existentes na sociedade e na economia, como veremos à frente.

### 3.2 As Ciências Sociais Computacionais (CSC)

As Ciências Sociais Computacionais (CSC) formam um campo multidisciplinar que tenta explicar fenômenos sociais em diferentes escalas, dos atores individuais a grandes grupos, utilizando instrumentos baseados em computadores. Seu objetivo é entender as transformações sociais e o comportamento humano nessa “era do *big data*”. Para isso, utiliza métodos das ciências computacionais para resolver problemas das ciências sociais (PILNY; POOLE, 2017).

Para Cioffi-Revilla (2014), é uma disciplina que parte da percepção de que a informação tem um papel central na maneira como os sistemas sociais operam e que assume a computação como instrumento metodológico fundamental para se modelar e entender os fenômenos sociais. Da mesma forma que na astronomia os telescópios permitiram uma nova percepção do universo, e que na linguística a matemática permitiu uma nova compreensão das linguagens humanas, nas ciências sociais o uso de computadores permite a compreensão de fenômenos tais como o comportamento de grupos face a desastres naturais, segregação racial, ausência de liderança, etc.

Esse tipo de abordagem, para o autor, permite a análise de situações complexas demais para serem analisadas de outra forma, como no caso de simulações de comportamento de multidões, por exemplo.

Por muito tempo, modelos estatísticos capazes de serem aplicados em contextos sociais eram considerados impossíveis de serem desenvolvidos por conta da complexidade dos sistemas sociais (HELBING, 2010). Entretanto, mais recentemente, as estruturas e relações sociais começaram a ser representáveis estatisticamente através de princípios tomados emprestados de outras áreas do conhecimento, como, por exemplo, da física, de onde surgem os modelos que explicam a aleatoriedade, e da teoria do caos, que nos permite compreender sistemas de equações não-lineares. Isso só se tornou possível, para Schutt e O’Neil (2014), porque, com os avanços da computação, o custo da capacidade de processamento dessas informações é quase inexistente. Ao mesmo tempo, conseguimos obter dados suficientes a respeito dos comportamentos dos indivíduos para que essas previsões tenham significância estatística.

Pesquisas que se encaixam no paradigma das Ciências Sociais Computacionais utilizam análise automatizada de texto, *machine learning*, algoritmos e simulações para compreender dinâmicas e questões sociais complexas. Com essas técnicas, é possível automatizar a coleta e a categorização de grandes volumes de informação, de forma sistematizada, com baixo custo e com margem reduzida de erros. São pesquisas não somente menos custosas, principalmente

em recursos humanos, mas, sobretudo, utilizam bases de análise mais confiáveis e permitem compreender melhor a complexidade social contemporânea (CIOFFI-REVILLA, 2014).

As Ciências Sociais Computacionais surgem no mesmo contexto da chamada Ciência de Dados (ou *data science*, em inglês), um termo guarda-chuva, utilizado mais frequentemente em ambientes corporativos, que inclui as diversas disciplinas e técnicas surgidas em torno do uso de dados quantitativos aplicados aos mais diversos setores.

Segundo Schutt e O’Neil (2014), o termo cientista de dados foi criado em 2008 por dois executivos, um do LinkedIn e outro do Facebook, para designar o tipo de profissional que estavam recrutando e que era, então, muito raro. Um cientista de dados é um pesquisador, independentemente do seu campo de origem, que

[...] trabalha com grandes quantidades de dados e deve lidar com os problemas computacionais originados na estrutura, tamanho, confusão, complexidade e natureza desses dados, ao mesmo tempo em que tenta resolver problemas empíricos do mundo (p.15).

Para Wickham e Grolemund (2016), o objetivo da Ciência de Dados é transformar dados brutos em compreensão e conhecimento. Um projeto típico envolve a obtenção de dados, sua organização e transformação em um formato que possa ser utilizado posteriormente, a apresentação desses dados sob forma visual, a modelagem estatística e, por fim, a construção de suportes de comunicação que permitam ao cientista apresentar seus achados para suas audiências.

É um trabalho analítico que demanda habilidades com softwares e programação (CADY, 2017). Além da habilidade em se construir os modelos estatísticos, um cientista de dados é capaz de escrever os códigos do software capazes de fazer as análises, eventualmente utilizar computação distribuída para processar bases de dados grandes demais, e, ao final, interpretar os dados de forma a resolver seus problemas de pesquisa.

O que não significa, para Cady (2017), que um cientista de dados seja um “estatístico supervalorizado” (p.2). Para ele, esse cientista passa a maior parte de seu tempo fazendo com que os dados estejam organizados de tal forma que as análises estatísticas sejam aplicáveis, o que envolve garantir que os problemas analisados estejam de acordo com os objetivos das pesquisas e que os sentidos extraídos dos dados brutos não estejam contaminados por problemas dos próprios dados. Um cientista de dados vai começar seu trabalho em fontes de informação que um estatístico não compreende, como sites em HTML, imagens, e-mails, *logs* de servidores, etc. Depois que esse trabalho é realizado é que se podem utilizar as ferramentas estatísticas.

### 3.3 CSC e pesquisa quantitativa

Na base das CSC e da ciência de dados estão os métodos quantitativos e as análises estatísticas. Para Agresti, Franklin e Klingenberg (2018), estatística é mais do que um emaranhado de números e fórmulas confusas, mas uma maneira de refletir a respeito dos dados e de tentar quantificar a incerteza. Mais do que isso, é uma ciência que se dedica a transformar dados brutos “em conhecimento e compreensão do mundo” (p. 30). Através da inferência, a estatística permite que se tome decisões e se faça previsões a respeito de questões relevantes para a sociedade.

Segundo Creswell (2014), a distinção entre pesquisas qualitativas e quantitativas é, em geral, apresentada de maneira simplificada; as quantitativas como sendo as que usam números ou questões fechadas, as qualitativas como as que usam palavras e questionários abertos. Para ele, no entanto, a diferença tem mais a ver com as estratégias, os pressupostos e os métodos utilizados pelos pesquisadores no decorrer de seus estudos. Tem, ainda, para o autor, raízes nos processos históricos que marcaram a evolução das duas abordagens.

Os métodos qualitativos são indutivos; buscam explorar e compreender os sentidos individuais de um problema social ou humano, prezam a complexidade de dado contexto, e se utilizam de um número pequeno de observações e de análises em profundidade de cada um dos casos. Já os métodos quantitativos trilham um percurso dedutivo; buscam identificar as variáveis mensuráveis de tais problemas, estabelecer as relações entre eles e generalizar os seus achados, enquanto vão construindo proteções contra os vieses e controlando as explicações alternativas (CRESWELL, 2014; STOCKEMER, 2018).

O que norteia a escolha entre uma ou outra estratégia de pesquisa, para Creswell (2014), são: o problema de pesquisa, as experiências pessoais do pesquisador e a audiência para a qual o estudo será escrito. Pesquisas quantitativas são mais indicadas, para o autor, quando o problema se relaciona “(a) à identificação de fatores que influenciam um resultado, (b) à utilidade de uma intervenção ou (c) preditores para um evento” (p.50). à compreensão dos melhores

Para Stockemer (2018), o processo de pesquisa quantitativa começa e termina com questões e hipóteses originadas nas teorias existentes. Idealmente, seu questionamento endereça uma questão que seja relevante e importante, e que traga uma contribuição teórica para a literatura. As hipóteses ajudam a identificar claramente as variáveis dependentes e as independentes, e de que forma as independentes influenciam o comportamento das dependentes.

Para Schutt e O’Neil (2014), os modelos estatísticos são tentativas de compreender e representar a natureza através de expressões matemáticas. Para eles, os cientistas de dados “capturam a incerteza e a aleatoriedade dos processos geradores de dados através de funções que expressam o formato e a estrutura dos dados em si” (p.28). Modelos, em outras palavras, são construtos artificiais dos quais os detalhes estranhos são removidos e que nos permitem compreender, ao menos em parte, problemas de determinadas naturezas.

Schutt e O’Neil (2014) alertam para o equívoco de se considerar os dados quantitativos como sendo elementos objetivos do mundo. Para eles, os dados são ecos imperfeitos dos eventos em nossa sociedade, necessariamente incompletos por nossa incapacidade de compilar todos os dados existentes de todos os fenômenos sociais, e certamente enviesados pelos problemas sociais e históricos. Como exemplo, um algoritmo que define sozinho o melhor candidato para uma vaga muito provavelmente vai apontar um homem ao invés de uma mulher, porque homens são mais promovidos que mulheres e recebem menos *feed-backs* negativos em ambientes similares. Além disso, os algoritmos, como nos alerta Noble (2018), refletem também os preconceitos e vieses de seus programadores e gestores. Pessoas pretas e latinas, por exemplo, tendem a pagar prêmios e juros maiores nas relações bancárias que seus vizinhos brancos, sobretudo se são originários de bairros pobres. A busca por termos associados aos corpos de pessoas pretas geralmente resulta em sites de pornografia. Essa discriminação, que a autora vai chamar de opressão algorítmica, é tão presente que Noble (2018) não a considera um erro do sistema, mas que o sexismo e o racismo são partes fundamentais da arquitetura e da linguagem da internet, e a adoção desses sistemas sem a devida reflexão crítica acaba por reforçar as relações de opressão econômicas e sociais.

No que diz respeito à experiência pessoal do pesquisador, Creswell (2014) considera que um indivíduo treinado nos aspectos técnicos, nas linguagens de programação e na escrita específica das revistas mais ligadas aos métodos quantitativos tende a escolher abordagens desta natureza. Já pesquisadores com habilidades para entrevistas pessoais, com facilidade para observar os fenômenos presencialmente, vão tender a pesquisas mais qualitativas.

No caso desta tese, a escolha por métodos mais quantitativos tem a ver, além da experiência pessoal do pesquisador, com a natureza e a quantidade de dados coletados a respeito dos processos de financiamento, e com o objetivo de identificar quais elementos destas campanhas são determinantes para o apoio a um projeto de financiamento coletivo.

Ao mesmo tempo, esta é uma pesquisa de natureza empírica. Para Stockemer (2018), o ganha-pão dos cientistas sociais são seus métodos de pesquisa. Para explicar os fenômenos do mundo, é necessário que um pesquisador domine das ferramentas mais simples às mais complexas, sobretudo quando se propõe a realizar pesquisas empíricas.

Para o autor, há uma diferença entre o que chama de pesquisa empírica e de normativa no que tange a seus objetivos principais. Segundo afirma, a pesquisa empírica está menos interessada em como as coisas deveriam ser, nas normas, e mais em como elas são, como o mundo funciona, como está estruturado.

O que não implica uma ordem de importância entre elas. Enquanto as pesquisas empíricas e normativas têm objetivos diferentes, para Stockemer (2018) elas são complementares. Para entender como o mundo é, o pesquisador empírico deve estabelecer um quadro de referências, um padrão, que sirva como ponto de partida para se compreender os dados coletados e os testes a serem realizados, elementos que serão fornecidos pelas pesquisas normativas.

Para Stockemer (2018), é importante buscar em pesquisas empíricas o que chama de conhecimento intersubjetivo, ou seja, caso dois pesquisadores empreguem o mesmo processo de coleta e análise de dados, seus resultados devem ser análogos. Essa intersubjetividade é possível caso as pesquisas sigam os seguintes critérios:

- a) *Falsificabilidade*: as hipóteses ou afirmações devem poder ser provadas ou refutadas;
- b) *Transmissibilidade*: qualquer descoberta pode ser testada novamente ou replicada, usando-se os mesmos dados coletados inicialmente ou novos dados que estejam de acordo com os referenciais empíricos e métodos originais;
- c) *Natureza acumulativa do conhecimento*: a base para descobertas relevantes ou métodos de pesquisa é o conhecimento anterior, ou seja, pesquisadores não começam do nada em suas pesquisas, mas tentam confirmar, emendar ou expandir o conhecimento a partir dos conhecimentos e projetos já realizados anteriormente;
- d) *Generalização*: a partir de um número de casos específicos, é possível encontrar elementos universais. Em outras palavras, têm um peso maior as descobertas que se encaixam em mais casos.

### 3.4 Pesquisas Quantitativas em Comunicação

No campo da Comunicação, pesquisas empíricas com metodologia quantitativa não são as mais comuns. Ao analisar as estratégias metodológicas das pesquisas do campo da comunicação, Figaro (2014) nota que estas têm se concentrado nos produtos culturais de meios de comunicação, e tentam compreender sua função, seu poder e sua influência na sociedade e na cultura. Para a autora, são estudos que têm referenciais teóricos e metodologias derivadas,

em geral, das escolas funcionalista, crítica e semiótica, com larga produção e consolidação teórica e metodológica.

As pesquisas empíricas como forma de explicar os processos comunicacionais, por outro lado, são pouco realizadas. Para a autora, aparecem como exceção tradicionalmente, somente Robert Ezra Park, representante da Escola de Chicago, que, por sua técnica de entrevista, influenciou fortemente as ciências sociais dos anos 1920 e 1930, e os Estudos Culturais, de onde derivam-se as técnicas de observação participante e etnografia das audiências.

Uma explicação possível, para Figaro (2014), tem a ver com uma “crítica cientificista de origem positivista” (p.125), que considera menos importantes as pesquisas que não oferecem condições de verificação e controle dos dados obtidos. Essas críticas, para a autora, inibiram os pesquisadores da robustos.

Nesse sentido, acreditamos que existe uma lacuna no que diz respeito à utilização de métodos quantitativos, em especial das ferramentas de estatística, nas pesquisas de comunicação. O exemplo que Figaro (2014) utiliza para defender a possibilidade da utilização de métodos quantitativos em pesquisas deste campo é a de um questionário fechado de múltipla escolha que serviu de base para entrevistas face a face e grupos de discussão. Em outras palavras, não há indicação de que inferências estatísticas ou técnicas mais elaboradas tenham sido utilizadas na pesquisa em questão.

Da mesma forma, em um dos manuais mais utilizados a respeito de métodos e técnicas de pesquisa em comunicação, o de Duarte e Barros (2005), as técnicas de pesquisa quantitativa apresentadas são:

- a) *a pesquisa de opinião*, ou *survey*, sendo que a discussão limitou-se à construção de um questionário, sem um aprofundamento nas formas de análise;
- b) *a análise semiótica*, na qual as técnicas quantitativas são apontadas como possibilidade;
- c) *a análise de conteúdo*, único procedimento no qual a inferência estatística é mencionada, mas com a ressalva de que há controvérsia em sua utilização “excessiva” (p. 282);
- d) *a auditoria de imagem*, com os aspectos quantitativos sendo criticados por ignorar “instâncias que condicionam a própria inserção de notas, notícias ou reportagens [...] ou o peso e importância dos veículos” (p. 348).

Em todo o livro, a palavra estatística, no plural ou singular, é utilizada 16 vezes em quase 400 páginas, sendo que em várias vezes associada a ressalvas como:

Se a quantidade de material a ser analisado for muito grande, isso exigirá a adoção de procedimentos de estatística para se obter uma visão de conjunto, devendo-se optar

pela análise quantitativa. **Neste caso, ganha-se em abrangência, mas perde-se em profundidade.** Se o objetivo for aprofundar o conteúdo, torna-se impossível o processamento de uma grande quantidade de dados, devendo ser realizada uma análise qualitativa. (DUARTE; BARROS, 2005, p. 293, grifo nosso)

Para Epstein (2005), os procedimentos quantitativos têm sido menos valorizados nas ciências sociais por serem considerados reducionistas, o que não permitiria a um pesquisador, através dessas ferramentas, dar real dimensão à “complexidade do real” (p.26). Para o autor, no entanto, não há procedimento analítico que não seja, em algum grau, reducionista. Ao contrário, pensa Epstein (2005), os procedimentos quantitativos são indispensáveis por sua possibilidade de generalização, ainda que resguardadas as condições nos quais os casos se diferem.

Na mesma linha, Lopes (2004) acredita que há uma falsa dicotomia entre pesquisas de natureza quantitativa e qualitativa, sobretudo pela dificuldade de se determinar um limite preciso entre uma e outra. Nem as pesquisas quantitativas são meramente descritivas, nem as qualitativas, somente interpretativas. Para a autora, “as operações quantitativas se apoiam em dados qualitativos originalmente coletados e em seguida transformados” (p. 22). Da mesma forma, Creswell (2014) afirma que, em geral, os estudos iniciam-se com pesquisas quantitativas que dão um panorama geral do fenômeno e terminam com abordagens mais qualitativas, que ajudam a explicar os números. Da mesma forma, Stockemer (2018) considera que os métodos quantitativos apontam para as relações entre as variáveis observáveis de um fenômeno, enquanto as pesquisas qualitativas vão dar sentido para essas relações, o que os torna, a seu ver, os métodos complementares, e não excludentes.

### 3.5 Extração automatizada de dados e implicações éticas

Para que esse tipo de pesquisa seja possível de ser realizada, é necessário ter dados. De preferência, muitos dados. Por isso, existe todo um conjunto de técnicas para extrair e estruturá-los em bases, genericamente chamadas de *data-mining*<sup>19</sup>.

De acordo com Provost e Fawcett (2013), a expressão refere-se aos procedimentos utilizados para extrair conhecimento de dados brutos, pouco estruturados. De acordo com Cioffi-Revilla (2014), essas técnicas tiveram origem na extração automatizada de texto, inventada nos anos 1960 com o nome de análise quantitativa de texto. Apesar disso, até o

<sup>19</sup> Há diversas denominações distintas para técnicas relativamente próximas. Em Cioffi-Revilla (2014), *data-mining* aparece como uma expressão guarda-chuva para as diversas técnicas de análise de dados, independentemente da origem e do formato. Em Johnson e Gupta (2012), *web-mining* é apresentado como sendo o trabalho de “descoberta” de conhecimento a partir de dados obtidos na internet. Feldman e Sanger (2007) apresenta o *text-mining* como um conjunto de ferramentas análogas ao *data-mining*, mas específicas para a análise de texto não-estruturado, através da identificação de padrões em coleções de documentos.

surgimento de computadores, o *data-mining* era uma tarefa realizada por codificadores, pesquisadores especialmente treinados para registrar e identificar manualmente informações, na maior parte das vezes, textuais, provenientes, sobretudo, da imprensa e de fontes governamentais. A execução desse trabalho por um humano, segundo Cioffi Revilla (2014), demanda muito tempo, treinamento e está mais susceptível a erros do que se realizada por um algoritmo<sup>20</sup>.

Em termos gerais, a extração de dados a partir de um site pode ser feita de forma automatizada ou não. Os métodos não automatizados envolvem a participação do pesquisador em fóruns, grupos, sites de redes sociais, etc., escutando ou participando das conversas ou interações a respeito dos temas pesquisados (ALIM, 2014; FRAGOSO; RE- CUERO; AMARAL, 2011).

As técnicas de extração automatizadas são aquelas em que não há interferência humana ao longo do processo e é possível realizar esse trabalho com fontes em que as informações estejam estruturadas, ou seja, organizadas e facilmente arranjadas em bases de dados, ou não, quando é necessário realizar um trabalho posterior de limpeza e arrumação dessas informações, nos mais diversos formatos, sejam eles textos, sons ou imagens. De maneira geral, os métodos automatizados podem utilizar APIs ou *crawlers*.

As *Application Programming Interfaces* (APIs) são canais através dos quais os sistemas conseguem interagir e trocar informações ou executar serviços de forma estruturada. Um software para publicar mensagens no Twitter, por exemplo, utiliza a sua API para enviar e receber informações para a plataforma de forma simplificada. Aplicativos como o Tinder acessam a API do Facebook para extrair os *likes* e os contatos dos usuários e, com base nessas informações, estabelecer quem são seus potenciais parceiros (HELMOND, 2015).

Nos sites que as disponibilizam, o melhor, segundo Munzert et al. (2014), é montar as bases de dados a partir das APIs, já que elas fornecem exatamente as informações que são demandadas, de forma padronizada, respeitando as regras estabelecidas pelos Termos de Serviços<sup>21</sup> no que diz respeito ao tipo de informação que pode ser extraída, e em que volume,

---

<sup>20</sup> Cioffi-Revilla (2014) define um algoritmo como sendo um conjunto de etapas a serem computadas com vistas a um resultado. Um exemplo seria um algoritmo escrito para extrair informações de uma página web e estruturá-las sob forma de base de dados para posterior análise. Algoritmos de recomendação utilizam informações a respeito de escolhas passadas, tanto de um indivíduo como de outros com perfil ou comportamento similares, para indicar um produto ou um conteúdo.

<sup>21</sup> Termos de Serviços são documentos que estipulam as regras de utilização de serviços na internet, geralmente apresentadas aos usuários no primeiro contato entre ele e as plataformas. Em geral, as políticas de privacidade e de utilização de dados pessoais pelos sites são apresentadas nos termos e, ao final, é exigido do usuário que concorde com a utilização descrita no documento para que tenha acesso aos serviços oferecidos (OBAR; OELDORF-HIRSCH, 2020).

formato e tempo. Para se utilizar APIs, é necessário criar um algoritmo simples, utilizando linguagens de programação, como o R ou o Python, ou programas especialmente desenvolvidos para esse fim.

APIs, no entanto, podem ser um empecilho para o desenvolvimento de determinadas pesquisas por ser muito restritivas, demandar pagamento por determinados dados, ou não dar acesso a nenhuma informação da plataforma (MUNZERT et al., 2014), o que é o caso de empresas pouco transparentes como o Facebook. A empresa, após o escândalo da Cambridge Analytica (CONVERSATION, 2015), limitou severamente o acesso aos dados de usuários<sup>22</sup>, a tal ponto que tornou praticamente inviável qualquer pesquisa em larga escala. Do ponto de vista acadêmico, uma API restrita impede que o pesquisador obtenha uma amostra mais representativa de comportamentos, eventos, campanhas, o que pode comprometer a validade dos achados<sup>23</sup>. Além disso, promove uma desigualdade de acesso aos dados e limita a capacidade de programas ou estudos que não tenham recursos financeiros ou acesso privilegiado.

