

Capítulo

7

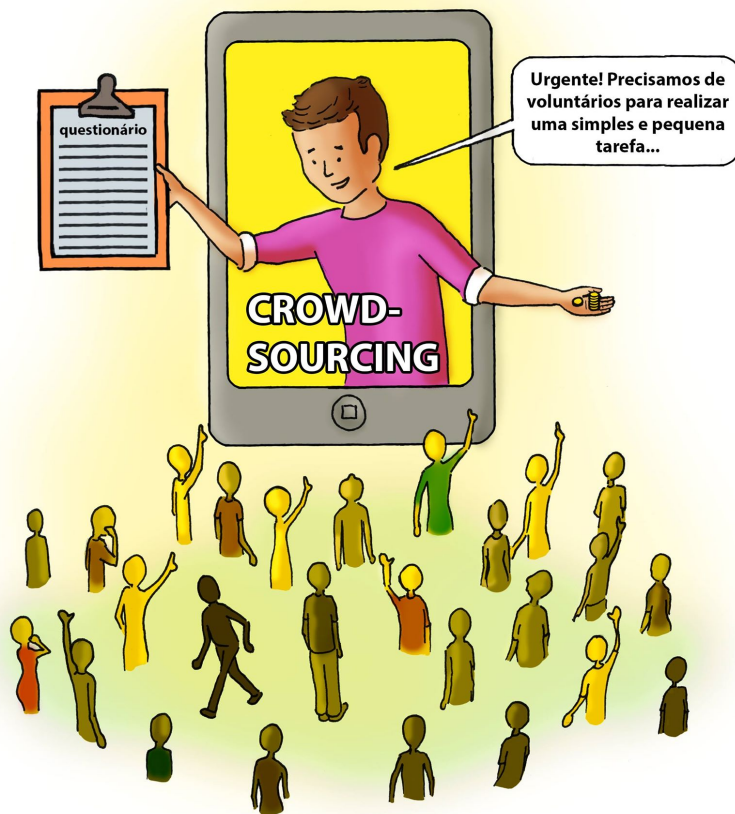
Coleta de dados através de *Crowdsourcing* em Pesquisas em Informática na Educação

Ujwal Gadiraju (L3S Research Center),
Bernardo Pereira Nunes (PUC-Rio e UNIRIO)

Objetivo do Capítulo

Este capítulo tem o objetivo de apresentar o conceito de *crowdsourcing* bem como uma reflexão sobre os fatores-chave a serem considerados em experimentos que exijam inteligência humana em larga escala. Ao final da leitura deste capítulo, você deve ser capaz de:

- Conhecer um novo paradigma de avaliações baseada em *crowdsourcing*.
- Conhecer um processo padrão para realizar tarefas *crowdsourcing*.
- Identificar oportunidades de executar tarefas baseadas no paradigma de *crowdsourcing*.
- Construir e modelar tarefas apropriadas para serem realizadas por trabalhadores *crowdsourcing*.
- Conhecer os prós e contras de tarefas realizadas através do *crowdsourcing*.



Em um lugar longínquo e isolado... havia um aluno de doutorado muito empenhado, que se dedicava a provar uma nova teoria que mudaria a área de Informática na Educação para sempre. A sua teoria era muito bem fundamentada e não tinha como estar errada, mas seu orientador insistia que ele deveria prová-la, verificar sua validade e aplicação na prática. Incomodado pelo comentário de seu orientador e certo de que sua teoria estava correta, o aluno saiu em busca de indivíduos que pudessem participar de sua pesquisa de apenas e, tão somente, 15 minutos. Quem poderia se recusar? Os que aceitavam participar não possuíam o perfil demográfico exigido. Os que recusavam diminuíam a chance do magnífico aluno de doutorado de obter o número mínimo de indivíduos para validar sua teoria. Sem saber o que fazer para convencer o seu orientador e provar que sua teoria estava correta, ele recorreu a internet! Decidiu recompensar aqueles que participassem de sua pesquisa oferecendo-lhes um pequeno prêmio para provar sua teoria. Incentivados pela recompensa, uma multidão de respondentes se formou. Resultado: uma grande massa de dados para analisar.

1 Introdução

Crowdsourcing refere-se à terceirização de tarefas bem definidas para uma multidão de indivíduos aptos a sua realização. A noção de *crowdsourcing* nasceu em 2006 e desde então tem sido amplamente adotado em vários domínios, beneficiando-se da inteligência e conhecimento de uma multidão de indivíduos. Na indústria, por exemplo, as soluções de *crowdsourcing* têm sido propostas e implementadas como uma forma de superar obstáculos que exigem inteligência humana em larga escala. Mais recentemente, adotado em pesquisas científicas, o *crowdsourcing* apresenta novas oportunidades para estudos qualitativos e quantitativos, fornecendo um meio para expandir estudos realizados em laboratório e experimentos controlados (RAND, 2012). Com o *crowdsourcing* é possível, por exemplo, construir *groundtruths* para avaliações de pesquisas acadêmicas, além de possibilitar o acesso a uma grande variedade de participantes com perfis demográficos desejados (IPEIROTIS, 2010) dentro de um período razoavelmente curto de tempo. Contudo, apesar dos benefícios trazidos pelo *crowdsourcing*, há de se considerar uma série de desafios como, por exemplo, a qualidade dos dados gerados.