Já os *crawlers* são pequenos programas escritos para acessar um site, encontrar a informação desejada e estruturar os dados resultantes em uma base. A técnica, frequentemente chamada de *web scraping*, ou raspagem, é mais trabalhosa e demanda mais recursos e conhecimentos dos pesquisadores.

Essa técnica é especialmente eficiente para se obter informações a partir do código HTML, em sites que estejam disponíveis de forma pública, sem proteção por senha. Serve também para se extrair dados de perfis em redes sociais, principalmente aqueles sem as restrições de privacidade configuradas, o que levanta alguns questionamentos legais e éticos (MUNZERT et al., 2014).

Do ponto de vista legal, a extração de dados sem autorização prévia pode ser considerada uma violação dos direitos de propriedade intelectual dos sites já que raspar dados é, em termos simples, copiar o conteúdo de algum lugar (MUNZERT et al., 2014). Não encontramos nenhum estudo, no Brasil, a respeito de violações dos direitos de propriedade intelectual por raspagem

---

<sup>22</sup> Somente um grupo de 60 pesquisadores, de 11 países, em todo o mundo tem, hoje, acesso aos dados da API do Facebook. Eles foram selecionados por um conselho, o Social Science Research Council (SSRC), que analisou seus projetos de pesquisa e definiu aqueles que poderiam auxiliar a entender melhor “como as notícias a respeito de política são compartilhadas [...], como as plataformas podem influenciar a opinião pública” (SSRC, 2019). No Brasil, somente dois pesquisadores da USP participam do programa.

<sup>23</sup> A API do Twitter, considerada a plataforma mais aberta e, por isso, uma das preferidas por pesquisadores, dá acesso somente a até 18 mil *posts* a cada 15 minutos, sendo que essas postagens são selecionadas “por relevância”, ou seja, o conjunto de mensagens não representa a totalidade do que foi comentado usando as palavras-chave escolhidas, naquele período, a não ser que se contrate uma assinatura que começa em US\$ 149,00 por mês (TWITTER, 2020a).

de dados. Munzert et al. (2014), no entanto, afirmam que houve, nos EUA, alguns casos assim, mas, sempre que houve judicialização havia também interesses comerciais envolvidos, ou seja, empresas ou indivíduos que raspam dados de outras empresas com fins econômicos, ou para vender os dados ou para oferecer serviços.

Em artigo publicado no Jota.info, Longhi e Petiti (2018) argumentam que os limites entre o público e o privado são pouco claros na internet (os dados pertenceriam aos usuários ou às plataformas?) e que há poucos casos judiciais ou administrativos no Brasil para que se determine a responsabilização ou a ilegalidade da prática, ainda que considerem, em princípio, que a Lei de Direitos Autorais (lei nº 9.610/1998) define um banco de dados como obra intelectual, da mesma natureza que as coletâneas, e que a Lei de Propriedade Industrial (Lei nº 9.279/96) considera crime “a utilização, venda, exposição ou oferecimento de obra intelectual, sem a devida autorização do autor, nem a indicação de sua autoria”. Da mesma forma que Munzert et al. (2014), Longhi e Petiti (2018) afirmam que os casos conhecidos de judicialização envolvem interesses econômicos e que uma eventual responsabilização legal por uso indevido dependeria “do tipo de acesso obtido pelo *scraper*, da quantidade de informações acessadas e copiadas, do grau em que o acesso afeta adversamente o sistema do proprietário do *website* e dos tipos e forma de proibições em tal conduta”.

Do ponto de vista ético, a principal questão refere-se ao chamado consentimento informado, ou consentimento livre esclarecido, quando um participante declara estar de acordo em participar de uma pesquisa. O princípio básico, para Flick (2016), é que qualquer procedimento que fuja da expectativa de comportamento normativo tem que ser consentido. Como exemplo, cortar a barriga de alguém (o que acontece em um procedimento cirúrgico) não é um comportamento que se espera corriqueiramente, em um contexto que não seja médico. Logo, demanda uma autorização clara de quem terá a barriga cortada (ou dos seus responsáveis) para ser realizado.

Esse tipo de declaração, que tem origem nas ciências humanas, em especial nas pesquisas biomédicas, garante que os participantes de um estudo ou de uma pesquisa sejam respeitados em seus direitos e no seu bem-estar, aí incluído o direito à privacidade. Para Flick (2016), esse consentimento demanda uma comunicação clara e simples, que permita ao participante compreender os riscos, os benefícios e os objetivos da pesquisa da qual está participando, assim como as maneiras como os seus dados serão armazenados e tratados ao longo do estudo. Desde a resolução do Conselho Nacional de Saúde nº 510 de 2016, é uma exigência também para pesquisas das ciências humanas e sociais (CNS, 2016).

Segundo Alim (2014), a preocupação com o consentimento em pesquisas que utilizam extração automatizada é recente e há diversos estudos relevantes que não obtiveram a autorização explícita de seus participantes, o que pode ser especialmente sensível sobretudo porque há muitos menores de idade participando de sites de redes sociais, ainda que desrespeitando os Termos de Serviços. Flick (2016) relembra, como exemplo, uma pesquisa intitulada *Experimental Evidence Of Massive-Scale Emotional Contagion Through Social Networks* (KRAMER; GUILLORY; HANCOCK, 2014), realizada em conjunto entre o Departamento de Comunicação e Ciências da Informação da Universidade de Cornell e o Facebook, cujo objetivo era determinar se os *posts* dos usuários poderiam ser influenciados por exposição e conteúdo intencionalmente positivo ou negativo, ou seja, se é possível modificar o humor dos usuários da plataforma através do que chamaram de “contágio emocional”. Nenhum dos 689,003 usuários que foram alvo de manipulação consentiu com a estudo nem teve a opção de sair dele (em inglês, *opt-out*)<sup>24</sup>.

Para Flick (2016), um dos principais problemas do estudo de Kramer, Guillory e Hancock (2014) é que ele desrespeitou a expectativa dos usuários a respeito do comportamento da plataforma. Para a pesquisadora, quem usa o serviço do Facebook espera<sup>25</sup>:

- a) *Users expect their news feeds in Facebook to be published faithfully by the people, communities, and organizations they follow and not manipulated.*
- b) *Users expect to be explicitly asked to participate in any experimental research conducted by Facebook.*
- c) *Users expect Facebook to conduct human research under the auspices of some sort of independent ethics committee. (FLICK, 2016, p.22)*

Neste caso, para Flick (2016), ainda que exista um contrato entre usuário e empresa que dê uma autorização prévia para utilização de dados de uma determinada maneira, isso não é suficiente para que se constitua consentimento livre esclarecido pela falta de clareza e transparência nos procedimentos.

---

<sup>24</sup> O argumento defendido pelos pesquisadores e pelo conselho de ética da Universidade de Cornell, de acordo com Flick (2016) para a realização e validade do estudo é que toda a pesquisa foi acompanhada pelo conselho de ética da instituição e que o Facebook, enquanto empresa privada, não está obrigada a se submeter aos códigos de ética de pesquisa acadêmica no que diz respeito à sua relação com seus usuários. Já o Facebook, ainda de acordo com Flick (2016), afirma que seus termos de serviço deixam claros que os dados dos usuários podem ser utilizados para pesquisa acadêmica sem aviso prévio e que não considera o estudo problemático.

<sup>25</sup> Tradução nossa: 1) usuários esperam que as informações no Facebook sejam publicadas fielmente pelas pessoas, comunidades e organizações que eles seguem, sem manipulações. 2) usuários esperam ser consultados explicitamente se desejam participar de pesquisas conduzidas pelo Facebook. 3) usuários esperam que o Facebook conduza pesquisas com humanos com a supervisão de algum tipo de comitê de ética independente.

Para Alim (2014), as pesquisas que fazem uso de APIs tendem a respeitar mais o desejo dos indivíduos porque, em vários casos, o usuário tem que autorizar a extração para que a interface disponibilize as informações. Além disso, essas pesquisas respeitam os Termos de Serviço dos sites e as condições de utilização dos dados acordadas entre indivíduos e empresas. Esse não é o caso com os *web crawlers*, que vão “raspando” as informações que encontram pelo caminho.

Alim (2014) observa que, na maior parte das vezes, pesquisadores utilizam os Termos de Serviço para saber que tipo de informação pode ser extraída e qual o uso permitido desses dados. O Facebook, por exemplo, deixa claro em seus termos que ninguém pode acessar os dados da plataforma com meios automatizados sem autorização explícita (FA- CEBOOK, 2019). Da mesma forma, o Twitter afirma, em seus termos:

Você não pode realizar nenhuma das seguintes ações ao acessar ou utilizar os Serviços: [...] (iii) acessar ou pesquisar, ou tentar acessar ou pesquisar os Serviços por quaisquer meios (automatizados ou de qualquer outra maneira) de outra forma que não por meio das interfaces publicadas e atualizadas disponibilizadas pelo Twitter (e sujeito aos respectivos termos e condições), exceto quando você tenha sido especificamente autorizado a realizar esses atos por meio de um acordo separado com o Twitter (OBS.: o crawling ou recurso a indexadores automáticos nos Serviços é admissível desde que efetuado de acordo com os termos e condições constantes do arquivo robots.txt; não obstante, é expressamente proibido o scraping dos Serviços sem o consentimento prévio do Twitter. (TWITTER, 2020b)

No caso do Catarse, os Termos de Serviço não impedem a obtenção automatizada dos dados de sua plataforma. Há restrições a respeito da coleta de informações que possam identificar os usuários individualmente, ou da utilização de equipamentos ou softwares que interfiram no funcionamento do site. No entanto, o Catarse deixa claro que todas as ofertas feitas pelos realizadores são de caráter público, e podem ser acessadas livremente CATARSE (2018d).

Há pesquisadores, no entanto, de acordo com Alim (2014), que consideram que informações publicamente acessíveis, como é o caso dos perfis em sites de redes sociais que não tenham restrições de acesso<sup>26</sup>, estão em domínio público e a invasão da privacidade desses usuários ocorre somente em poucos casos. Alim (2014) relembra, ainda, que sites, de qualquer natureza, podem ser considerados documentos eletrônicos e não indivíduos humanos. Logo,

<sup>26</sup> A questão das restrições de privacidade em sites de redes sociais é polêmica. Não há consenso a respeito da capacidade de compreensão por parte dos usuários do quanto são realmente visíveis nessas plataformas (ALIM, 2014). Logo, ainda que não tenham restringido o acesso, isso não significa, necessariamente, que têm consciência dos riscos da exposição pública. Por outro lado, ainda que um usuário configure seu perfil como sendo privado, existe a possibilidade de ter suas informações acessadas através de sua rede de amigos, como foi o caso no escândalo da Cambridge Analytica.

pesquisas que usam extração automatizada estariam, por essa perspectiva, realizando uma outra forma de pesquisa documental, o que não demandaria uma autorização formal prévia.

Uma preocupação indicada por Alim (2014) refere-se à viabilidade de se obter consentimento explícito de milhares ou milhões de usuários que são objeto desse tipo de pesquisa. No nosso caso, como exemplo, a base com os usuários cadastrados na plataforma tem mais de 950 mil entradas, sendo que esses perfis não são completos, o que torna inviável a coleta desse tipo de consentimento individualmente.

Ainda no que diz respeito à resolução do CNS de 2016, dentre as pesquisas que não demandam obrigatoriamente registro nos Conselhos de Ética em Pesquisa e, consequentemente, não tem a necessidade de obtenção de consentimento explícito dos participantes, estão:

- I - pesquisa de opinião pública com participantes não identificados;
- II - pesquisa que utilize informações de acesso público, nos termos da Lei nº 12.527, de 18 de novembro de 2011;
- III - pesquisa que utilize informações de domínio público;**
- IV - pesquisa censitária;
- V - **pesquisa com bancos de dados, cujas informações são agregadas, sem possibilidade de identificação individual;** e
- VI - pesquisa realizada exclusivamente com textos científicos para revisão da literatura científica;
- VII - pesquisa que objetiva o aprofundamento teórico de situações que emergem espontânea e contingencialmente na prática profissional, desde que não revelem dados que possam identificar o sujeito; e
- VIII - atividade realizada com o intuito exclusivamente de educação, ensino ou treinamento sem finalidade de pesquisa científica, de alunos de graduação, de curso técnico, ou de profissionais em especialização. (CNS, 2016, p.2, grifos do autor)

Em outras palavras, pesquisas que não identifiquem individualmente os participantes ou que se apoiem em informações públicas, o que consideramos ser o caso no que se refere aos dados dos projetos, não demandam consentimento prévio ou registro no Conselho de Ética da instituição na qual o estudo está sendo realizado. No que diz respeito aos usuários, as informações coletadas que poderiam servir para identifica-los, como nome, perfil em sites de redes sociais e apresentação individual foram eliminados após a consolidação das bases. As únicas variáveis desta base utilizadas neste estudo, e somente dos realizadores, foram número de projetos com os quais contribuiu, número de projetos publicados, número de seguidores e número de pessoas que o realizador segue.

Uma última preocupação descrita por Alim (2014) refere-se à disseminação das informações extraídas. Dependendo da maneira como os participantes estão identificados nas bases de dados, é possível, em alguns casos, traçar sua origem, raça, endereço, comportamentos sexuais, etc. Por isso, a disponibilização pública das bases utilizadas, apesar de ser

uma das recomendações para que se possa verificar e checar os achados posteriormente aos estudos concluídos, deve ser feita de tal forma que os dados individuais sejam protegidos, ou por anonimização ou por exclusão de determinadas informações, ainda que exista um questionamento sobre se o esforço de anonimização é suficiente para garantir a privacidade dos participantes. Alim (2014) defende, ainda, que se deve controlar o acesso aos sistemas que contenham informações individuais no decorrer da pesquisa para evitar vazamentos ou usos indevidos.

## 4 METODOLOGIA E ANÁLISE DE DADOS

### 4.1 Problema de pesquisa e escopo do estudo

Como discutimos anteriormente, o objetivo deste estudo é identificar quais são os determinantes do apoio nos projetos de financiamento coletivo publicados no Catarse, ou seja, que elementos fazem com que os apoiadores escolham uma proposta e a ajudem a ser viabilizada. Acreditamos que esses elementos são identificáveis nas características dos projetos, nas estratégias de comunicação utilizadas e no comportamento dos empreendedores ao longo das campanhas.

Esse objeto interessa porque projetos de financiamento coletivo fazem parte das novas formas de sociabilidade proporcionadas pelo surgimento da internet. Neste cenário, novos processos e formas de vinculação social se estruturam, o que é objeto de interesse especial dos estudos de comunicação.

É um objeto interessante, ainda, porque obriga o pesquisador a se aventurar em novos campos de conhecimento, dada a complexidade do fenômeno e a dificuldade de dar sentido a uma quantidade muito grande de informações. Por isso, acreditamos que existem, ainda, contribuições relevantes do ponto de vista da metodologia, em especial dos métodos quantitativos, para o campo da comunicação.

### 4.2 Coleta de dados

Para realizar as análises, consolidamos informações a respeito de 15.618 projetos cadastrados no Catarse, o que corresponde à totalidade dos projetos já finalizados até 25 de maio de 2018. Além disso, coletamos informações individuais a respeito de cada membro cadastrado no site, pouco mais de 952 mil usuários, o que representa a totalidade dos usuários inscritos na plataforma até 26 de julho de 2018. As informações das duas bases foram cruzadas e consolidadas, e todas as informações que poderiam identificar os usuários foram suprimidas<sup>27</sup>.

A coleta dos dados foi realizada por *web scrapping*, utilizando duas soluções, uma comercial, chamada import.io (figuras 5 e 6), baseada em rede, e outra de código aberto, instalada em um computador do pesquisador.

---

<sup>27</sup> Após os dados consolidados, retiramos da base 23 projetos que solicitaram mais de R\$ 3 milhões. Nenhum deles recebeu contribuições relevantes (somente dois receberam 1 doação cada, de R\$ 10,00 e de R\$ 50,00, respectivamente) e nenhum tinha expectativas realistas. Um deles, por exemplo, solicitou R\$ 800 milhões em doações para construir um *shopping center*. Com isso, a base utilizada nas análises tem 15.595 observações. Antes de realizar a modelagem dos dados, unimos as categorias Gastronomia e Pessoais, que têm características similares, para aumentar a robustez das estimativas.

Figura 5 - Tela de captura de informações do import.io



Fonte: Acervo do autor

A solução comercial é uma ferramenta baseada inteiramente em interface visual. Sua função principal é converter grandes volumes de dados de *websites* em dados estruturados, sem formatação, o que é fundamental para que softwares de análise estatística ou editores de planilhas eletrônicas, como o Microsoft Excel, possam ser utilizados na interpretação desses dados.

Existem duas vantagens importantes ao se utilizar soluções comerciais como o import.io. A primeira é a facilidade de se escolher a informação que se precisa extrair e de se determinar o arquivo de destino sem ter nenhum tipo de conhecimento em programação. A segunda é que arquivos grandes demais frequentemente demandam muito poder de processamento e muita memória, o que não costuma ser o caso de computadores pessoais. A desvantagem óbvia é justamente ser uma ferramenta comercial, logo paga. O import.io cobrava, no momento da coleta, por *query*, ou seja, por cada interrogação ou contato que estabelece com o site que está sendo raspado.

Figura 6 - Dados de projetos do Catarse, estruturados pelo import.io.

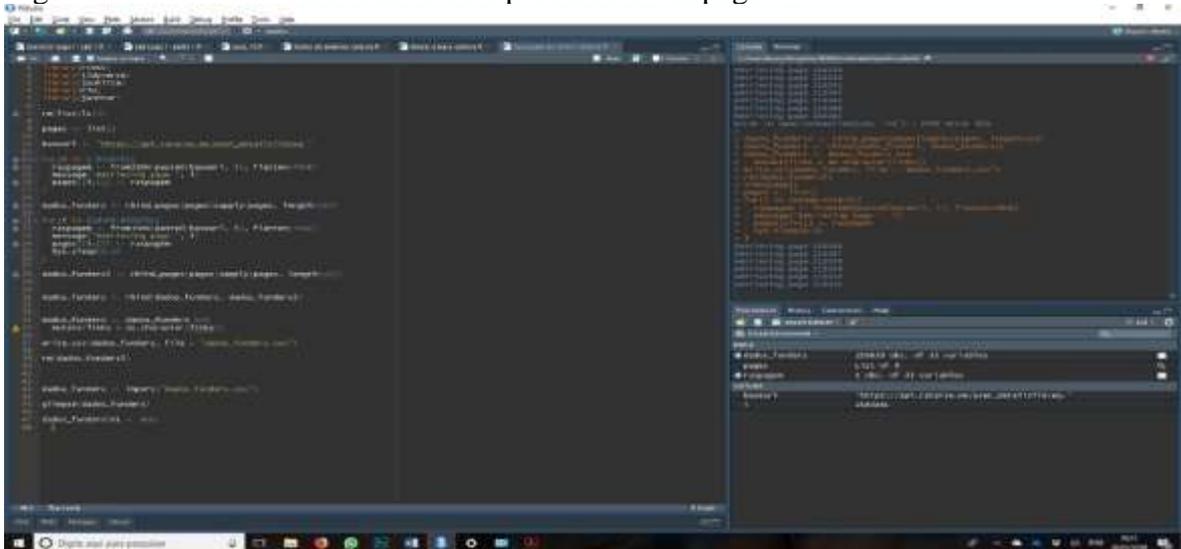
Nome	Autor	Valor arrecadado	aportadores	Porcentagem
Natal Invisível 2017 - Natal Para Todos	por CP Invisível	R\$ 60.523	1027 pessoas	79%
Documentário: História do Casarão de Pelotas	por Diego Pereira	R\$ 22.278	263 pessoas	96%
Tudo é sobre o Brasil	por Victoriano	R\$ 21.470	183 pessoas	167%
Superlatos - A Sua Fonte de Inspiração 4C	por Teresiano Rodrigues Teresiano Perery	R\$ 2.505	34 pessoas	2%
Projeto Professor Thiago	por Thiago Emergentis	R\$ 21.991	173 pessoas	100%
Música Escrita (sua história)	por Lucas Assunção	R\$ 21.119	119 pessoas	37%
AMTUB	por Helton Marcelo	R\$ 8.210	82 pessoas	100%
Contos e Histórias de Natal - O Livro de Natal	por Camilla Mendes Araújo - O Ca de Leite	R\$ 3.170	17 pessoas	7%
Um que não espera? - Rapto Show	por Elton Carlos Pereira	R\$ 8.000	87 pessoas	40%
Illego Cinema - A História	por Rogério Calmon	R\$ 544	17 pessoas	13%
Projeto Memórias que Sobram	por Ivo Vinício	R\$ 0	0 pessoas	0%
Documentário: Minha escola, minha região	por Raimundo	R\$ 3.275	56 pessoas	87%
No Caro Vinte e Cinco / Rapagem Festas	por Rapagem Festas	R\$ 0	0 pessoas	0%
Carreira: A História	por P. da Silva	R\$ 999	13 pessoas	40%
Case de São	por Roberto Almeida	R\$ 50	3 pessoas	1%
Serviço de Audo com Audo High	por P. da Silva	R\$ 4.200	17 pessoas	37%
Quem Falei na Janela	por Carolina Pastorek e Helena Inácio	R\$ 10	1 pessoa	0%
PROJ - Movimento Urbano em 3 Dim.	por Renato	R\$ 20	1 pessoa	0%

Fonte: Acervo do autor

A solução *opensource* foi o R (figura 7), uma linguagem de programação desenvolvida inicialmente para realizar análises estatísticas, mas que, rapidamente, se expandiu, através de seus pacotes, para a realização de diversas tarefas, tais como edição de texto, publicação e editoração. A principal vantagem do R, além do fato de ser distribuído gratuitamente e de suas capacidades, é o fato de possuir uma grande comunidade de usuários e de programadores que trocam informações e expandem suas funcionalidades. Além disso, todos os *scripts* utilizados durante a pesquisa podem ser compartilhados, o que confere transparência à pesquisa e permite sua reprodutibilidade. O próprio pesquisador pode revisitar sua metodologia e atualizar sua pesquisa posteriormente com facilidade. Por esse motivo, linguagens de programação, o R em especial, vêm se tornando populares entre pesquisadores que utilizam dados quantitativos (MUNZERT et al., 2014).

Uma dificuldade que encontramos ao utilizar o R, além da necessidade de se aprender a programar, é o tempo necessário para se obter os dados por raspagem. A base com informações a respeito dos usuários precisou de 11 dias (pouco menos de 265 horas) de processamento ininterrupto, dadas as limitações impostas pelo site do Catarse e pelo tipo de equipamento utilizado. Além de ter sido utilizado na coleta de dados, utilizamos o R na organização e estruturação da base, nas análises estáticas, e na elaboração dos gráficos e tabelas<sup>28</sup>.

Figura 7 - Interface RStudio durante o processo de raspagem.



Fonte: Acervo do autor

<sup>28</sup> Os pacotes utilizados foram Jsonlite (OOMS, 2014) e Rvest (WICKHAM, 2019) para o scrapping, arm (GELMAN; SU, 2018), Tidyverse (WICKHAM et al., 2019), Rio (CHAN et al., 2018), Haven (WICKHAM; MILLER, 2019) e Janitor (FIRKE, 2020) para a limpeza, organização e análise dos dados. Na modelagem, foram utilizados os pacotes glmmTMB (BROOKS et al., 2017), nnet (VENA-BLES; RIPLEY, 2002), effects (FOX; HONG, 2009), ggeffects (LU'DECKE, 2018), sjPlot (LU'DECKE, 2020), countreg (KLEIBER; ZEILEIS, 2016), DHARMa (HARTIG, 2020), pscl (ZEILEIS; KLEIBER; JACKMAN, 2008) e AER (KLEIBER; ZEILEIS, 2008).

Em seguida, utilizamos o algoritmo de processamento de linguagem natural do Google<sup>29</sup> para analisar os textos das descrições. Esse algoritmo é capaz de identificar a estrutura e o sentido dos textos, através de *machine learning*. Em língua portuguesa, o algoritmo era capaz de realizar, à época da pesquisa, somente análise de sentimento, com duas pontuações: a magnitude, ou seja, o quão emocional é um texto, e o score, que indica a inclinação emocional do texto, se mais negativo ou positivo (GOOGLE, 2020). Esses dois índices foram utilizados nas regressões como variáveis de controle.

Por fim, para interpretar os dados, utilizamos estatística descritiva e modelos de regressão para identificar que variáveis determinam o apoio de um usuário aos projetos.

### 4.3 Modelos binomiais negativos e modelos multinível

Segundo Cameron e Trivedi (2013), os modelos binomiais negativos, que fazem parte da família dos modelos de contagem, referem-se ao número de vezes que um evento ocorre e especifica a distribuição da probabilidade do número de ocorrências deste evento. São uma extensão dos modelos de Poisson e são utilizados quando a variável dependente (no nosso caso, o número de apoiadores de um projeto) tem uma distribuição não linear, não negativa, tem como respostas muito elevadas eventos raros e, diferentemente do que tolera o modelo de Poisson, há problemas de sobredispersão da distribuição (BAR-RETO, 2011; CAMERON; TRIVEDI, 2013). Os coeficientes estimados são expressos em logaritmos e são interpretados, após exponenciados, como probabilidades.

Para definir qual modelo de contagem se ajustaria melhor aos dados, fizemos estimativas utilizando os modelos de Poisson, Binomial Negativo, Hurdle, Hurdle Binomial Negativo e Zero-inflated e comparamos os resultados graficamente, através de rootogramas (KLEIBER; ZEILEIS, 2016) (apêndice B) e por AIC (tabela 5).

Tabela 5 - Graus de liberdade e índices AIC dos modelos de contagem

	DF	AIC
Binomial Negativo	19	114.802,9
Hurdle	36	908.466,3
Poisson	18	116.431,7
Zero Inflated	36	908.464,6
Hurdle Binomial Negativo	36	908.466,3

Fonte: Elaboração do autor

<sup>29</sup> O pacote utilizado no R foi o googleLanguageR (EDMONDSON, 2020).

Já modelos multinível, também chamados de hierárquicos, assumem que o contexto no qual cada um dos fenômenos acontece influencia o resultado. O exemplo clássico, como apontam Field, Miles e Field (2012), é o da análise do desempenho de estudantes, que é influenciado pelo contexto da sala de aula. Um professor que estimula experimentação vai ter resultados diferentes de outro mais cauteloso, ainda que estejam em uma mesma instituição de ensino.

Do ponto de vista das análises estatísticas, isso significa que os resíduos não são independentes, o que seria uma violação dos pressupostos dos modelos não hierárquicos. Modelos lineares, como apontam Finch, Bolin e Kelley (2019), assumem que não há relação entre as observações de uma amostra para a variável dependente, uma vez que todas as variáveis independentes estejam consideradas. Em uma amostra coletada aleatoriamente da população, geralmente este é o caso.

No entanto, muitas vezes as observações têm alguma relação com os grupos dos quais esses indivíduos ou fenômenos pertencem. Para verificar se há dependência entre as observações e o contexto, ou seja, entre as categorias e os apoiadores dos projetos, devemos utilizar uma técnica chamada de correlação intraclasse (ICC, da sigla em inglês) que mede a proporção da variação em uma variável que ocorre entre grupos em comparação com o total da variação. O ICC varia de 0 (nenhuma variância entre grupos) a 1 (variância entre grupos, mas não dentro dos grupos) e, em um modelo de dois níveis, o ICC aponta o total de variabilidade dos resultados atribuído a uma determinada categoria (FIELD; MILES; FIELD, 2012; FINCH; BOLIN; KELLEY, 2019).

No caso que estamos analisando, consideramos os dados coletados como sendo hierarquizados em dois níveis, o dos projetos (nível 1) e o das categorias (nível 2) nas quais estão inscritos, o que pressupõe que projetos inscritos nas mesmas categorias sofrem efeitos similares entre si. O ICC estimado é de 0,22, o que significa que 22% da variação encontrada no número de apoiadores pode ser atribuída somente à categoria na qual são inscritos os projetos.

Outra maneira de se identificar se a variação entre distintos contextos deve ser considerada no modelo a ser construído, de acordo com Field, Miles e Field (2012), é através do ajuste de modelos simples, somente estimando o intercepto, e os comparando utilizando índices AIC e BIC e o método de máxima verossimilhança (-2LL). Esses indicadores nos mostram o quanto podemos confiar nos modelos ajustados a dada amostra. A comparação entre os modelos, detalhada na tabela 6, nos indica que aquele que se ajusta melhor aos dados é o multinível, com intercepto variando.

Tabela 6 - Diagnóstico de ajuste dos modelos multinível

Modelos	AIC	BIC	-2LL	Df	L. ratio..	p-value
Intercepto simples	65105.70	65121.01	65101.7 2			
Intercepto variável	62096.85	62119.81	62090.85	3	3010.851	$p < 0.001$

Fonte: Elaboração do autor

Posteriormente, o modelo foi construindo seguindo o processo sugerido por Field, Miles e Field (2012), partindo de um modelo simples, somente com o intercepto variando, e, incrementalmente, acrescentando efeitos fixos e efeitos variáveis.

À medida que os efeitos foram sendo incluídos, sua adequação foi analisada usando os métodos de máxima verossimilhança (log likelihood, -2LL), e a comparação de AICs e BICs. Ao final, foram ajustados cerca de 450 modelos diferentes, ainda que reportemos somente os três que nos parecem mais adequados para explicar os fenômenos estudados.

#### 4.4 Tratamento dos dados e descrição das variáveis

Após a coleta, percebemos que havia dados faltantes em algumas variáveis na base final (tabela 7). Considerando que esses dados eram *missing at random*, procedemos à imputação múltipla utilizando MICE (van Buuren; GROOTHUIS-OUDSHOORN, 2011)<sup>30</sup>. A estratégia utilizada foi a múltipla imputação multinível (MLMI, na sigla em inglês), utilizando como modelo o linear generalizado com efeitos mistos (2l.lmer), com efeitos como fixos e os interceptos aleatórios (JOLANI et al., 2015). As médias das variáveis após a imputação não sofreram alterações significativas (tabela 8).

Tabela 7 - Relação de dados faltantes

Tipo	Variável	Total	Percentual
numérico	projetos_contrib	1190	7,6%
numérico	projetos_publi	1190	7,6%
numérico	seguindo	1190	7,6%
numérico	seguidores	1190	7,6%
numérico	duracao	1061	6,8%
numérico	ano	1061	6,8%

Legenda: Relação de variáveis, com tipo de dado, total faltante e percentual de faltantes.

Fonte: Elaboração do autor

Tabela 8 - Resultados da imputação múltipla

	Médias antes	Médias depois
projetos_contrib	3.66	3.68
projetos_publi	2.78	2.76
seguindo	9.33	9.29
seguidores	4.90	4.97
duracao	65.24	65.37
ano	2015.71	2015.70

Legenda: Comparação entre as médias da variáveis que sofreram imputação múltipla, antes e depois do processo. Fonte: Elaboração do autor

<sup>30</sup> O diagnóstico da imputação pode ser encontrado no apêndice A

Depois da imputação, todas as variáveis contínuas foram centralizadas tal como sugerido por Gelman (2008), dividindo seus valores por duas vezes o seu desvio padrão. Com esse procedimento, a comparação entre a magnitude dos coeficientes se torna menos sensível às suas escalas, e reduz-se a colinearidade causada por termos de interação e por interceptos aleatórios fortemente correlacionados (FINCH; BOLIN; KELLEY, 2019). Além disso, como todas as variáveis estão centralizadas, é possível interpretar o intercepto, caso tenha significância, como sendo a média predita da variável dependente caso todas as outras variáveis estejam também em suas médias.

#### 4.5 Definição de variáveis e desenvolvimento das hipóteses.

Após a coleta e tratamento dos dados, testamos que elementos são determinantes para o apoio nos projetos. As variáveis analisadas foram definidas a partir do referencial bibliográfico, discutido com mais profundidade no segundo capítulo. Diferentemente da literatura, no entanto, estamos considerando como variável dependente o número de apoiadores por projeto. Nos parece mais adequado analisar a motivação dos apoiadores do que se as metas estabelecidas foram atingidas ou não. Um projeto com muitas doações, por exemplo, pode ser ou não bem-sucedido, mas tem elementos intrínsecos relevantes a ponto de mobilizar a comunidade, o que nos parece mais interessante do que o resultado financeiro.

Separamos as variáveis em quatro grupos:

- a) **Estratégias de relacionamento e divulgação.** Nossa expectativa é que os realizadores que desenvolvem planos de comunicação e relacionamento com a comunidade mais robustos tendem a ter mais apoiadores. Neste conjunto de variáveis, avaliaremos:
- **Atualizações:** refere-se ao número de vezes em que os realizadores entraram diretamente em contato com os apoiadores. Planos mais bem elaborados têm uma maior frequência de atualizações. Logo, estamos considerando ( $H_1$ ) que quanto maior a frequência de atualizações, maior o número de apoiadores.
  - **Perfis em sites de redes sociais:** os realizadores podem ampliar a visibilidade de seus projetos através do apoio de perfis em sites de redes sociais como o Facebook, o Twitter e o Instagram. Logo, um planejamento mais robusto considera o uso dessas plataformas ao longo das campanhas. Codificamos essa variável de duas maneiras: como *dummy* (0 como não tendo nenhum outro perfil, 1 como tendo) ou em relação ao número de perfis incluídos na descrição dos projetos, entre 0 e 5. Desta forma,

esperados que ( $H_2$ ) ter perfis em SRS, ou ter mais perfis em SRS, aumenta o número de apoiadores.

- **Descrição dos projetos:** Textos mais longos, com mais detalhes a respeito dos projetos, podem servir como medida para o preparo dos realizadores. Na estruturação da variável, medimos o tamanho do texto contando o número de palavras. Esperamos ( $H_3$ ) que quanto maior a descrição, mais apoiadores o projeto atrai.
- Estamos utilizando como variáveis de controle a magnitude e o score dos textos, obtidas pelo processamento das descrições pelo algoritmo de PNL do Google.

b) **Características dos projetos.** Algumas definições iniciais podem ter consequências significativas para o aumento no número de apoios recebidos ao longo das campanhas.

- **Número de recompensas:** Ainda que na essência do financiamento coletivo esteja a mobilização comunitária, acreditamos que as recompensas são parte importante dos incentivos para as doações. Por isso, acreditamos que ( $H_4$ ) quanto mais recompensas forem oferecidas, mais apoios recebem os projetos.
- **Meta:** Na literatura, a meta tem o potencial de servir como sinal de qualidade dos projetos indicando aos apoiadores se o realizador compreende o que está se propondo a fazer ou não. Com isso, ( $H_5$ ) metas muito altas teriam o potencial de afastar os apoiadores.
- **Modalidade:** Sabendo que na modalidade tudo-ou-nada os realizadores só obtêm os fundos se conseguirem um número maior de apoiadores, esperamos que eles apresentem planejamentos mais robustos do que os projetos flex. Desta forma, ( $H_6$ ), imaginamos que projetos tudo-ou-nada atraiam mais apoios, comparativamente.
- **Duração:** Considerando que os apoios tendem a acontecer mais no início e no final das campanhas, ( $H_7$ ) esperamos que projetos curtos demais ou longo demais tenham menos apoiadores.

c) **Características dos realizadores.** Consideramos que determinados realizadores têm características individuais que os tornam capazes de atrair mais apoios que os demais, tais como:

- **Número de seguidores:** o número de pessoas inscritas no Catarse que seguem seus projetos, independentemente de apoiá-los ou não, pode ser um indicador de importância dentro da comunidade. Assim, ( $H_8$ ) acreditamos que projetos de realizadores com mais seguidores têm mais apoios.

- **Número de projetos com que contribuiu:** Uma outra forma de medir o engajamento de um realizador com a comunidade de *crowdfunding* é através do número de projetos dos quais participou como apoiador. Por isso, ( $H_9$ ) acreditamos que um realizador que tenha contribuído com mais projetos atrai mais apoiadores que outros.
  - **Número de projetos publicados:** Espera-se que um realizador experiente, com mais campanhas já realizadas ( $H_{10}$ ), seja capaz de conseguir mais apoios para os seus projetos, não só por uma capacidade maior de se planejar, mas porque também tem uma rede de relacionamentos já construída na comunidade do Catarse.
- d) **Efeitos de geografia.** Mesmo que o financiamento coletivo seja uma prática típica da internet, e, em princípio, tenha em sua base a possibilidade de relacionamentos entre indivíduos desconhecidos e distantes uns dos outros, acreditamos que a geografia tem efeitos significativos na mobilização de apoios.
- **Coefficiente de Gini:** Acreditamos que projetos que tenham como origem as cidades com maior concentração de renda tenham mais apoiadores ( $H_{12}$ ).
  - **Densidade populacional:** Da mesma forma, acreditamos que cidades com mais densidade populacional tenham projetos com mais apoiadores ( $H_{13}$ ).

#### 4.6 Descrição das variáveis.

Os dados utilizados no estudo foram estruturados em uma base com 16 variáveis, descritas no apêndice C, e 15.595 observações. As matrizes de correlação, utilizando os coeficientes de Pearson e Spearman, podem ser encontradas no apêndice D.

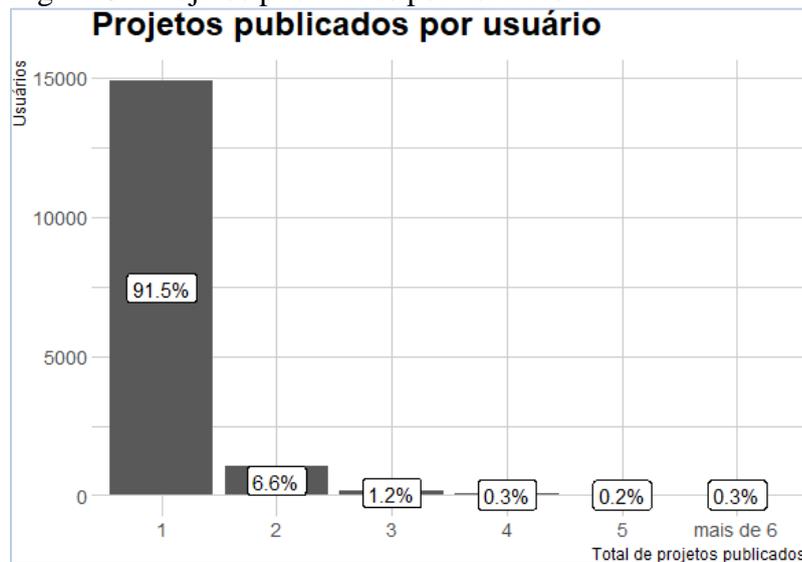
##### 4.6.1 Características dos usuários da plataforma

A maior parte dos que são cadastrados no Catarse tem muito pouco envolvimento com a plataforma, seja no que se refere à participação em campanhas ou na apresentação de projetos. São, em sua maioria, apoiadores<sup>31</sup> (54,3% do total) e não realizadores (1,7% do total), que tendem a participar somente de uma campanha (42,7% do total) ou publicar somente um projeto (91,5% dos realizadores) e não se envolvem mais com o site.

---

<sup>31</sup> Dos 952 mil usuários cadastrados, somente 517.471 fizeram contribuições.

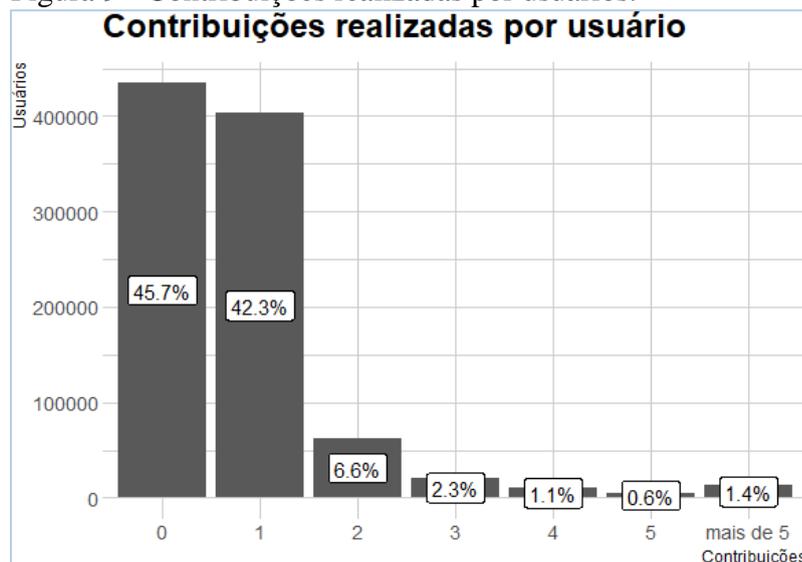
Figura 8 - Projetos publicados por usuários.



Legenda: Gráfico de contagem relacionando o número de projetos apresentados por usuários cadastrados no Catarse, filtrados por aqueles que apresentaram ao menos um, de 2011 a 2018. N=16.299

Fonte: Elaboração do autor.

Figura 9 - Contribuições realizadas por usuários.



Legenda: Gráfico de contagem relacionando o número de contribuições realizadas por usuários em projetos do Catarse, de 2011 a 2018. N = 952.772

Fonte: Elaboração do autor.

De maneira geral, os projetos recebem poucos apoios (5, considerando a mediana), mas esse comportamento varia muito de acordo com os projetos e com as categorias (tabela 9). Pouco mais de 71% das campanhas relacionadas a gastronomia, por exemplo, ficaram sem nenhuma contribuição, enquanto somente 7% dos projetos da categoria quadrinho não recebeu nenhum apoio. Essa diferença no comportamento pode ser percebida, ainda, ao observarmos a distribuição dos apoios por categoria (figura 10).