Em 1907, em uma feira no condado de Plymouth na Inglaterra, um grupo de indivíduos foi encorajado a participar de um concurso para estimar o peso de um boi após ser abatido e ter seus órgãos removidos. Ao todo, 800 pessoas participaram do concurso, incluindo açougueiros e agricultores, alguns dos quais especialistas na tarefa proposta; enquanto muitos outros não possuíam experiência alguma. Ao analisar as estimativas, Sir Francis Galton, um estatístico inglês, observou que a estimativa média de 787 participantes (13 estimativas foram descartadas por serem ilegíveis) era muito próxima do peso exato. Notou-se também que qualquer estimativa individual escolhida aleatoriamente tinha uma chance de 50% de estar dentro de um intervalo $[-3,7\%, +2,4\%]$ da estimativa média do peso (GALTON, 1907). Esta descoberta sugeriu a noção, agora bastante popular, da *sabedoria das multidões* (também conhecida em inglês por *wisdom of crowds*).

Groundtruth

O termo *groundtruth*, muito utilizado na área de Computação, refere-se a dados obtidos através de observação e não por inferência.

A origem deste termo é comumente associada à de tomadas de fotos aéreas, uma vez que a foto aérea retrata a verdade (*truth*) que se encontra no solo (*ground*) em um momento específico de uma dada observação.

É possível, também, encontrar na literatura o termo *crowdtruth* para indicar a construção de *groundtruths* realizadas a partir do processo de *crowdsourcing*.

Um século após as descobertas surpreendentes do Sir Francis Galton, um esforço notável de *crowdsourcing* não obteve o mesmo sucesso em seu propósito – a busca e resgate de um notável cientista da informática, Jim Gray (HELLERSTEIN;

TENNENHOUSE, 2011). Em janeiro de 2007, Jim Gray desapareceu no mar enquanto navegava sozinho com planos para dispersar as cinzas de sua mãe a cerca de 30 milhas de distância da Golden Gate em São Francisco nos Estados Unidos. Ao perceber o seu desaparecimento, amigos e colegas de Jim Gray começaram a buscar formas de ajudar as autoridades a localizá-lo e resgatá-lo. A busca por Gray evoluiu extraordinariamente envolvendo aviões privados, satélites, análises de imagens automatizadas, simulações de corrente oceânica e computação humana multidimensional, juntamente com a Guarda Costeira dos Estados Unidos. A equipe de busca era composta por estudantes de pós-graduação, engenheiros, cientistas da computação, oceanógrafos, astrônomos e empresários, muitos dos quais sequer conheciam o cientista desaparecido. Os esforços obtiveram apoio financeiro, suporte de tecnologias de ponta, entre outros recursos, além de uma vontade de ajudar altruísta de diversas pessoas que nem conheciam o cientista desaparecido. Apesar do esforço, infelizmente, após aproximadamente três semanas infrutíferas, o time de busca concordou em encerrar a busca por Jim Gray.

Já em 2010, um grupo de pesquisadores da Universidade de Washington nos Estados Unidos mostrou que *Foldit*, um jogo online de multijogadores que envolve indivíduos não-cientistas na resolução de problemas de predições complexos, forneceu resultados extremamente úteis que eram compatíveis ou mesmo superavam o estado-da-arte (COOPER et al., 2010). A Figura 1 apresenta uma captura de tela da interface do jogo em execução. Em 2011, um grupo de jogadores do *Foldit* ajudou a decifrar a estrutura cristalina de um vírus de macacos que causa infecção pelo vírus da imunodeficiência humana e síndrome da imunodeficiência adquirida (HIV/AIDS) (KHATIB et al., 2011), um problema científico que permaneceu sem solução por quase 15 anos. Após aproximadamente três semanas, jogadores (vale frisar, não-cientistas) do “quebra-cabeça” foram capazes de produzir um modelo 3D preciso da enzima de forma colaborativa.



Figura 1. Captura de tela da interface do jogo Foldit em ação. Fonte: <http://lgdb.org/game/foldit>

Mais recentemente, em março de 2014, o voo MH370 da companhia aérea *Malaysia Airlines* desapareceu logo após decolar de Kuala Lumpur, Malásia, em direção a Pequim, China. Apenas algumas horas após o desaparecimento do avião, os esforços de busca já haviam sido realizados por aproximadamente 2,5 milhões de pessoas na Web usando o site Tomnod¹ em uma campanha de *crowdsourcing* que se tornou viral (FISHWICK, 2014). A Figura 2 apresenta a interface do site Tomnod que foi usado para tentar localizar o voo desaparecido. A campanha obteve a ajuda de mais de 8 milhões de usuários que verificaram surpreendentes 1.007.750 quilômetros quadrados de imagens de satélite de alta resolução fornecidos pela empresa de satélites comerciais, DigitalGlobe². Os usuários da Web anotaram manualmente milhões de imagens que poderiam ajudar a localizar o avião desaparecido. Equipes especializadas investigaram todas as anotações promissoras dos usuários na Web mas, apesar da mobilização e participação *crowdsourced* das pessoas, o avião desaparecido não foi encontrado.