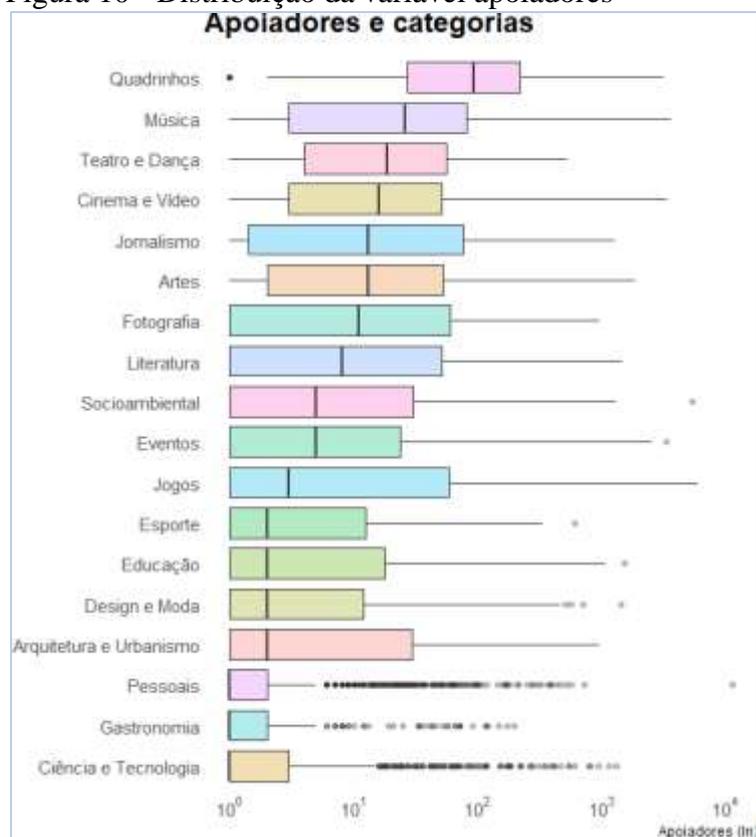
Tabela 9 - Relação de apoios por categoria

Categoria	Total	Sem apoio	% sem apoio	Com apoio	% com apoio
Gastronomia	259	186	71.8%	73	28.2%
Pessoais	1396	966	69.2%	430	30.8%
Ciência e tecnologia	822	544	66.2%	278	33.8%
Arquitetura e urbanismo	186	87	46.8%	99	53.2%
Design e moda	303	140	46.2%	163	53.8%
Esporte	491	221	45.0%	270	55.0%
Educação	1373	605	44.1%	768	55.9%
Jogos	441	179	40.6%	262	59.4%
Eventos	765	277	36.2%	488	63.8%
Socioambiental	1371	432	31.5%	939	68.5%
Fotografia	395	123	31.1%	272	68.9%
Literatura	1402	357	25.5%	1045	74.5%
Jornalismo	287	72	25.1%	215	74.9%
Artes	797	198	24.8%	599	75.2%
Música	2025	358	17.7%	1667	82.3%
Cinema e vídeo	1626	280	17.2%	1346	82.8%
Teatro e dança	954	128	13.4%	826	86.6%
Quadrinhos	702	51	7.3%	651	92.7%

Legenda: Relação entre o total de projetos, o total de projetos sem nenhum apoio e o total de projetos com ao menos uma contribuição, separados por categoria, entre 2011 e 2018 na plataforma Catarse

Fonte: Elaboração do autor

Figura 10 - Distribuição da variável apoiadores



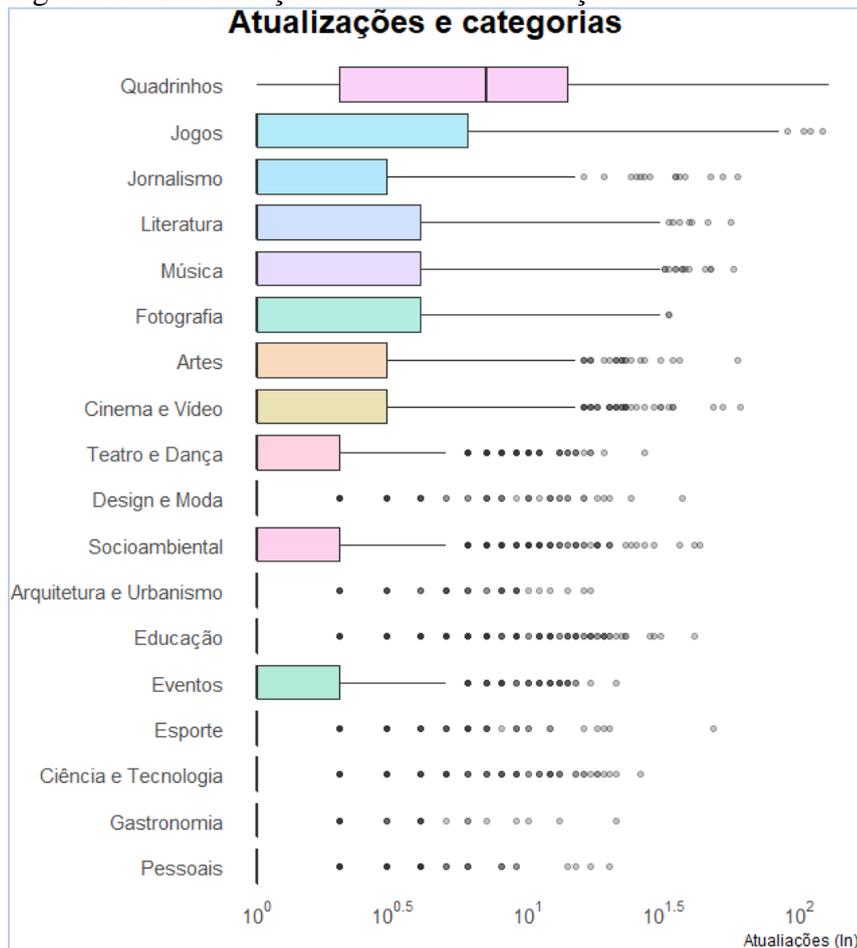
Legenda: Gráfico de caixa, com o número de apoiadores, considerando cada categoria, em escala logarítmica. N = 15.595.

Fonte: Elaboração do autor.

#### 4.6.2 Estratégias de divulgação

No que se refere às estratégias de divulgação, os realizadores tendem a enviar poucas atualizações (75% não envia nenhuma ao longo da campanha), quase sempre disponibilizam perfis em sites de redes sociais (69% têm ao menos um perfil na descrição) e têm descrições de até 640 palavras<sup>32</sup>. Há diferenças, no entanto, na maneira como os realizadores de diferentes categorias realizam sua divulgação. As atualizações dos projetos de quadrinhos, por exemplo, são mais frequentes (figura 11), com mais perfis em sites de redes sociais (figura 12) e descrições mais longas (figura 13).

Figura 11 - Distribuição da variável atualizações.



Legenda: Gráfico de caixa, com o número de atualizações de cada projeto, considerando cada categoria, em escala logarítmica. N = 15.595.  
Fonte: Elaboração do autor.

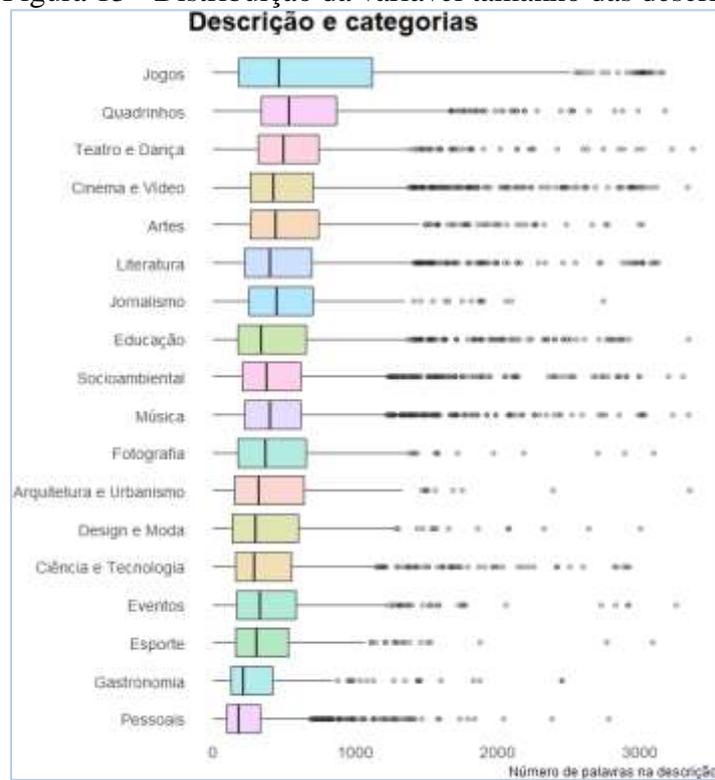
<sup>32</sup> Há, no entanto, uma dificuldade em se determinar exatamente o tamanho das descrições de todos os casos. Vários realizadores escolhem colocar uma imagem com a descrição do projeto ao invés de utilizar a ferramenta de edição de texto da plataforma, por conta da falta de recursos estéticos. Isso impede o *script* que raspou as informações de estruturar os textos, já que seria necessário importar todas as imagens e processá-las em algoritmos de reconhecimento de caracteres para depois se estimar o número de palavras, o que não era possível considerando os recursos disponíveis no momento deste estudo. Nos casos em que os usuários fizeram isso, as descrições vão aparecer na base como tendo 0, ou muito poucas, palavras.

Figura 12 - Distribuição do número de perfis em sites de redes sociais.



Legenda: Gráfico de caixa, com o número de perfis disponíveis em sites de redes sociais, considerando cada categoria. N = 15.595. Fonte: Elaboração do autor.

Figura 13 - Distribuição da variável tamanho das descrições.



Legenda: Gráfico de caixa, com tamanho das descrições dos projetos, em número de palavras, considerando cada categoria. N = 15.595. Fonte: Elaboração do autor.

#### 4.6.3 Características dos projetos

A respeito de suas características, em geral, considerando as medianas, os projetos oferecem 5 recompensas, pedem R\$ 10 mil reais para sua realização, duram o tempo recomendado pela plataforma, 60 dias, e não parece haver uma preferência entre as duas modalidades principais. No entanto, se observarmos cada uma dessas características de acordo com as categorias, percebemos que há diferenças significativas. Projetos de arquitetura e urbanismo, por exemplo, têm, em média, metas 10 vezes maiores e oferecem 2,5 x menos recompensas que a categoria quadrinhos (tabela 10). Enquanto os projetos Pessoais tendem a ser inscritos na modalidade Flexível, os realizadores da categoria Jogos preferem Tudo-ou-Nada (tabela 11).

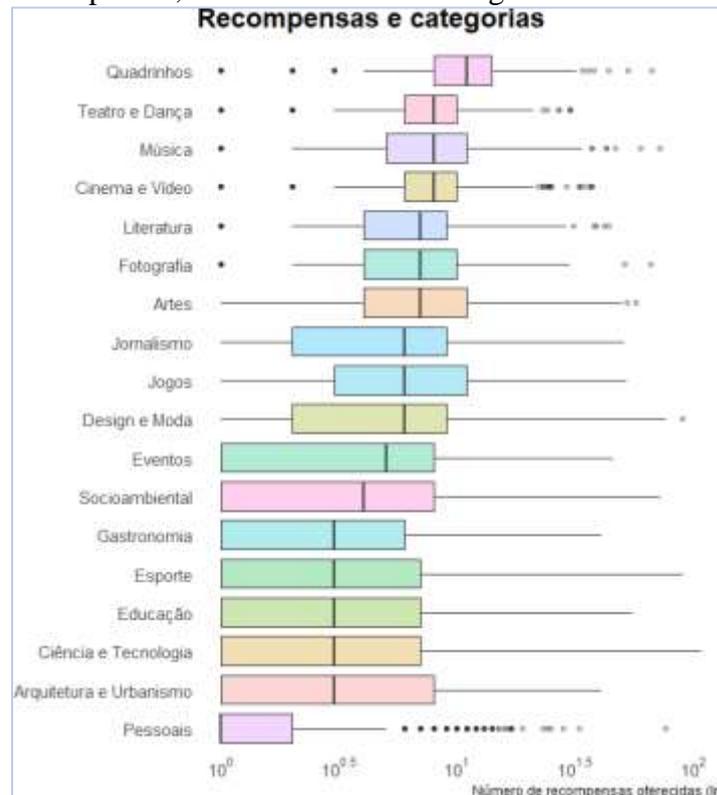
Tabela 10 - Características dos projetos

Categoria	Recompensas (n)	Meta (R\$)	Duração (dias)
Arquitetura e Urbanismo	4.2	R\$ 100.924,88	73.3
Artes	7.4	R\$ 25.289,35	64.8
Ciência e Tecnologia	3.9	R\$ 79.500,73	70.6
Cinema e Vídeo	6.9	R\$ 22.004,61	59.6
Design e Moda	6.3	R\$ 32.314,72	63.0
Educação	3.8	R\$ 38.662,57	69.6
Esporte	3.6	R\$ 24.910,37	68.1
Eventos	4.5	R\$ 28.718,29	51.6
Fotografia	6.8	R\$ 19.059,02	64.2
Gastronomia	3.6	R\$ 61.939,67	71.7
Jogos	7.0	R\$ 25.072,68	61.2
Jornalismo	6.0	R\$ 21.718,80	56.2
Literatura	6.1	R\$ 15.870,91	67.0
Música	7.8	R\$ 18.105,30	64.9
Pessoais	1.6	R\$ 46.084,71	73.0
Quadrinhos	10.5	R\$ 11.951,51	56.1
Socioambiental	4.3	R\$ 36.318,57	75.5
Teatro e Dança	7.1	R\$ 15.610,09	56.2

Legenda: Relação dos valores médios das variáveis relacionadas às características dos projetos, separados por categoria, dos projetos do Catarse, de 2011 a 2018.

Fonte: Elaboração do autor.

Figura 14 - Distribuição da variável número de recompensas, considerando cada categoria



Legenda: Gráfico de caixa com a distribuição do número de recompensas oferecidas, em escala logarítmica. N = 15.595.

Fonte: Elaboração do autor.

Tabela 11 - Modalidade dos projetos por categoria

Categoria	Flexível	Tudo-ou-nada
Quadrinhos	17.9% (126)	82.1% (576)
Jogos	31.3% (138)	68.7% (303)
Cinema e Vídeo	36.5% (594)	63.5% (1032)
Literatura	36.8% (516)	63.2% (886)
Música	38.4% (778)	61.6% (1247)
Teatro e Dança	40.3% (384)	59.7% (570)
Design e Moda	41.9% (127)	58.1% (176)
Jornalismo	43.2% (124)	56.8% (163)
Fotografia	43.8% (173)	56.2% (222)
Total	48.9% (7629)	51.1% (7966)
Artes	52.2% (416)	47.8% (381)
Socioambiental	54.4% (746)	45.6% (625)
Ciência e Tecnologia	54.5% (448)	45.5% (374)
Arquitetura e Urbanismo	54.8% (102)	45.2% (84)
Eventos	60.8% (465)	39.2% (300)
Esporte	63.5% (312)	36.5% (179)
Gastronomia	64.1% (166)	35.9% (93)
Educação	66.6% (914)	33.4% (459)
Pessoais	78.8% (1100)	21.2% (296)
Total	48.9% (7629)	51.1% (7966)

Legenda: Relação da modalidade na qual os projetos são inscritos, separados por categoria, na plataforma Catarse, de 2011 a 2018. N = 15.595.

Fonte: Elaboração do autor.

#### 4.6.4 Características dos realizadores

Já os realizadores têm um envolvimento restrito com a plataforma. Em geral, seguem poucos usuários (80% segue até 1 outro inscrito), são pouco seguidos (75% tem até 2 seguidores), contribuem pouco com outros projetos (75% participou de até três campanhas) e publicam apenas uma vez. Novamente, esse comportamento varia de acordo com as categorias. Enquanto um realizador da categoria quadrinho, em média, tem 42 seguidores, contribui com 21 campanhas e publica dois projetos, o da categoria jornalismo tem 3 seguidores, segue 0,7 usuários, contribui com 3 campanhas e publica 1,5 projetos (tabela 12).

Tabela 12 - Características dos realizadores

Categoria	Seguidores	Contribuições	Seguindo	Projetos Publicados
Arquitetura e Urbanismo	2.5	7.2	11.4	1.5
Artes	3.6	3.3	3.8	1.4
Ciência e Tecnologia	1.7	1.8	3.6	1.4
Cinema e Vídeo	3.2	3.2	2.1	2.3
Design e Moda	2.8	4.4	7.4	1.3
Educação	1.6	2.1	2.9	1.4
Esporte	1.8	1.7	5.1	1.8
Eventos	2.4	2.5	2.2	2.0
Fotografia	5.4	3.0	19.9	1.6
Gastronomia	1.0	1.5	2.0	1.2
Jogos	22.8	4.4	11.2	2.3
Jornalismo	3.3	3.0	0.7	1.5
Literatura	4.9	3.7	7.8	1.6
Música	2.0	2.7	2.7	1.4
Pessoais	1.1	1.9	5.3	1.5
Quadrinhos	42.3	21.0	109.4	2.2
Socioambiental	3.5	3.5	5.1	14.3
Teatro e Dança	1.4	2.7	2.0	1.4

Legenda: Valores médios das variáveis associadas às características dos realizadores da plataforma Catarse, considerando cada categoria, de 2011 a 2018. Seguidores refere-se ao número de usuários que seguem os realizadores; Contribuições refere-se ao número de projetos com os quais os realizadores contribuíram; Seguindo refere-se ao número de outros usuários que são seguidos pelos realizadores; Projeto publicados refere-se ao número de projetos já propostos pelos realizadores.

Fonte: Elaboração do autor

#### 4.6.5 Efeitos de geografia

Do ponto de vista da distribuição geográfica, identificamos projetos oriundos de 1.214 municípios, de todos os estados brasileiros e do Distrito Federal (tabela 13). Há, no entanto, mais projetos das cidades de São Paulo e do Rio de Janeiro (38% do total, somados) e, mais especificamente, nas capitais dos estados.

Tabela 13 - Projetos por cidade

Cidade	Projetos	Percentual
São Paulo	4227	27%
Rio de Janeiro	1702	11%
Porto Alegre	629	4%
Belo Horizonte	589	4%
Curitiba	529	3%
Brasília	499	3%
Florianópolis	289	2%
Salvador	283	2%
Fortaleza	275	2%
Recife	260	2%
Campinas	234	2%
Goiânia	159	1%
Natal	130	1%
Manaus	126	1%
Outros	5664	36%
<b>Total</b>	<b>15595</b>	<b>100%</b>

Legenda: Número de projetos apresentados, separados por município, de 2011 a 2018 na plataforma Catarse. Fonte: Elaboração do autor.

Como a literatura indica, há uma concentração de projetos em regiões de maior participação no PIB (tabela 14 e figura 15), como São Paulo, Rio de Janeiro e Minas Gerais, locais com maior densidade demográfica e cujos usuários têm mais disponibilidade financeira, estão mais habituados a realizar transações online. Além disso, são áreas com mais atividade da indústria criativa.

Figura 15 - Localização e concentração geográfica dos projetos



Legenda: Mapas com a localização dos municípios de origem e concentração por estado dos projetos publicados no Catarse, de 2011 a 2018

Fonte: Elaboração do autor.

Tabela 14 - Projetos por Estado

Estado	Projetos	Percentual	% PIB
SP	6.288	40%	31,4%
RJ	2.066	13%	10,0%
MG	1.273	8%	8,9%
RS	1.038	7%	7,0%
PR	792	5%	6,8%
SC	648	4%	4,1%
BA	520	3%	4,2%
DF	499	3%	2,6%
PE	404	3%	2,4%
CE	366	2%	2,0%
GO	266	2%	2,9%
ES	198	1%	2,2%
RN	190	1%	9,0%
PB	168	1%	0,8%
PA	159	1%	2,5%
AM	140	1%	1,5%
AL	88	1%	0,9%
MT	85	1%	2,3%
MA	80	1%	1,4%
MS	73	0%	1,7%
PI	70	0%	0,6%
SE	64	0%	0,6%
RO	43	0%	0,6%
AC	33	0%	0,2%
TO	19	0%	0,5%
AP	16	0%	0,2%
RR	9	0%	0,1%
<b>Total</b>	<b>15595</b>	<b>100%</b>	

Legenda: Relação de projetos apresentados de 2011 a 2018 na plataforma Catarse, por estado e contribuição do PIB

Fonte: Elaboração do autor. Participação de cada UF na composição do PIB nacional, referência 2015, IBGE (2018a)

As características regionais se apresentam, ainda, na categoria dos projetos. Regiões nas quais a indústria criativa é mais importante, como Rio de Janeiro e São Paulo, têm projetos mais ligados às artes, como Cinema e Vídeo, Música e Teatro e Dança. No Amapá, que tem um grande número de áreas de conservação ambiental, 43% dos projetos são da categoria Socioambiental (figura 16).

Figura 16 - Principais categorias por estado



Legenda: Distribuição regional dos projetos publicados no Catarse, de 2011 a 2018, ordenados por categoria de maior incidência, nos estados da federação.

Fonte: Elaboração do autor.

#### 4.7 Determinantes do apoio em projetos do Catarse

O modelo que utilizamos para testar quais variáveis são determinantes para o apoio nos projetos pode ser formalizado da seguinte maneira:

$$\begin{aligned}
\text{apoiadores}_{ij} = & \beta_0j + \beta_{1j}(\text{n\_recompensas\_ln})_{ij} + \beta_2(\text{n\_recompensas\_ln}^2)_{ij} + \beta_{3j}(\text{meta\_ln})_{ij} + \\
& \beta_{4j}(\text{atualiza\c{c}oes\_ln})_{ij} + \beta_5(\text{atualiza\c{c}oes\_ln}^2)_{ij} + \beta_6(\text{dura\c{c}ao\_ln})_{ij} + \\
& \beta_7(\text{dura\c{c}ao\_ln}^2)_{ij} + \beta_8(\text{modalidade\_tudounada})_{ij} + \beta_9(\text{redes\_sociais\_sim})_{ij} + \\
& \beta_{10}(\text{n\_redes\_sociais}^2)_{ij} + \beta_{11}(\text{desc\_palavras})_{ij} + \beta_{12}(\text{desc\_palavras}^2)_{ij} + \\
& \beta_{13}(\text{seguidores\_ln})_{ij} + \beta_{14}(\text{seguidores\_ln}^2)_{ij} + \beta_{15j}(\text{seguindo\_ln})_{ij} + \\
& \beta_{16}(\text{seguindo\_ln}^2)_{ij} + \beta_{17}(\text{projetos\_contrib\_ln})_{ij} + \beta_{18}(\text{projetos\_contrib\_ln}^2)_{ij} + \\
& \beta_{19}(\text{projetos\_publi\_ln})_{ij} + \beta_{20j}(\text{magnitudo\_ln})_{ij} + \beta_{21}(\text{score})_{ij} + \\
& \beta_{22}(\text{gini\_cidade})_{ij} + \beta_{23}(\text{densidade})_{ij} + \beta_{24}(\text{densidade}^2)_{ij} + \\
& \beta_{25}(\text{ano}_{2012})_{ij} + \beta_{26}(\text{ano}_{2013})_{ij} + \beta_{27}(\text{ano}_{2014})_{ij} + \beta_{28}(\text{ano}_{2015})_{ij} + \\
& \beta_{29}(\text{ano}_{2016})_{ij} + \beta_{30}(\text{ano}_{2017})_{ij} + \beta_{31}(\text{ano}_{2018})_{ij} + E_{ij}
\end{aligned}$$

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\beta_{1j} = \beta_1 + u_{1j}$$

$$\beta_{3j} = \beta_3 + u_{3j}$$

$$\beta_{4j} = \beta_4 + u_{4j}$$

$$\beta_{15j} = \beta_{14} + u_{14j}$$

$$\beta_{20j} = \beta_{19} + u_{19j}$$

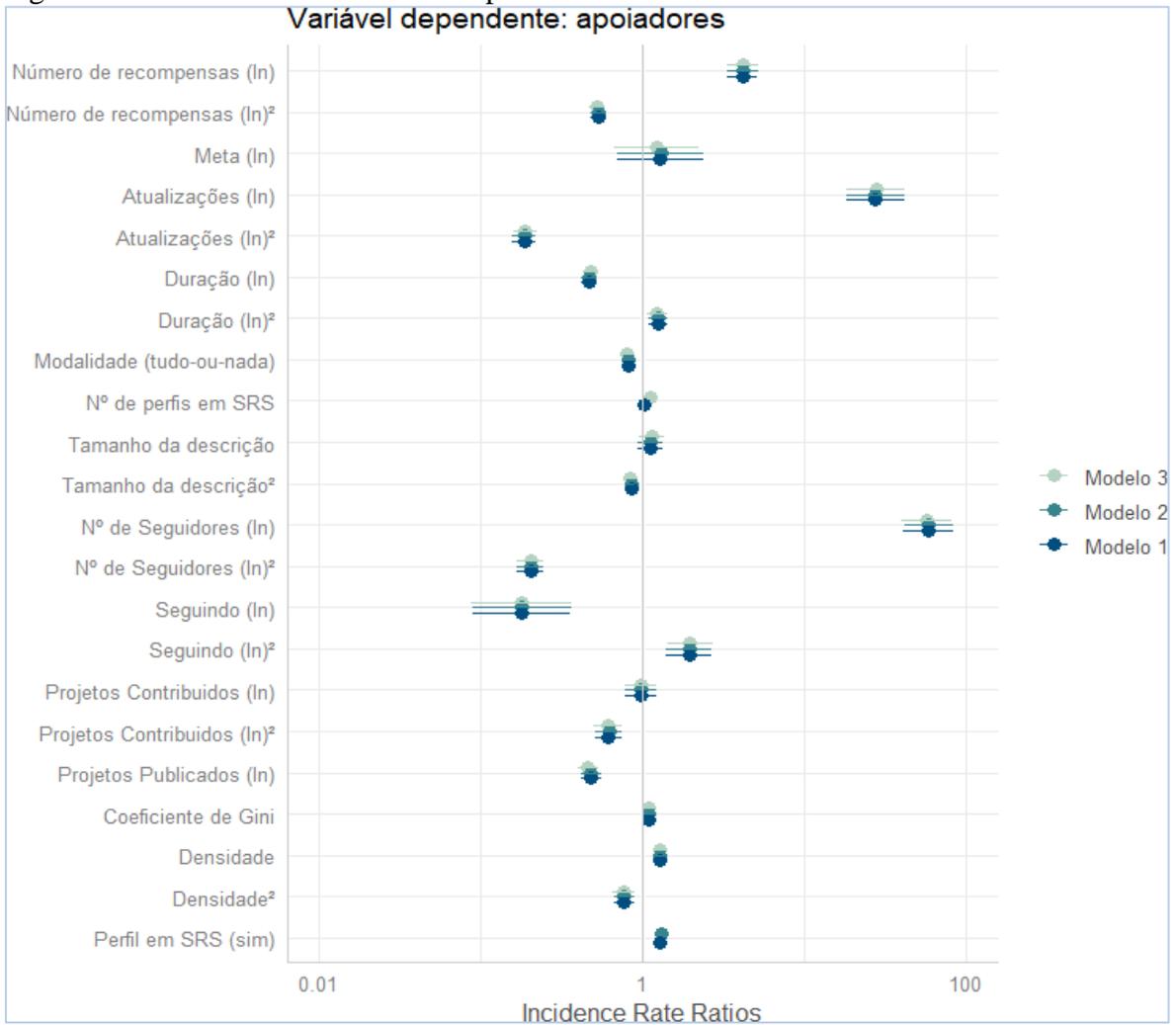
(1)

Neste caso, estamos assumido  $i$  como sendo os projetos,  $j$  as categorias e  $u$  a variabilidade dos preditores. Desta forma, o intercepto e os *slopes* das variáveis  $n$  *recompensas(ln)*, *meta(ln)*, *atualizacoes(ln)*, *seguindo(ln)* e *magnitudo(ln)* variam de acordo com as categorias nas quais os projetos estão inscritos, enquanto que as demais variáveis se mantêm fixas. As variáveis *ano*, *magnitudo* e *score* estão sendo utilizadas como controle.

Os modelos estimados conseguem explicar cerca de 70% da variação no número de apoiadores, o que é muito significativo. Na figura 17 estão apresentados os preditores principais, exceto as variáveis de controle. A tabela com todos os coeficientes pode ser encontrada no apêndice E<sup>33</sup>.

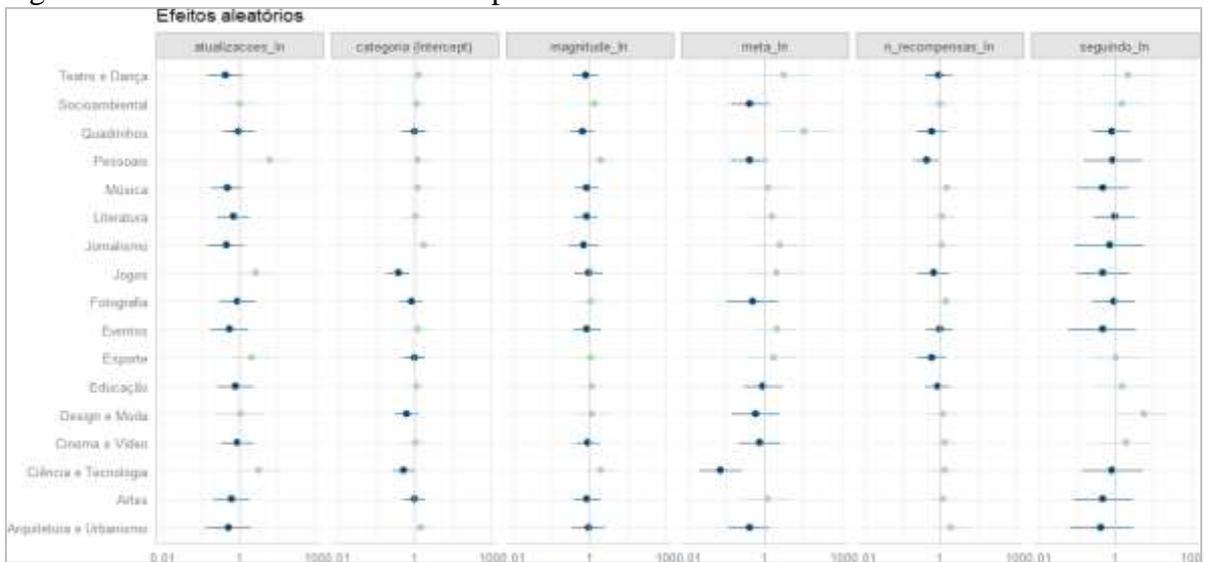
<sup>33</sup> Os modelos foram testados, ainda, para multicolinearidade e estimados a normalidade da distribuição dos efeitos aleatórios e a normalidade da distribuição dos resíduos. Os gráficos com os diagnósticos podem ser encontrados no apêndice G

Figura 17 - Preditores do número de apoiadores



Legenda: Preditores estimados do número de apoiadores dos projetos de *crowdfunding* no Catarse.  
 Fonte: Elaboração do autor.

Figura 18 - Preditores do número de apoiadores - efeitos aleatórios



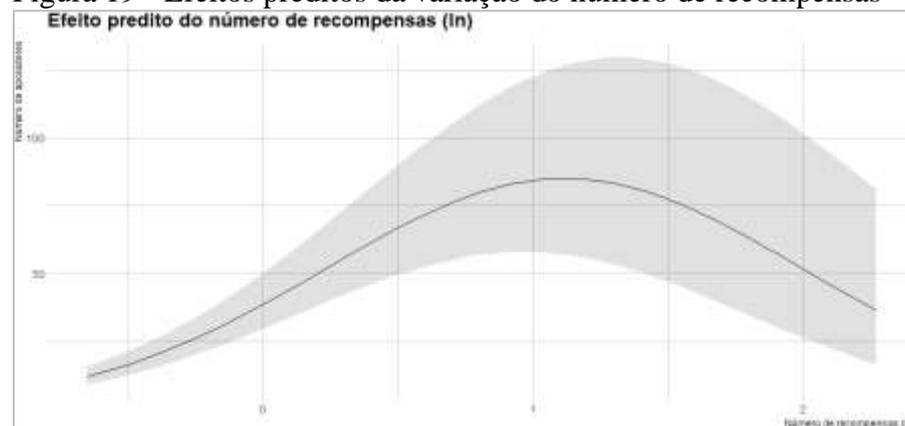
Legenda: Efeitos aleatórios estimados do número de apoiadores dos projetos de *crowdfunding* no Catarse, referentes ao modelo 2.  
 Fonte: Elaboração do autor.

Como todas as variáveis estão padronizadas, ou seja, com suas médias subtraídas e divididas por duas vezes seu desvio-padrão, o intercepto, que tem significância, torna-se interpretável. Nos nossos modelos, o intercepto corresponde à média predita de apoiadores considerando que todas as outras variáveis estão também suas médias. Neste caso, cada projeto tem entre 65 e 67 apoiadores, considerando um intervalo de confiança de 95%<sup>34</sup>.

O número de recompensas é um preditor significativo; se a oferta de recompensas aumenta, controlada por todas as demais variáveis, o número de apoiadores também aumenta (figura 19). No entanto, esse efeito vai perdendo força à medida em que o número de recompensas vai aumentando e se torna negativo a partir de certo ponto, o que significa que o incremento de recompensas a partir daí é acompanhado de uma redução do número de apoiadores.

Mais especificamente, se considerarmos um projeto que oferece até -0,05 desvio- padrão da média de recompensas, o que corresponde a 5 recompensas, ou percentil 50, os modelos preveem que ele obterá entre 28 e 47 apoiadores, aproximadamente. Da mesma forma, um projeto que oferece 0,25 desvio-padrão da média de recompensas, o que corresponde a 9 recompensas, pode obter entre 41 e 71 apoiadores. O ponto de inflexão da curva, ou seja, o ponto em que a relação entre as variáveis passa ser negativa, situa-se em 1,17 desvio-padrão, ou 32 recompensas. Neste caso, os modelos preveem entre 57 e 128 apoiadores.

Figura 19 - Efeitos preditos da variação do número de recompensas



Legenda: Valores estimados do efeito da variação do número de recompensas, em logaritmo, no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

A variável meta não tem significância em nenhum dos modelos estimados, diferentemente do que acontece em modelos que utilizam a variável sucesso como dependente.

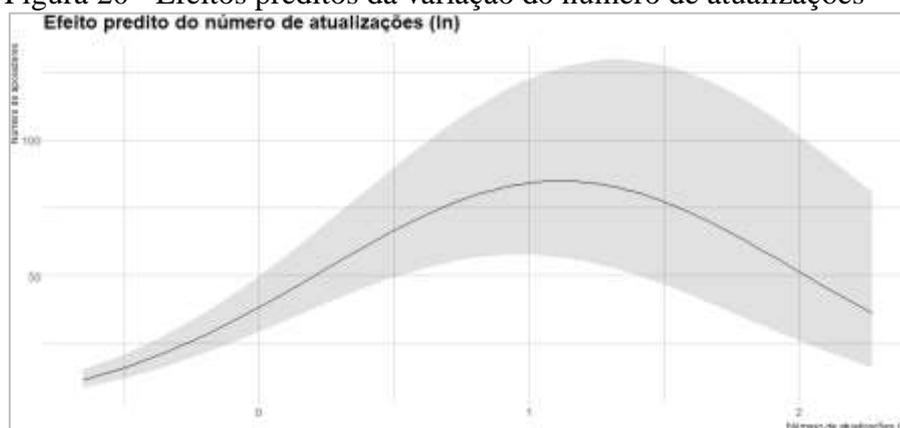
<sup>34</sup> Todos os intervalos apresentados na interpretação dos modelos consideram 95% de confiança.

Significa que quando consideramos o comportamento dos apoiadores, não se pode afirmar que a meta os atrai ou não, o que faz sentido se, como nos aponta a literatura, os apoiadores estiverem interessados, acima de tudo, nas causas ou nas recompensas oferecidas. Nesse caso, não importa muito a meta, mas as demais qualidades das campanhas. Nos modelos que utilizam sucesso como variável dependente, a meta ter uma correlação negativa com o sucesso pode ter relação com a dificuldade, crescente, em se obter montantes maiores.

O número de atualizações enviadas pelos realizadores, por outro lado, tem efeito significativo no número de apoiadores (figura 20). Controlando por todas as demais variáveis, quanto mais atualizações tem um projeto, mais apoios recebe. Esse efeito, no entanto, também vai perdendo força à medida em que o número de atualizações aumenta, e se torna negativo a partir de certo ponto.

Um projeto que tenha -0,01 desvio-padrão da média de atualizações, ou 1 atualização, o que corresponde ao percentil 70 da distribuição, receberá, de acordo com os modelos, entre 31 e 52 apoios. Já aqueles que publicaram 0.35 desvio-padrão de atualizações, ou 7 atualizações, o que corresponde ao percentil 90, recebem entre 84 e 142 apoios, considerando o mesmo intervalo de confiança. Os modelos consideram que o número ótimo de atualizações, ou seja, o ponto de inflexão da reta, situa-se em 0,97 desvio-padrão, ou 15 atualizações, situação em que o projeto pode receber entre 142 e 319 apoios.

Figura 20 - Efeitos preditos da variação do número de atualizações



Legenda: Valores estimados do efeito da variação do número de atualizações, em logaritmo, no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

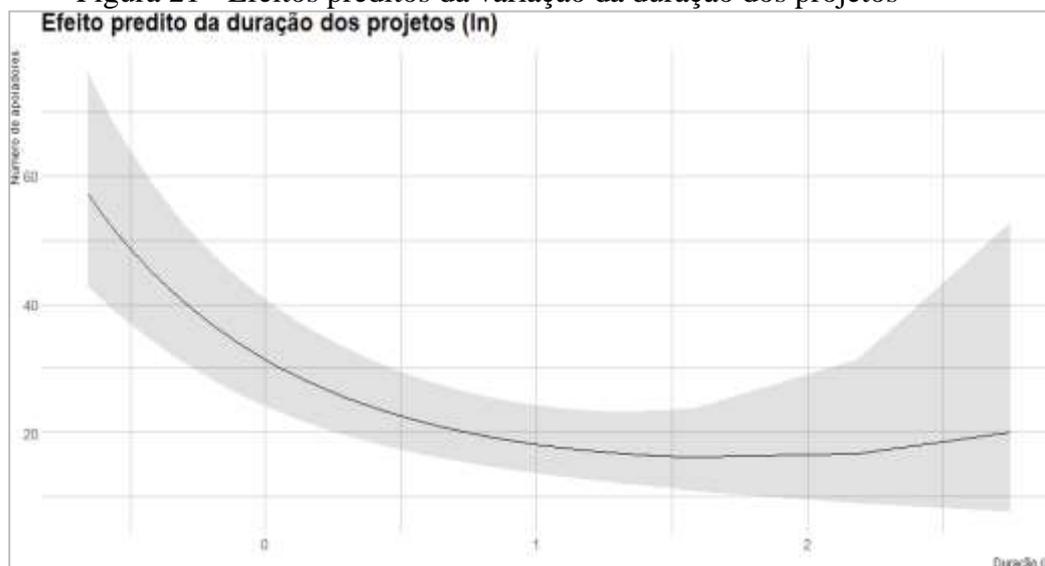
Nos modelos estimados, a duração dos projetos é também uma variável significativa. Se controlarmos por todas as demais variáveis, um projeto com duração menor tende a ter mais apoiadores e esse efeito é atenuado à medida em que aumenta a duração, ou seja, quanto mais tempo passa, o efeito da duração se torna positivo e o número de apoios tende a aumentar (figura

21). Isso nos indica que os apoiadores tendem a contribuir no início ou no final da duração dos projetos, o que confirma achados similares em outros estudos.

Projetos com duração de  $-0,41$  desvio-padrão da média, ou seja, 20 dias, o que corresponde ao percentil 10 da distribuição, deverão ter entre 35 e 61 apoiadores. Já os projetos com duração de 60 dias, ou seja, que estejam em 0 desvio-padrão, o que representa o percentil 80, tendem a ter entre 25 e 43 apoiadores.

A escolha da modalidade é, também, significativa. Em comparação com os projetos do tipo Flexível, escolher Tudo-ou-nada, se controlarmos por todas as demais variáveis, tem um efeito negativo no número de apoiadores. Mais especificamente, projetos Tudo- ou-nada têm, aproximadamente, 18% menos chances de ter apoiadores do que os projetos Flex.

Figura 21 - Efeitos previstos da variação da duração dos projetos



Legenda: Valores estimados do efeito da variação da duração, em logaritmo, no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

Ter perfis em sites de redes sociais (SRS) é, também, uma variável significativa. Nos modelos em que codificamos como variável binária, ter um perfil, em comparação com não ter, se controlada por todas as outras variáveis, tem uma relação positiva com o número de apoiadores. Especificamente, projetos que apresentam perfis têm 34% de chance de ter mais apoiadores do que aqueles que não têm.

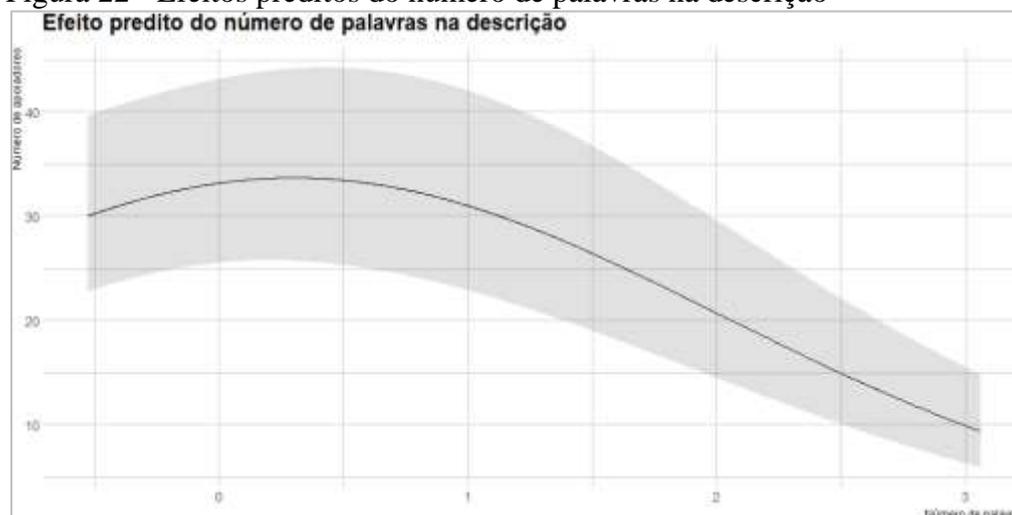
No modelo em que codificamos a variável como contínua, ou seja, em que o número de perfis variou de 0 a 5, apresentar mais perfis tem efeito positivo no número de apoiadores, se controlarmos as demais variáveis. A cada perfil apresentado, o projeto tem 10% de chances de ter mais apoiadores. Quando incluímos ambas as variáveis no mesmo modelo, o número de

perfis em sites de redes sociais perde a significância, enquanto ter ou não um perfil se mantém como significativa.

Para avaliar a importância do texto de descrição dos projetos, estimamos as variáveis tamanho (em número de palavras), magnitude (o quão emocional é o texto) e score (se positivo, negativo ou neutro), sendo que as duas últimas, estimadas pelo algoritmo de PLN do Google, são consideradas variáveis de controle.

O tamanho da descrição só tem significância a partir de determinado tamanho e uma relação negativa com o número de seguidores, ou seja, textos longos demais tendem a reduzir o número de apoiadores, se controlarmos por todas as variáveis, mas não podemos afirmar se há algum efeito para os casos em que as descrições são muito curtas (figura 22).

Figura 22 - Efeitos preditos do número de palavras na descrição



Legenda: Valores estimados do efeito do tamanho das descrições, em número de palavras, no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

O tamanho ótimo para a descrição situa-se a 0,25 desvio-padrão da média do número de palavras, ou 732 palavras, o que corresponde ao percentil 80 da distribuição. Esse projeto teria, segundo os modelos, entre 25 e 44 apoiadores. Já um projeto que esteja 3 desvios-padrão da média do número de palavras, o que representa 3.369 palavras, teria entre 5 e 14 apoiadores.

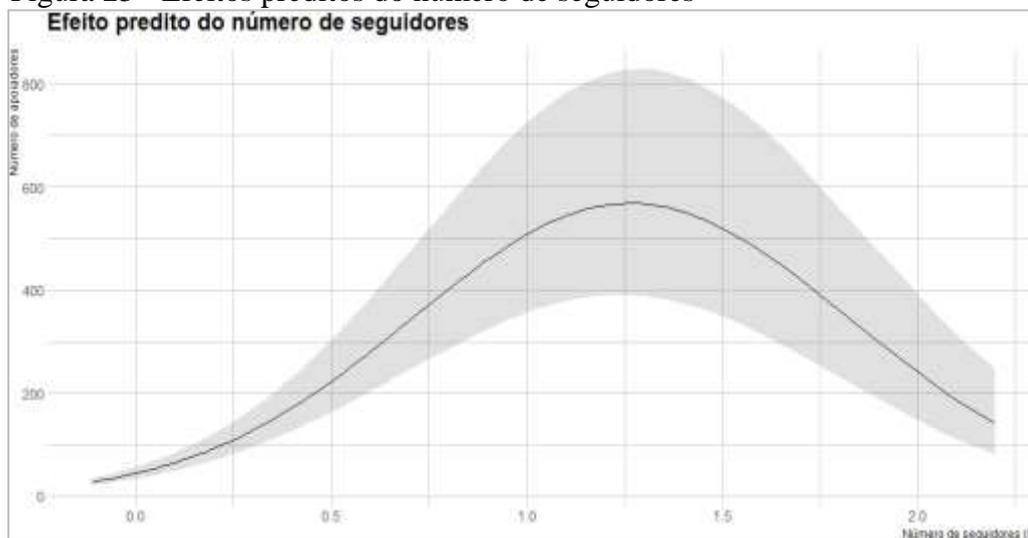
Para avaliar a importância das redes pessoais do realizador, três variáveis foram analisadas, o número de seguidores que cada um deles possui na plataforma do Catarse, o número de pessoas que cada realizador segue e com quantos projetos o realizador contribuiu na plataforma. Em tese, quanto mais envolvido com a comunidade em torno da plataforma, mais apoiadores teria um projeto deste realizador.

Nos nossos modelos, quanto mais seguidores tem o realizador, se controladas por todas as outras variáveis, mais aumenta o número de apoiadores de seus projetos (figura 23). Essa

relação também não é linear, o que significa que, à medida em que aumenta o número de seguidores, o efeito da variável diminui e se torna negativo a partir de certo ponto.

Um realizador que tenha 3 seguidores, o que é o equivalente a  $-0,04$  desvio-padrão da média e corresponde ao percentil 80 da distribuição, tende a ter de 25 a 42 apoiadores em seu projeto. Já aquele que tenha 7 seguidores, que se situa a  $0,04$  desvio-padrão, tende a ter entre 35 e 60 apoios. A partir de 65 seguidores, o que representa um desvio-padrão de  $1,28$ , o número de apoiadores tende ser entre 357 a 757. A partir deste ponto, a relação se inverte. Um realizador com  $2.20$  desvio-padrão da média do número de seguidores, o que corresponde a  $384$ , teria entre 75 e 230 apoiadores em seus projetos.

Figura 23 - Efeitos preditos do número de seguidores



Legenda: Valores estimados do efeito do número de usuários, em logaritmo, que seguem os realizadores no número de apoiadores dos projetos de *crowdfunding* no Catarse.

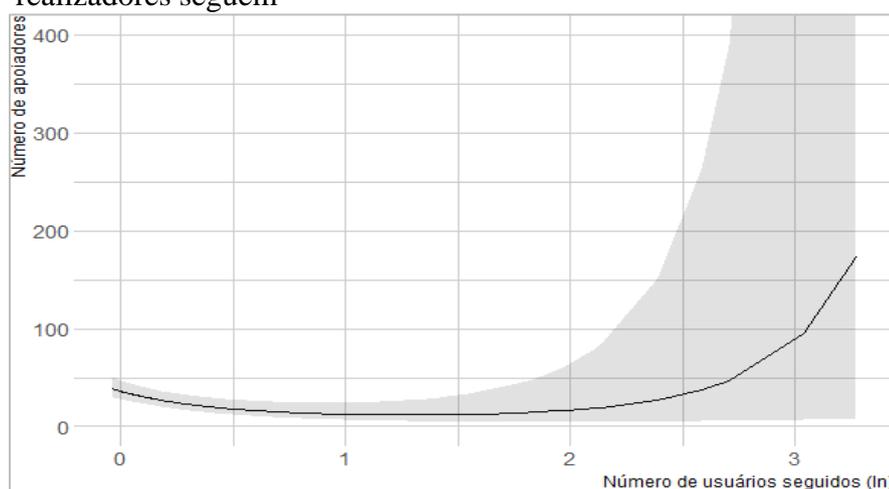
Fonte: Elaboração do autor.

Já o número de usuários que o realizador está seguindo tem uma relação negativa com o número de apoiadores em seus projetos, e essa relação se inverte à medida em que o realizador segue mais usuários (figura 24).

Um realizador que segue 1 usuário, o que representa  $-0,03$  desvio-padrão da média e o percentil 90 da distribuição, teria entre 29 e 49 apoiadores. A relação se inverte a partir do momento em que os realizadores mais de 584 usuários, ou  $1,21$  desvio-padrão da média. Neste caso, seus projetos teriam entre 5 e 25 apoiadores.

Por outro lado, o número de projetos com os quais o realizador já contribuiu só tem significância a partir de certo ponto, quando a relação se torna negativa, ou seja, quando o número de contribuições tem efeito negativo no número de apoiadores (figura 25).

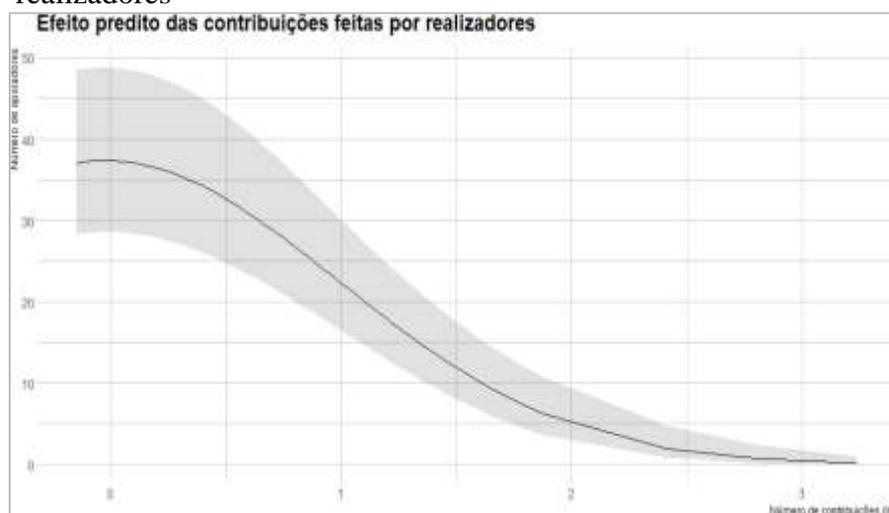
Figura 24 - Efeitos preditos do número de usuários que os realizadores seguem



Legenda: Valores estimados do efeito do número de usuários, em logaritmo, que os realizadores seguem no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

Figura 25 - Efeitos preditos do número de contribuições feitas por realizadores



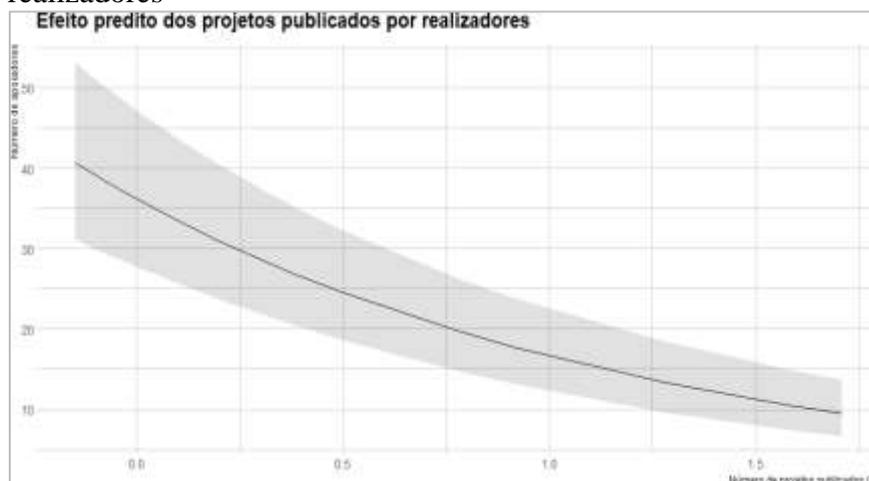
Legenda: Valores estimados do efeito do número de contribuições, em logaritmo, feitas por realizadores no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

Um realizador que tenha contribuído com 9 projetos, ou seja, no percentil 90 da distribuição, teria entre 25 e 43 apoiadores. Já um realizador que tenha contribuído com 16 projetos, o que representa 0.39 desvio-padrão e se situa no percentil 95, teria entre 24 e 41 apoiadores.

A variável número de projetos publicados tem uma relação negativa com o número de apoiadores, o que significa dizer que, se controlados pelas demais variáveis, o incremento no número de projetos publicados reduz em 51% as chances de se atrair apoiadores (figura 26).

Figura 26 - Efeitos preditos do número de projetos já publicados por realizadores



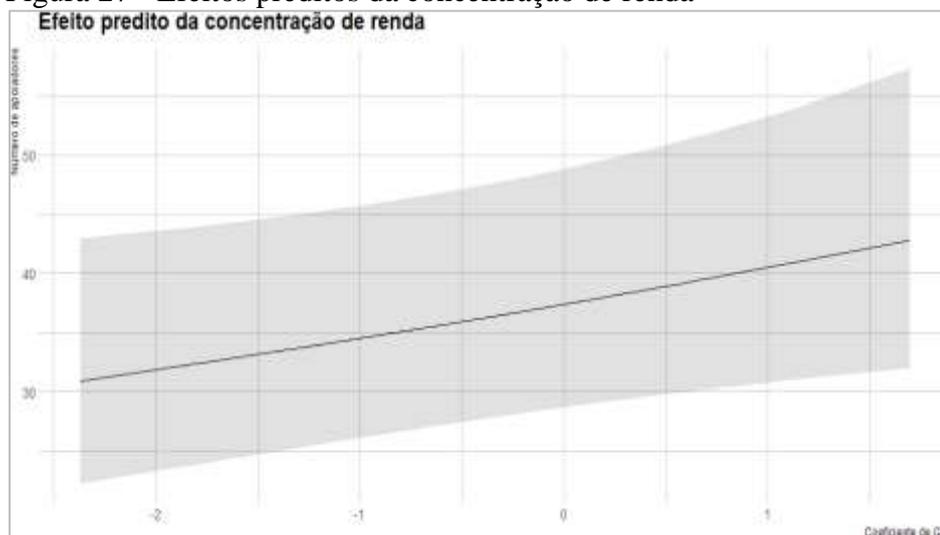
Legenda: Valores estimados do efeito do número de projetos, em logaritmo, publicados anteriormente por realizadores no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

Essa relação nos sugere que os apoiadores se interessam mais por projetos que tragam novidades do que por atualizações dos mesmos eventos, produtos ou serviços. Ainda que a possibilidade de ser bem-sucedido aumente com a experiência do realizador, com sugere a literatura, seus projetos subsequentes tendem a ter menos apoiadores do que os primeiros.

No caso da concentração de renda, o incremento de 1 desvio-padrão no coeficiente de Gini das cidades de origem dos projetos aumenta em cerca de 10% o número de apoiadores, se controladas todas as demais variáveis (figura 27). Isso significa que projetos de cidades com mais concentração de renda tendem a ter mais apoiadores que projetos menos desiguais.

Figura 27 - Efeitos preditos da concentração de renda

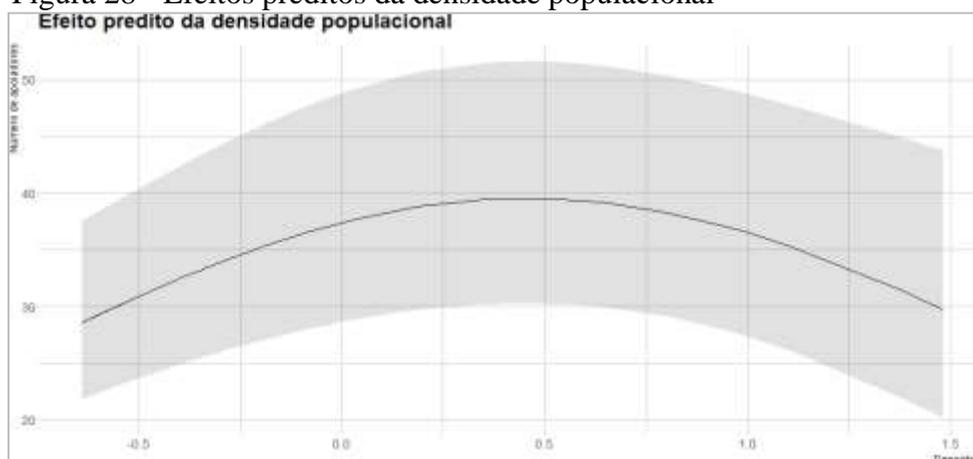


Legenda: Valores estimados do efeito da concentração de renda nas cidades de origem dos projetos no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

Por fim, a densidade populacional tem uma relação inicialmente positiva com o número de apoiadores, mas essa relação se inverte a partir de determinado ponto. Projetos de cidades com densidade média, ou 0 desvio-padrão, tendem a ter entre 28 e 48 apoiadores, controladas todas as demais variáveis. Cidades com 1,5 desvio-padrão da média de densidade teriam entre 19 e 43 apoiadores, assim como aquelas com -1 desvio-padrão da média de densidade deverão ter entre 16 e 30 apoios.

Figura 28 - Efeitos preditos da densidade populacional



Legenda: Valores estimados do efeito da densidade populacional das cidades de origem dos projetos no número de apoiadores dos projetos de *crowdfunding* no Catarse.

Fonte: Elaboração do autor.

## CONCLUSÃO

Quando iniciamos este estudo, nos questionávamos sobre o que leva alguém a dar seu apoio a um projeto de financiamento coletivo e não a outro. Nossa hipótese era a de que havia elementos constitutivos dos projetos e características dos propositores que tornavam sua proposta mais atraente do que outras, e que poderíamos identificar essas características analisando as informações disponíveis nos *hotsites* das campanhas publicados na plataforma do Catarse.

Essa formulação foi possível por conta do momento em que vivemos. A internet e o avanço dos computadores pessoais permitiram que não só houvesse uma produção imensa de dados a respeito do comportamento dos indivíduos, mas que, individualmente, um pesquisador seja capaz de coletar, organizar e processar essas informações, utilizando técnicas inacessíveis há poucos anos. Nesse sentido, vivemos uma transformação drástica no ambiente de pesquisa, que abre caminho para inúmeras inovações e oportunidades possíveis nas ciências humanas.

Esse contexto é, nesse sentido, desafiador, porque exige que o pesquisador que se aventure em campos que não são os seus de origem. Quando um estudante de comunicação se depara com formulações estatísticas, é automaticamente desincentivado a se embrenhar na complexidade e na dureza dos números (como se as questões sociais e humanas não fossem, também, complexas e duras), o que é um equívoco considerando que estamos caminhando em direção a um mundo mais “datafocado”, permeado por dados quantitativos, e não o contrário. Ao ser desincentivado, esse estudante perde a oportunidade de desenvolver novas formas de refletir a respeito do mundo. Como defende Regis (2006), precisamos pensar para além dos limites das ciências porque o mundo não respeita as fronteiras epistemológicas, artificiais, que a academia criou.

As pesquisas dedicadas aos fenômenos ligados à internet e à comunicação digital nos parecem, em especial, cada vez mais demandantes desses conhecimentos transdisciplinares, justamente porque deixam ainda mais claros os limites dos métodos “puros”.

Esta não é, no entanto, uma defesa dos métodos quantitativos como sendo superiores aos qualitativos, uma discussão que nos parece ultrapassada, mas uma contatação de que é necessário, para darmos conta de explicar os fenômenos contemporâneos, abraçar caminhos e possibilidades metodológicas que são oferecidas pelos mais diversos campos do conhecimento científico.

Escolher o financiamento coletivo como objeto de estudo não só nos permitiu, mas exigiu que nos embrenhássemos em conceitos e ferramentas que não são típicos da

comunicação, o que resultou em uma contribuição que não se limita à comunicação. As conclusões de nossas análises podem auxiliar em uma melhor compreensão de um fenômeno econômico recente, além de ter o potencial de servir como inspiração para a formulação de políticas públicas de fomento à inovação, de desenvolvimento regional e de novas articulações comunitárias.

O que podemos compreender a respeito dos comportamentos dos apoiadores é que eles são atraídos de forma diferente de acordo com a categoria dos projetos. Há uma comunidade mais densa de interessados em quadrinhos e em produção audiovisual no Catarse do que em projetos pessoais e jornalismo, o que vai demandar dos realizadores dessas categorias novas estratégias de engajamento para aumentar sua atratividade.

Independentemente das categorias, este estudo mostra, ainda, que realizadores que planejam melhor suas estratégias de comunicação atraem mais realizadores. Mais atualizações, a utilização de outras plataformas de comunicação e uma descrição não muito longa, nem muito curta, são importantes. Nos parece que há espaço, neste campo, para se desenvolver estudos que aprofundem a compreensão a respeito de características dos textos das recompensas e de elementos específicos dos textos das descrições, o que demandaria recursos como reconhecimento de caracteres técnicas de análise automatizada de texto. Além disso, há espaço para estudos mais qualitativos a respeito da percepção dos apoiadores a determinados tipos de descrição. Considerando a natureza comunitária do *crowdfunding*, acreditamos que as descrições mais emocionais, com temas ligados a questões sociais, ambientais ou dramas pessoais tendam a mobilizar mais os apoiadores.

As características dos projetos também têm um papel relevante na determinação dos apoios a um projeto. Aqueles que oferecem mais recompensas, têm a duração recomendada pela plataforma e escolhem a categoria Flex, atraem mais apoiadores. Neste ponto acreditamos também que há espaço para novas frentes de pesquisa porque nos parece contraintuitivo que um projeto Tudo-ou-nada atraia menos apoios.

As características dos realizadores são também relevantes. Ter uma rede pessoal de contatos e se engajar com outros projetos na plataforma são importantes para se atrair apoios. Já a experiência na realização de projetos anteriores diminui o número de apoiadores, o que nos parece ser sinal de uma espécie de otimização do número de apoiadores, ou seja, os projetos podem precisar de menos apoiadores para ser bem-sucedidos quanto mais experientes forem os realizadores.

Por fim, ainda que a internet tenha uma promessa de desterritorialização e liberdade na circulação das informações, das ideias e, no caso do financiamento coletivo, das doações, a

geografia é um fator relevante na atratividade dos projetos. Apoiadores tendem a doar mais em locais com mais densidade populacional, o que é um indicativo de que tendem a se relacionar mais com projetos realizados a pouca distância, e naqueles com mais concentração de renda. A esse respeito, acreditamos que há espaço para estudos que se aprofundem em como as diferenças culturais, sociais e econômicas dos estados brasileiros afetam o comportamento dos apoiadores.

## REFERÊNCIAS

ACEVEDO, D. D. *Beyond the Algorithm: Qualitative Insights for Gig Work Regulation*. Cambridge: Cambridge University Press, 2020.

AGRAWAL, A.; CATALINI, C.; GOLDFARB, A. *Some simple economics of crowdfunding*. Cambridge: Cambridge University Press, 2013.

AGRAWAL, A. K.; CATALINI, C.; GOLDFARB, A. *The Geography of Crowdfunding*. Cambridge: Cambridge University Press, 2011. (Working Paper Series, 16820). Disponível em: <http://www.nber.org/papers/w16820>.

AGRESTI, A.; FRANKLIN, C.; KLINGENBERG, B. *Statistics the Art and Science of Learning from data*. Florida: Pearson Education Limited, 2018.

ALIM, S. An initial exploration of ethical research practices regarding automated data extraction from online social media user profiles. *First Monday*, 2014.

ANDERSON, C. *A cauda longa: do mercado de massa para o mercado de nicho*. Rio de Janeiro: Elsevier, 2006.

ANDERSON, C. *Makers: The new industrial revolution* New York. NY: *Crown Business*, 2012.

BARABÁSI, A.-L.; FRANGOS, J. *Linked: the new science of networks science of networks*. New York: Basic Books [E-book], 2002.

BARRETO, A. S. *Modelos De Regressao: teoria e aplicações com o programa estatístico R*. Brasília: Ed. do Autor, 2011. 176 p. ISBN: 978-85-911728-0-1.

BEST, J. et al. *Crowdfunding's potential for the developing world*. World, 2013.

BRASIL. *Lei no 10.406, de 10 de janeiro de 2002*. 2002. Institui o Código Civil. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2002/L10406.htm](http://www.planalto.gov.br/ccivil_03/leis/2002/L10406.htm). Acesso em: 5 de dezembro de 2017.

BROOKS, M. E. et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, v. 9, n. 2, p. 378–400, 2017. Disponível em: (<https://journal.r-project.org/archive/2017/RJ-2017-066/index.html>).

BRÜNTJE, D.; GAJDA, O. *Crowdfunding in Europe: State of the Art in Theory and Practice. State of the Art in Theorey and Practice*, Springer International Publishing, 2016.

CADY, F. *The data science handbook*. New Jersey: John Wiley & Sons, 2017.

CALLON, M. Entrevista com michel callon: dos estudos de laboratório aos estudos de coletivos heterogêneos, passando pelos gerenciamentos econômicos. *Sociologias*, n. 19, p. 302–321, 2008.

CAMERON, A. C.; TRIVEDI, P. K. *Regression analysis of count data*. Cambridge: Cambridge university press, 2013. v. 53.

CARR, S. Collective action and the financing of innovation: Evidence from crowdfunding. *Darden Business School Working Paper*, 2013.

CASTELLS, M. *A Galáxia Internet: reflexões sobre a Internet, negócios e a sociedade*. Rio de Janeiro: Zahar, 2003.

CASTELLS, M. *The network society: a cross-cultural perspective*. Cheltenham; Northampton: Edward Elgar, 2004.

CASTELLS, M. *Redes de indignação e esperança: movimentos sociais na era da internet*. Rio de Janeiro: Zahar, 2013.

CASTELLS, M. *Ruptura: a crise da democracia liberal [e-book]*. Rio de Janeiro: Jorge Zahar Editor, 2018.

CATARSE: Projeto cafeteira aram. 2017. Disponível em: <https://www.catarse.me/cafeteiraaram>. Acesso em: 7 de agosto de 2017.

CATARSE: Projeto quem mexeu na minha merenda? 2017. Disponível em: <https://www.catarse.me/quemmexeunaminhamerenda> . Acesso em: 5 de dezembro de 2017.

CATARSE: Sala de imprensa. 2017. Disponível em: [https://www.catarse.me/pt/press?ref=ctrse\\_footer](https://www.catarse.me/pt/press?ref=ctrse_footer). Acesso em: 7 de janeiro de 2017.

CATARSE: Projeto delico. a frigideira de ferro repensada. 2017. Disponível em: [https://www.catarse.me/frigideira\\_delico](https://www.catarse.me/frigideira_delico). Acesso em: 5 dez. 2017.

CATARSE: Catarse assinaturas. 2017. Disponível em: (<https://crowdfunding.catarse.me/assinaturas>). Acesso em: 5 de dezembro de 2017.

CATARSE: Apoiadores. 2017. Disponível em: [https://suporte.catarse.me/hc/pt-br/articles/202031703-O-que-fazer-quando-o-projeto-que-eu-apoiei- %C3%A9-financiado-mas-n %C3 %A3o- %C3 %A9-realizado-de-acordo-com-a-proposta-](https://suporte.catarse.me/hc/pt-br/articles/202031703-O-que-fazer-quando-o-projeto-que-eu-apoiei-%C3%A9-financiado-mas-n%C3%A3o-%C3%A9-realizado-de-acordo-com-a-proposta-) . Acesso em: 5 de dezembro de 2017.

CATARSE: Diretrizes para criação de projetos. 2018. Disponível em: <https://suporte.catarse.me/hc/pt-br/articles/202387638-Diretrizes-para-criacao-de-projetos> . Acesso em: 24 jul. 2018.

CATARSE: 3 dicas para definir o prazo de sua campanha de financiamento coletivo. 2018. Disponível em: (<http://fazum.catarse.me/prazo>). Acesso em: 25 jul. 2018.

CATARSE: Projeto online. 2018. Disponível em: <https://suporte.catarse.me/hc/pt-br/articles/115002214463-projeto-ONLINE>. Acesso em: 25 jul. 2019.

CATARSE: Termos de serviço. 2018. Disponível em: <https://crowdfunding.catarse.me/legal/termos-de-uso>. Acesso em: 5 dez. 2017.

CATARSE: Retrospectiva 2019. 2019. Disponível em: <https://ano.catarse.me/2019>. Acesso em: 2 de agosto de 2020.

CHAN, C. Hong et al. *Rio: a Swiss-army knife for data file I/O*. [S.l.], 2018. R package version 0.5.16.

CHEN, H.; CHIANG, R. H.; STOREY, V. C. Business intelligence and analytics: From big data to big impact. *MIS quarterly*, JSTOR, p. 1165–1188, 2012.

CHEN, X.-P.; YAO, X.; KOTHA, S. Entrepreneur passion and preparedness in business plan presentations: a persuasion analysis of venture capitalists' funding decisions. *Academy of Management journal*, Academy of Management Briarcliff Manor, NY, v. 52, n. 1, p. 199–214, 2009.

CIOFFI-REVILLA, C. Introduction to computational social science. *London and Heidelberg: Springer*, Springer, 2014.

CNS. *Resolução nº 510. 2016*. Disponível em: <https://conselho.saude.gov.br/resolucoes/2016/Reso510.pdf>. Acesso em: 30 de julho de 2020.

COCATE, F. M.; JÚNIOR, C. P. Crowdfunding: estudo sobre o fenômeno virtual. *LÍBERO*, n. 29, p. 135–144, 2016.

CONVERSATION: Facebook's data lockdown is a disaster for academic researchers. 2015. Disponível em: <https://theconversation.com/facebooks-data-lockdown-is-a-disaster-for-academic-researchers-94533> . Acesso em: 30 de julho de 2020.

COSTA, R. B. Sobre o papel da confiança e das tecnologias digitais de comunicação nas experiências de economia colaborativa. *UFRJ. Anais dos Seminários dos Alunos do PPGAS-Museu Nacional-UFRJ*, v. 7, 2015.

COSTA, R. B. *Economia da Confiança: sobre a construção dos vínculos sociais no Airbnb e no Couchsurfing*. Tese (Doutorado) - Universidade do Estado do Rio de Janeiro, 2017.

CRESWELL, J. W. *A concise introduction to mixed methods research*. University of Nebraska - Lincoln: SAGE publications, 2014.

CVM: Instrução cvm 588. 2017. Disponível em: <http://www.cvm.gov.br/legislacao/instrucoes/inst588.html>. Acesso em: 5 de dezembro de 2017.

CÂMARA DOS DEPUTADOS: Novidade nas eleições deste ano, financiamento coletivo já está valendo. 2018. Disponível em: <http://www2.camara.leg.br/camaranoticias/noticias/POLITICA/560530-NOVIDADE-NAS-ELEICOES-DESTE-ANO,-FINANCIAMENTO-COLETIVO-JA-ESTA-VALENDO.html> . Acesso em: 20 jul. 2018.

DRESNER, S. *Crowdfunding: a guide to raising capital on the internet*. New Jersey: John Wiley & Sons, 2014.

DUARTE, J.; BARROS, A. Métodos e técnicas de pesquisa em comunicação. Atlas, São Paulo, 2005.

EDMONDSON, M. *googleLanguageR: Call Google's 'Natural Language' API, 'Cloud Translation' API, 'Cloud Speech' API and 'Cloud Text-to-Speech' API*. [S.l.], 2020. R package version 0.3.0. Disponível em: <https://CRAN.R-project.org/package=googleLanguageR>.

EPSTEIN, I. Ciência, poder e comunicação. In: EPSTEIN, I. *Métodos e técnicas de pesquisa em comunicação*. São Paulo: Atlas, 2005. p. 15–30.

FACEBOOK. *Facebook terms of service*. 2019. Disponível em: <https://en-gb.facebook.com/legal/terms>. Acesso em: 2 de julho de 2019.

FARIETA, A. *Aproximaciones éticas al problema del free rider: consecuencialismo, deontología y ética de la virtud*. 2015.

FELDMAN, R.; SANGER, J. *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge: Cambridge university press, 2007.

FELIPE, I. J. d. S. *Determinantes do sucesso de campanhas de equity e de reward crowdfunding*. Tese (Doutorado), 2017.

FIELD, A. P.; MILES, J.; FIELD, Z. *Discovering statistics using R*. London: Sage, 2012.

FIGARO, R. A triangulação metodológica em pesquisas sobre a comunicação no mundo do trabalho. *Fronteiras-estudos midiáticos*, v. 16, n. 2, p. 124–131, 2014.

FINCH, W. H.; BOLIN, J. E.; KELLEY, K. *Multilevel modeling using R*. New York: Crc Press, 2019.

FIRKE, S. *janitor: Simple Tools for Examining and Cleaning Dirty Data*. [S.l.], 2020. R package version 2.0.1. Disponível em: <https://CRAN.R-project.org/package=janitor>.

FISK, C. L. Foreword. In: ACEVEDO, D. D. (Ed.). *Beyond the Algorithm: Qualitative Insights for Gig Work Regulation*. Cambridge: Cambridge University Press, 2020. cap. 1, p. ix–xi.

FLICK, C. Informed consent and the facebook emotional manipulation study. *Research Ethics*, SAGE Publications Sage, UK: London, England, v. 12, n. 1, p. 14–28, 2016.

FOX, J.; HONG, J. Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software*, v. 32, n. 1, p. 1–24, 2009. Disponível em: <http://www.jstatsoft.org/v32/i01/>.

FRAGOSO, S.; RECUERO, R.; AMARAL, A. Métodos de pesquisa para internet. *Sulina*, Porto Alegre, v. 1, 2011.

GELMAN, A. Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, Wiley Online Library, Cambridge, v. 27, n. 15, p. 2865–2873, 2008.

GELMAN, A.; SU, Y.-S. *arm: Data Analysis Using Regression and Multilevel/Hierarchical Models*. [S.l.], 2018. R package version 1.10-1. Disponível em: <https://CRAN.R-project.org/package=arm>.

GIERCZAK, M. M. et al. Crowdfunding: Outlining the New Era of Fundraising. In: BRUNTJE, D.; GAJDA, O. (Ed.). *Crowdfunding in Europe: State of the Art in Theory and Practice*. Cham: Springer International Publishing, 2016, (FGF Studies in Small Business and Entrepreneurship). p. 7–23.

GIZMODO: The 9 most disgraceful crowdfunding failures of 2015. 2015. Disponível em: <https://gizmodo.com/the-9-most-disgraceful-crowdfunding-failures-of-2015-1747957776>. Acesso em: 5 de dezembro de 2017.

GLOBO: Maracanã é construído em tempo recorde para sediar a copa do mundo de 1950. 2013. Disponível em: <http://acervo.oglobo.globo.com/rio-de-historias/maracana-construido-em-tempo-recorde-para-sediar-copa-do-mundo-de-1950-8906021> . Acesso em: 5 de dezembro de 2017.

GOOGLE: Análise de sentimentos. 2020. Disponível em: <https://cloud.google.com/natural-language/docs/basics#interpreting-sentiment-analysis-values> . Acesso em: 15 de março de 2020.

HARTIG, F. *DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. [S.l.], 2020. R package version 0.2.7. Disponível em: <https://CRAN.R-project.org/package=DHARMA>.

HELBING, D. *Quantitative sociodynamics: stochastic methods and models of social interaction processes*. Berlim: Springer Science & Business Media, 2010.

HELMOND, A. The platformization of the web: Making web data platform ready. *Social Media+ Society*, Sage Publications, UK: London, England, v. 1, n. 2, 2015.

BGE: Sistema de contas regionais. 2018. Disponível em: <https://www.ibge.gov.br/estatisticas-novoportal/economicas/contas-nacionais/9054-contas-regionais-do-brasil.html?=&t=o-que-e>). Acesso em: 25 de julho de 2018.

IBGE, D. d. P. *Coordenação de População e Indicadores Sociais, Estimativas da população residente com data de referência 1 de julho de 2017*. 2018.

INVESTOPIA: Definição de security. 2018. Disponível em: <https://www.investopedia.com/terms/s/security.asp>. Acesso em: 30 de julho de 2018.

JOHNSON, F.; GUPTA, S. K. Web content mining techniques: a survey. *International Journal of Computer Applications*, Foundation of Computer Science, v. 47, n. 11, 2012.

JOLANI, S. et al. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using mice. *Statistics in medicine*, Wiley Online Library, v. 34, n. 11, p. 1841–1863, 2015.

KICKSTARTER. *Kickstarter basics*. 2017. Disponível em: <https://www.kickstarter.com/help/faq/kickstarter%20basics>. Acesso em: 5 dez. 2017.

KICKSTARTER. *Prohibited itens*. 2018. Disponível em: (<https://www.kickstarter.com/rules/prohibited?ref=handbook-reward>). Acesso em: 25 de julho de 2018.

KICKSTARTER. *Prototypes renderings*. 2018. Disponível em: <https://www.kickstarter.com/rules/prototypes?ref=rules>. Acesso em: 25 de julho de 2018.

KLEIBER, C.; ZEILEIS, A. *Applied Econometrics with R*. New York: Springer-Verlag, 2008. Disponível em: <https://CRAN.R-project.org/package=AER>.

KLEIBER, C.; ZEILEIS, A. Visualizing count data regressions using rootograms. *The American Statistician*, v. 70, n. 3, p. 296–303, 2016.

- KRAMER, A. D.; GUILLORY, J. E.; HANCOCK, J. T. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 111, n. 24, p. 8788–8790, 2014.
- KUPPUSWAMY, V.; BAYUS, B. L. *Crowdfunding creative ideas: The dynamics of project backers in Kickstarter*. 2015.
- LANDSTRÖM, H. *Handbook of research on venture capital*. Massachusetts: Edward Elgar Publishing, 2007.
- LATOUR, B. Where are the missing masses? the sociology of mundane artifacts. *Technology and Society: Building our Sociotechnical Future*, MIT Press, 2008.
- LEVITSKY, S.; ZIBLATT, D. *Como as democracias morrem*. Rio de Janeiro: Jorge Zahar Editor, 2018.
- LÉVY, P. *Cibercultura*. Editora 34, São Paulo, 1999.
- LONGHI, C.; PETITI, M. *JOTA: Riscos envolvendo a prática de scraping à luz do ordenamento jurídico nacional*. 2018. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/riscos-envolvendo-a-pratica-de-scraping-a-luz-do-ordenamento-juridico-nacional-08122018> . Acesso em: 30 de julho de 2020.
- LOPES, M. I. V. de. Pesquisa de comunicação: questões epistemológicas, teóricas e metodológicas. *Intercom-Revista Brasileira de Ciências da Comunicação*, v. 27, n. 1, 2004.
- LU·DECKE, D. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, v. 3, n. 26, p. 772, 2018.
- LU·DECKE, D. *sjPlot: Data Visualization for Statistics in Social Science*. [S.l.], 2020. R package version 2.8.2. Disponível em: <https://CRAN.R-project.org/package=sjPlot>.
- MACMILLAN, I. C.; SIEGEL, R.; NARASIMHA, P. S. Criteria used by venture capitalists to evaluate new venture proposals. *Journal of Business venturing*, Elsevier, v. 1, n. 1, p. 119–128, 1985.
- MARTINS, N. P. M.; MARTINS, Y. M. *A influência da comunicação digital para o financiamento de projetos em crowdfunding: um estudo sobre projetos bem-sucedidos da plataforma catarse*. 2016.
- MOLLICK, E. The dynamics of crowdfunding: An exploratory study. *Journal of business venturing*, Elsevier, v. 29, n. 1, p. 1–16, 2014.
- MOLLICK, E. R. Delivery rates on kickstarter. *Available at SSRN 2699251*, 2015.
- MUNZERT, S. et al. *Automated data collection with R: A practical guide to web scraping and text mining*. United Kingdom: John Wiley & Sons, 2014.
- NEYLAND, D. *The everyday life of an algorithm*. London: Springer Nature, 2019.
- NOBLE, S. U. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press, 2018.

- OBAR, J. A.; OELDORF-HIRSCH, A. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, Taylor & Francis, v. 23, n. 1, p. 128–147, 2020.
- OOMS, J. The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv:1403.2805 [stat.CO]*, 2014. Disponível em: <https://arxiv.org/abs/1403.2805>.
- PILNY, A.; POOLE, M. S. *Group processes: Data-driven computational approaches*. Lexington: Springer, 2017.
- POPULACIONAL, C. Censo populacional 2010. *Instituto Brasileiro de Geografia e Estatística (IBGE)(29 de novembro de 2010)*, v. 11, n. 12, p. 2010, 2010.
- PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. *Big data*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 1, n. 1, p. 51–59, 2013.
- REGIS, F. Comunicação, sistemas complexos e transdisciplinaridade: um comunicar intercientífico. *Revista Contracampo*, n. 15, p. 151–164, 2006.
- SCHUTT, R.; O'NEIL, C. *Doing data science: straight talk from the frontline*. California: O'Reilly, 2014.
- SHIRKY, C. *Here comes everybody: the power of organizing without organizations*. New York, USA: Penguin Books, 2008.
- SHIRKY, C. *Cognitive surplus: creativity and generosity in a connected age* [E-book]. New York, USA: Penguin Books, 2010.
- SIMONDON, G. *El modo de existencia de los objetos técnicos*. Buenos Aires, Argentina: Prometeo Libros Editorial, 2007.
- SODRÉ, M. *A ciência do comum: notas para o método comunicacional*. Petrópolis, RJ: Editora Vozes, 2015.
- SSRC: Social media and democracy research grants. 2019. Disponível em: <https://www.ssrc.org/fellowships/view/social-media-and-democracy-research-grants/grantees/> . Acesso em: 30 de julho de 2020.
- STOCKEMER, D. *Quantitative methods for the social sciences: A Practical Introduction with examples in SPSS and Stata*. Ottawa, Canada: Springer, 2018.
- TAPSCOTT, D.; WILLIAMS, A. D. *Wikinomics: como a colaboração em massa pode mudar o seu negócio*. Rio de Janeiro: Nova Fronteira, 2007.
- TSE. *Financiamento coletivo*. 2018. Disponível em: <http://www.tse.jus.br/eleicoes/eleicoes-2018/prestacao-de-contas-1/financiamento-coletivo>. Acesso em: 25 jul. 2018.
- TWITTER. *Twitter api reference*. 2020. Disponível em: <https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets> . Acesso em: 4 ago. 2020.

TWITTER. *Twitter termos de serviços*. 2020. Disponível em: (<https://twitter.com/pt/tos>). Acesso em: 30 de julho de 2020.

VAN BUUREN, S.; GROOTHUIS-OUDSHOORN, K. mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, v. 45, n. 3, p. 1–67, 2011. Disponível em: <https://www.jstatsoft.org/v45/i03/>.

VENABLES, W. N.; RIPLEY, B. D. *Modern Applied Statistics with S*. Fourth. New York: Springer, 2002. ISBN 0-387-95457-0. Disponível em: <http://www.stats.ox.ac.uk/pub/MASS4>.

WICKHAM, H. *rvest: Easily Harvest (Scrape) Web Pages*. [S.l.], 2019. R package version 0.3.5. Disponível em: <https://CRAN.R-project.org/package=rvest>.

WICKHAM, H. et al. Welcome to the tidyverse. *Journal of Open Source Software*, v. 4, n. 43, p. 1686, 2019.

WICKHAM, H.; GROLEMUND, G. *R for data science: import, tidy, transform, visualize, and model data*. California: O'Reilly Media, 2016.

WICKHAM, H.; MILLER, E. *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. [S.l.], 2019. R package version 2.2.0. Disponível em: (<https://CRAN.R-project.org/package=haven>).

XU, A. et al. Show me the money! an analysis of project updates during crowdfunding campaigns. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2014. (CHI '14), p. 591–600. ISBN 9781450324731. Disponível em: (<https://doi.org/10.1145/2556288.2557045>).

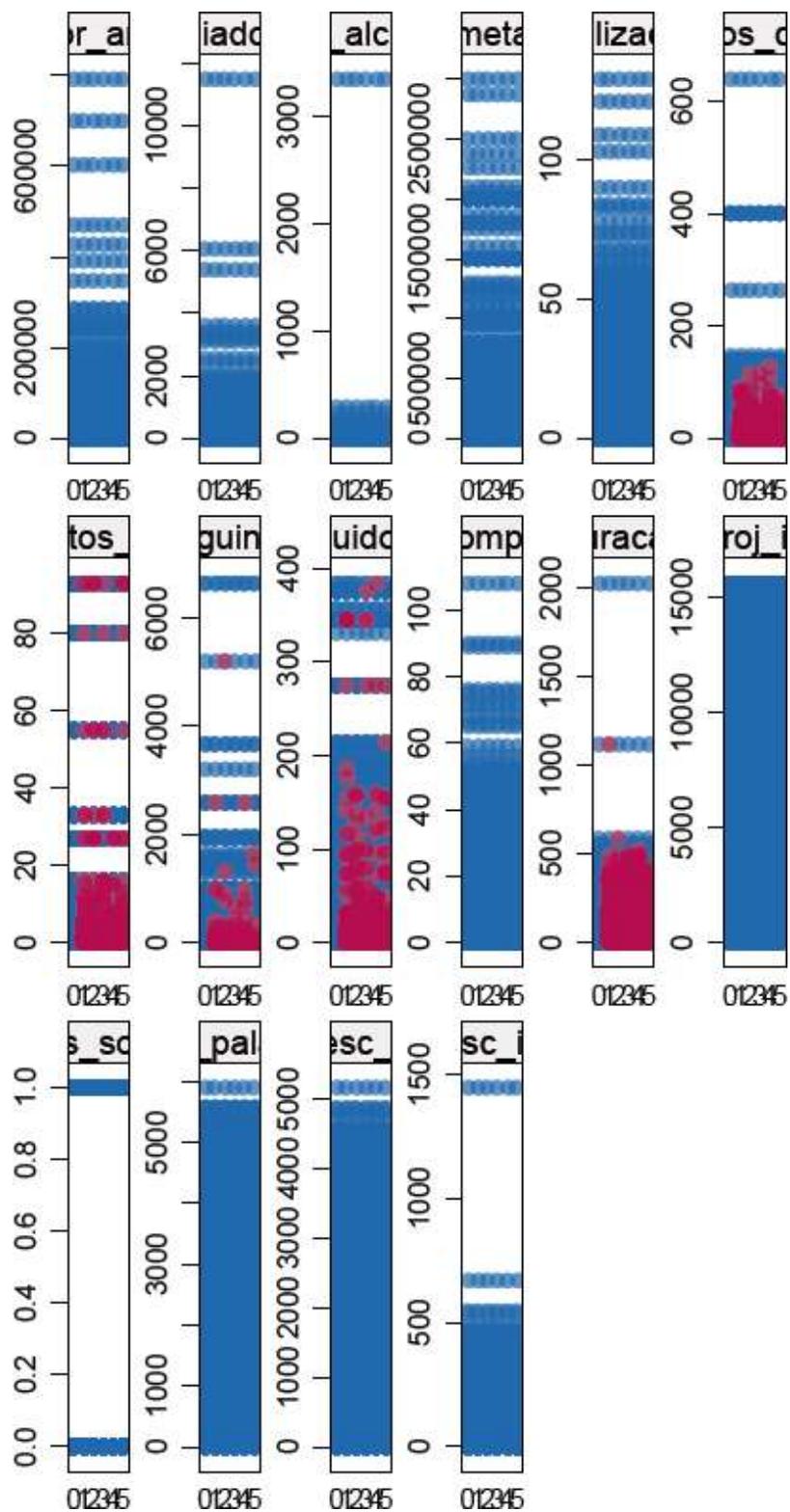
ZEILEIS, A.; KLEIBER, C.; JACKMAN, S. Regression models for count data in R. *Journal of Statistical Software*, v. 27, n. 8, 2008. Disponível em: <http://www.jstatsoft.org/v27/i08/>.

## GLOSSÁRIO

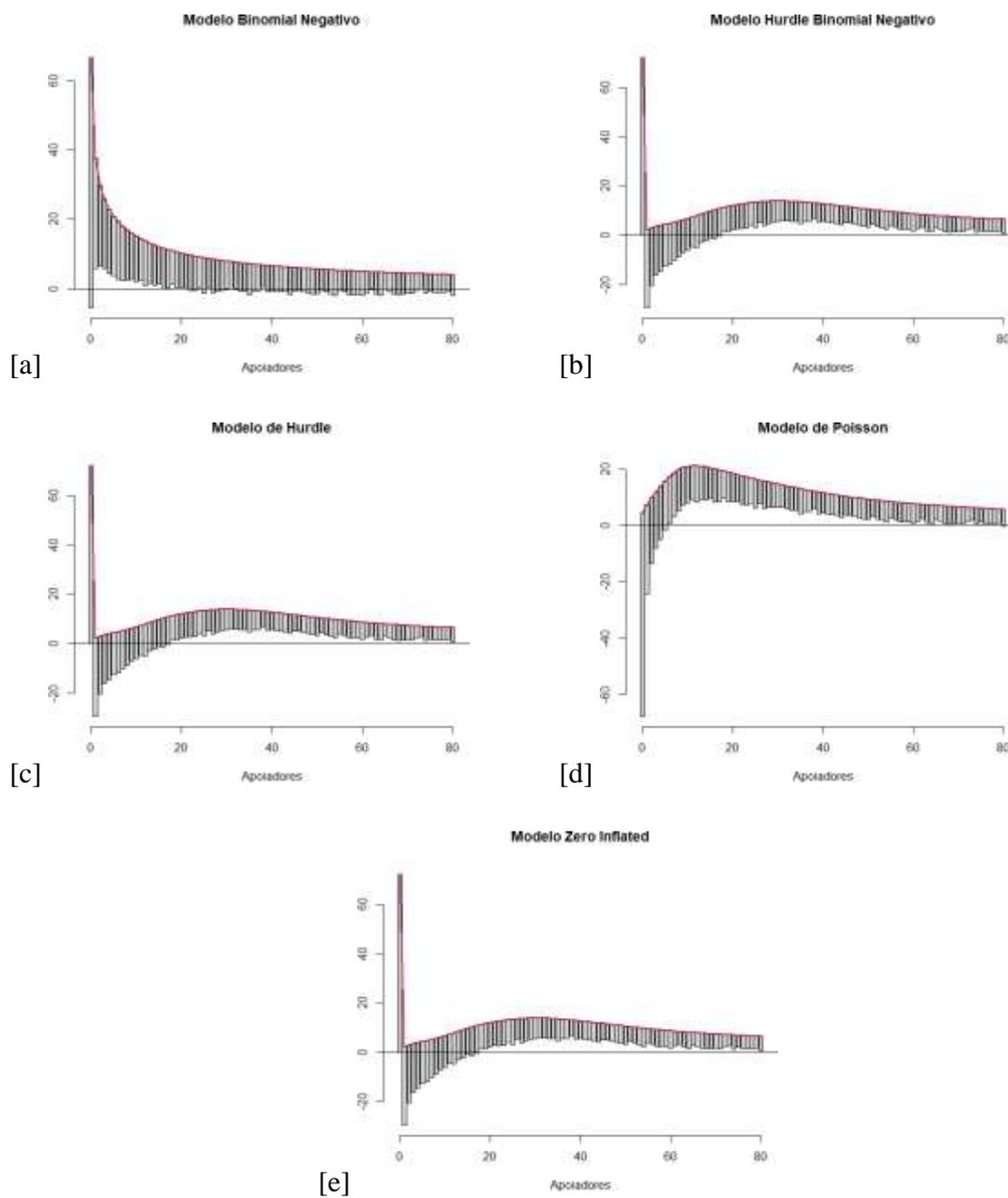
Accountability	Na economia, refere-se à necessidade de se manter registros confiáveis das operações financeiras que serão utilizadas, posteriormente, em um processo de prestação de contas, nos quais os gestores são responsabilizados por suas decisões.
Algoritmo	Sequência de passos ou regras a ser seguida em operações com o objetivo de resolver problemas. Na computação, refere-se às etapas que serão seguidas por um computador para resolver uma tarefa, seja recomendar um livro, um caminho a ser seguido, ou conhecidos em um site de redes sociais.
API	Application Programming Interface (API) é uma denominação utilizada para uma gama ampla de aplicações. No caso deste estudo, estamos usando APIs de serviços web, ou seja, a interface através da qual um aplicativo ou um serviço se relaciona com um servidor web, o que facilita a extração dos dados. Resumidamente, é uma interface de troca de informações de forma estruturada.
Cluster	A tradução direta é grupo, ou agrupamento. Ao ser utilizada no contexto dos estudos das redes, refere-se a uma concentração de conexões
Due-diligence	Refere-se à investigação necessária para que se reduzam os riscos de uma transação financeira, geralmente compra ou venda de ativos. Considera-se que essa é uma reponsabilidade dos gestores, sobretudo em fusões e aquisições de empresas. Essa investigação envolve aspectos legais, financeiros, trabalhistas ou mesmo estratégicos dos ativos em questão.
Equity	No jargão do mercado financeiro, equity significa patrimônio líquido, ou o quanto um determinado investidor possui em ativos. Um financiamento coletivo que tem como recompensa equity significa que a recompensa que o investidor receberá em troca dos recursos aportados é uma parte da empresa, que se torna patrimônio do investidor.
GitHub	Principal plataforma de colaboração de programadores do mundo. Utilizada para o compartilhamento de códigos, trocas de experiências em programação, sugestões de melhorias, etc. Foi adquirida pela Microsoft no início de junho de 2018.

Machine Learning	Conjunto de técnicas e ferramentas que busca identificar padrões e extrair conclusões de forma automatizada a partir de grandes bases de dados.
Marketplaces	Plataformas de comércio eletrônico, como o eBay, Mercado-Livre, Amazon e Americanas.com, nas quais vendedores individuais podem oferecer seus produtos pagando comissão à plataforma.
Securities	São valores mobiliários, como títulos de dívida e de propriedade, que podem, em alguns casos, ser negociadas em mercados financeiros. Representam a propriedade em uma corporação negociada publicamente, o crédito com um governo ou uma empresa ou os direitos de propriedade representadas por opções de compra (INVESTOPIA, 2018).
R Studio	É um ambiente de desenvolvimento para o R. Em sua forma original, o R é uma linguagem de programação que roda linha a linha de código. Ambientes como o R Studio facilitam o trabalho do programador ao permiti-lo reeditar e salvar os códigos, salvar o histórico das modificar, plotar os gráficos diretamente dentro da plataforma, etc.
Text mining	Para (??), text mining pode ser definido de maneira geral como um processo através do qual um usuário interage com uma coleção de documentos utilizando-se de uma série de ferramentas analíticas para identificar, extrair e explorar padrões que julga relevantes.
Web scraping	Também chamado de raspagem de dados, refere-se à ação de extração automatizada de informações de uma página <i>web</i> .

## APÊNDICE A – Diagnósticos da Imputação múltipla



## APÊNDICE B – Diagnósticos dos Modelos de Contagem



## APÊNDICE C – Sumário estatístico

Tabela 15 - Sumário Estatístico.

<b>Número de observações = 15.595</b>	
<b>apoiadores</b>	<i>número de apoiadores</i>
média (sd <sup>1</sup> )	48.65 ± 172.18
mediana (Q1, Q3 <sup>2</sup> )	5.00 (0.00, 42.00)
min	0
max	11471
<b>n recompensas</b>	<i>número de recompensas oferecidas</i>
média (sd)	5.65 ± 5.91
mediana (Q1, Q3)	5.00 (1.00, 8.00)
min	0
max	108
<b>meta</b>	<i>valor a ser alcançado pelo projeto</i>
média (sd)	R\$ 30.414,66 ± R\$ 107.312,32
mediana (Q1, Q3)	R\$ 10.000,00 (R\$ 4.500,00, R\$ 25.000,00)
min	R\$ 9,00
max	R\$ 2.872.500,00
<b>atualizacoes</b>	<i>número de atualizações enviadas aos apoiadores</i>
média (sd)	2.10 ± 5.77
mediana (Q1, Q3)	0.00 (0.00, 1.00)
min	0
max	129
<b>duracao</b>	<i>número de dias em que o projeto recebeu contribuições</i>
média (sd)	65.37 ± 66.95
mediana (Q1, Q3)	60.00 (35.00, 60.00)
min	1
max	2.029
<b>modalidade</b>	<i>modalidade na qual o projeto se inscreve</i>
Flexível (n, N% <sup>3</sup> )	7,629 (48.92%)
Tudo-ou-Nada (n, N% )	7,966 (51.08%)
<b>Redes_sociais</b>	<i>se o projeto indica perfis em SRS em sua descrição</i>
Sim (1) (n, N% )	10,800 (69.25%)
Não (0) (n, N% )	4,795 (30.75%)
<b>Desc_palavras</b>	<i>tamanho da descrição, em número de palavras</i>
média (sd)	496.48 ± 468.70
mediana (Q1, Q3)	373.00 (188.00, 640.00)
min	0
max	3.369
<b>magnitude</b>	<i>intensidade do texto da descrição</i>
média (sd)	15.33 ± 15.35
mediana (Q1, Q3)	11.30 (5.30, 19.90)
min	0
max	186.3

Tabela 15 - Sumário Estatístico.

<b>Número de observações = 15.595</b>	
<b>score</b>	<i>sentimento do texto da descrição de -1 a +1</i>
média (sd)	0.14 ± 0.14
mediana (Q1, Q3)	0.10 (0.10, 0.20)
min	-0.9
max	0.9
<b>seguidores</b>	<i>número de seguidores do realizador</i>
média (sd)	4.97 ± 23.67
mediana (Q1, Q3)	0 (0.00, 2.00)
min	0
max	384
<b>seguindo</b>	<i>número de usuários seguidos pelo realizador</i>
média (sd)	9.29 ± 130.82
mediana (Q1, Q3)	0 (0.00, 0.00)
min	0
max	6642
<b>projetos_contrib</b>	<i>projetos com os quais o realizador contribuiu</i>
média (sd)	3.68 ± 12.90
mediana (Q1, Q3)	0 (0.00, 3.00)
min	0
max	640
<b>projetos_publici</b>	<i>projetos publicados pelo realizador</i>
média (sd)	2.76 ± 10.01
mediana (Q1, Q3)	1 (1.00, 1.00)
min	0
max	93
<b>gini_cidade<sup>4</sup></b>	<i>coeficiente de Gini das cidades de origem</i>
média (sd)	0.59 ± 0.06
mediana (Q1, Q3)	0.62 (0.55, 0.65)
min	0.2947
max	0.8082
<b>densidade<sup>5</sup></b>	<i>densidade populacional das cidades de origem</i>
média (sd)	4,225.51 ± 3,294.63
mediana (Q1, Q3)	4,406.96 (522.95, 8,005.25)
min	0.1669463
max	14006.77

Fonte: elaboração do autor.

1) *sd* refere-se ao desvio-padrão.

2) *Q1* e *Q3* são os percentis 25 e 75, respectivamente.

3) *n*, *N%* referem-se ao número de observações e ao percentual do total de observações

4) Fonte: Populacional (2010)

5) Fonte: IBGE (2018b)

## APÊNDICE D – Matrizes de correlação

Tabela 16 - Matriz de correlações - Coeficiente de Pearson

Variável	Apoiadores	Recompensas	Meta	Atualizações	Duração	SRS	Descrição	Seguidores	Seguindo	Contribuições	Publicações	Gini	Densidade	Magnitude
Apoiadores														
Recompensas	.27													
Meta	.03	-.04												
Atualizações	.39	.36	-.03											
Duração	-.05	-.07	.08	-.05										
SRS	.10	.23	-.04	.16	-.01									
Descrição	.22	.41	.02	.31	-.05	.21								
Seguidores	.27	.20	-.02	.41	-.03	.08	.16							
Seguindo	.03	.05	-.01	.08	.01	.03	.02	.27						
Contribuições	.13	.13	-.03	.21	-.04	.05	.07	.39	.20					
Publicações	-.01	-.07	-.04	-.00	.08	.08	-.04	.07	-.00	.03				
Gini	.08	.12	-.01	.06	-.03	.03	.08	.05	.01	.04	.07			
Densidade	.09	.12	-.01	.07	-.05	.03	.07	.05	.01	.06	.03	.68		
Magnitude	.23	.39	-.00	.32	-.04	.18	.84	.17	.03	.08	-.03	.07	.07	
Score	-.00	.03	-.01	.00	.01	.01	-.01	.00	.00	.02	-.01	.01	-.00	.04

Fonte: Elaboração do autor

Tabela 17 - Matriz de correlações - Coeficiente de Spearman

Variável	Apoiadores	Recompensas	Meta	Atualizações	Duração	SRS	Descrição	Seguidores	Seguindo	Contribuições	Publicações	Gini	Densidade	Magnitude
Apoiadores														
Recompensas	.63													
Meta	.06	.20												
Atualizações	.64	.47	.08											
Duração	-.09	.03	.23	-.04										
SRS	.26	.27	.07	.23	-.03									
Descrição	.47	.50	.20	.37	-.00	.25								
Seguidores	.45	.31	.01	.39	-.05	.13	.20							
Seguindo	.06	.07	.00	.13	.05	.05	.04	.32						
Contribuições	.42	.31	-.02	.38	-.04	.07	.20	.50	.23					
Publicações	.09	.04	-.10	.12	-.07	.10	.03	.29	.14	.20				
Gini	.18	.14	.05	.10	-.05	.02	.10	.10	-.01	.09	.04			
Densidade	.19	.16	.05	.12	-.06	.04	.10	.11	-.00	.11	.03	.78		
Magnitude	.47	.46	.14	.37	-.01	.21	.80	.21	.05	.21	.03	.10	.10	
Score	.04	.06	.02	.01	.02	.02	.01	.01	.01	.01	-.02	.01	.00	.10

Fonte: Elaboração do autor

## APÊNDICE E – Coeficientes dos modelos estimados

Tabela 18 - Modelos binomiais negativos para o número de apoiadores

<b>Variável Dependente: número de apoiadores</b>			
<b>Variáveis</b>	<b>Modelo 1</b>	<b>Modelo 2</b>	<b>Modelo 3</b>
Intercepto	4.2-*** (0.1)	4.2-*** (0.1)	4.2-*** (0.1)
Recompensas (l)	1.4*** (0.1)	1.4*** (0.1)	1.4*** (0.1)
Recompensas (ln)2	-0.6*** (0.1)	-0.6*** (0.1)	-0.6*** (0.1)
Meta (ln)	0.3 (0.3)	0.3 (0.3)	0.3 (0.3)
Atualizações (ln)	3.3*** (0.2)	3.3*** (0.2)	3.3*** (0.2)
Atualizações (ln)2	-1.7*** (0.1)	-1.7*** (0.1)	-1.7*** (0.1)
Duração (ln)	-0.8*** (0.1)	-0.8*** (0.1)	-0.7*** (0.1)
Duração (ln)2	0.2** (0.1)	0.2** (0.1)	0.2** (0.1)
Modalidade (tudo-ou-nada)	-0.2*** (0.0)	-0.2*** (0.0)	-0.2*** (0.0)
Redes Sociais (sim)	0.3*** 0.0	0.3*** 0.0	
Número SRS	0.0 (0.0)		0.1*** (0.0)
Número de palavras	0.1 (0.1)	0.1 (0.1)	0.1 (0.1)
Número de palavras2	-0.2*** (0.0)	-0.2*** (0.0)	-0.2*** (0.0)
Seguidores (ln)	4.1*** (0.2)	4.1*** (0.2)	4.1*** (0.2)
Seguidores (ln)2	-1.6*** (0.1)	-1.6*** (0.1)	-1.6*** (0.1)
Seguindo (ln)	-1.7*** (0.4)	-1.7*** (0.4)	-1.7*** (0.4)
Seguindo (ln)2	0.7*** (0.2)	0.7*** (0.2)	0.7*** (0.2)
Contribuições (ln)	-0.0 (0.1)	-0.0 (0.1)	-0.0 (0.1)
Contribuições (ln)2	-0.5*** (0.1)	-0.5*** (0.1)	-0.5*** (0.1)
Proj. Publicados (ln)	-0.7*** (0.1)	-0.7*** (0.1)	-0.8*** (0.1)
Gini cidades	0.1* (0.0)	0.1* (0.0)	0.1* (0.0)
Densidade	0.2*** (0.0)	0.2*** (0.0)	0.2*** (0.0)
Densidade2	-0.3*** (0.1)	-0.3*** (0.1)	-0.3*** (0.1)
<b>Controles</b>			
Magnitude (ln)	0.6*** (0.1)	0.6*** (0.1)	0.6*** (0.1)
Score	-0.0 (0.0)	-0.0 (0.0)	-0.0 (0.0)
Score2	-0.2*** (0.0)	-0.2*** (0.0)	-0.2*** (0.0)
Ano (2012)	0.2 (0.1)	0.2 (0.1)	0.2 (0.1)
Ano (2013)	-0.3** (0.1)	-0.3* (0.1)	-0.2' (0.1)
Ano (2014)	-0.3* (0.1)	-0.3* (0.1)	-0.2' (0.1)
Ano (2015)	-0.5*** (0.1)	-0.5*** (0.1)	-0.5*** (0.1)
Ano (2016)	-0.7*** (0.1)	-0.7*** (0.1)	-0.6*** (0.1)
Ano (2017)	-0.8*** (0.1)	-0.8*** (0.1)	-0.8*** (0.1)
Ano (2018)	-0.8*** (0.1)	-0.8*** (0.1)	-0.7*** (0.1)
***p < 0.001, **p < 0.01, *p < 0.05, 'p < 0.1			
Observações	15595	15595	15595
Marginal R2 / Conditional R2	0.639 / 0.714	0.639 / 0.714	0.637 / 0.713
AIC	112501.692	112499.866	112528.119
log-Likelihood	-56195.846	-56195.933	-56210.060

## APÊNDICE F – Coeficientes dos modelos estimados - efeitos aleatórios

Tabela 19 - Variância dos efeitos aleatórios - Modelo 1

	Groups	Name	Std.Dev.	Corr	V5	V6	V7	V8
X	categoria	(Intercept)	0.37886					
X.1		n_recompensas_ln	0.40763	0.097				
X.2		meta_ln	1.15071	0.252	-0.386			
X.3		atualizacoes_ln	0.78546	-0.523	-0.587	-0.382		
X.4		magnitude_ln	0.33488	-0.299	-0.214	-0.783	0.710	
X.5		seguindo_ln	0.88767	-0.086	-0.082	-0.084	0.058	0.163

Tabela 20 - Variância dos efeitos aleatórios - Modelo 2 (ATUALIZAR)

	Groups	Name	Std.Dev.	Corr	V5	V6	V7	V8
X	categoria	(Intercept)	0.37886					
X.1		n_recompensas_ln	0.40763	0.097				
X.2		meta_ln	1.15071	0.252	-0.386			
X.3		atualizacoes_ln	0.78546	-0.523	-0.587	-0.382		
X.4		magnitude_ln	0.33488	-0.299	-0.214	-0.783	0.710	
X.5		seguindo_ln	0.88767	-0.086	-0.082	-0.084	0.058	0.163

Tabela 21 - Variância dos efeitos aleatórios - Modelo 3 (ATUALIZAR)

	Groups	Name	Std.Dev.	Corr	V5	V6	V7	V8
X	categoria	(Intercept)	0.37886					
X.1		n_recompensas_ln	0.40763	0.097				
X.2		meta_ln	1.15071	0.252	-0.386			
X.3		atualizacoes_ln	0.78546	-0.523	-0.587	-0.382		
X.4		magnitude_ln	0.33488	-0.299	-0.214	-0.783	0.710	
X.5		seguindo_ln	0.88767	-0.086	-0.082	-0.084	0.058	0.163

## APÊNDICE G – Diagnósticos dos modelos estimados

Figura 29 - Normalidade dos resíduos

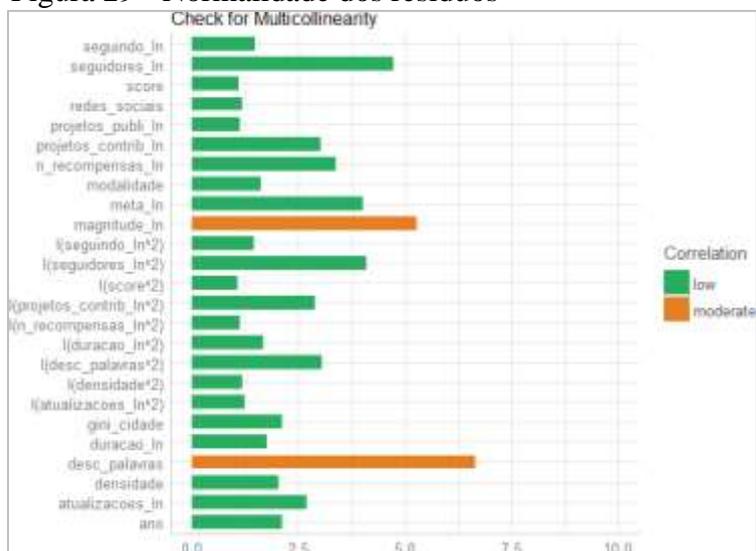


Figura 30 - QQ plot

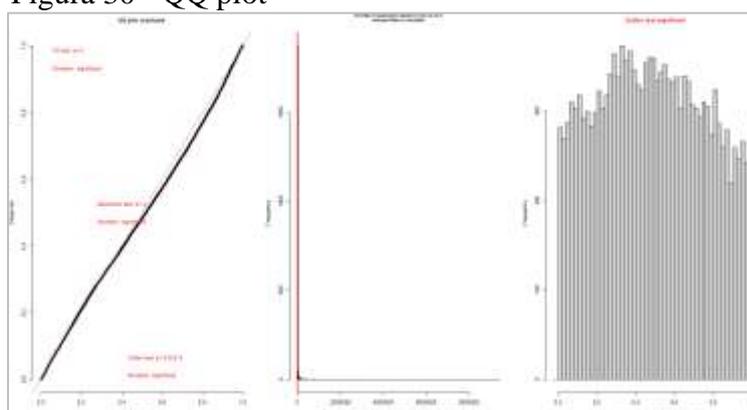


Figura 31 - Normalidade dos efeitos aleatórios