¹ Tomnod, Disponível em: <http://www.tomnod.com/> Acesso em: 03/03/2018

² Digital Globe, Disponível em: <http://www.digitalglobe.com/> - acesso em: 25/03/2019

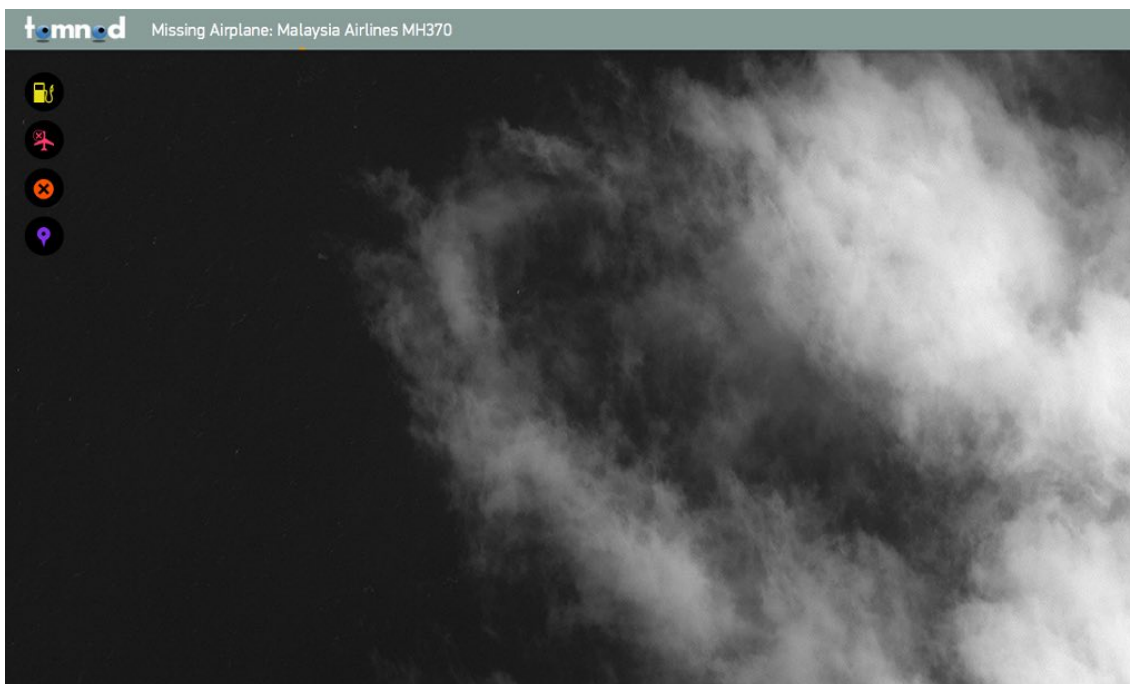


Figura 2. Captura de tela do site Tomnod mostrando a representação da interface de pesquisa para o voo MH370 da Malásia. No topo superior esquerdo da tela estão disponíveis ferramentas para identificação de áreas na imagem de satélite fornecida que possivelmente possam determinar a localização da aeronave desaparecida. O site Tomnod recebeu milhões de acesso de indivíduos com interesse em colaborar com as buscas através da Web, um verdadeiro esforço de uma multidão. Fonte (Fishwick 2014).

Como podemos notar a partir dos casos verídicos apresentados, a motivação por parte da participação *crowdsourced* pode variar muito – desde altruísmo a incentivos gamificados. Neste capítulo, abordaremos um domínio distinto do *crowdsourcing*, onde os participantes são principalmente motivados por recompensas monetárias associadas a realização de tarefas específicas, aqui referidas como paradigma de *crowdsourcing* pago. Este paradigma é amplamente utilizado em pesquisas científicas de forma a usufruir do conhecimento da multidão em determinado domínio para coleta de dados e construção de *groundtruths*.

2 Computação Humana e *Crowdsourcing* Pago

A computação humana, também conhecida como computação por humanos, pode ser definida como um paradigma para a utilização do poder de processamento humano na resolução de problemas que computadores ainda não são capazes de resolver (QUINN; BEDERSON, 2011). Este conceito foi introduzido pela primeira vez por Von Ahn através do jogo ESP (VON AHN; DABBISH, 2004; VON AHN, 2008). Von Ahn afirma que algumas tarefas como, por exemplo, a de reconhecimento de imagens, podem ser bastante triviais para seres humanos, mas muito desafiadoras para algoritmos sofisticados de reconhecimento de imagens. Embora seja uma tarefa árdua e difícil, Von Ahn sugere que a anotação manual de imagens é um dos métodos mais adequados para obter descrições de imagens precisas. Através do jogo ESP, Von Ahn notou que a experiência de anotação de imagens mediadas através de jogos poderia ser muito agradável e que seres humanos

realizavam as anotações de imagens sem perceber que esta era a tarefa a ser realizada.

O *input* humano pode ser adquirido através de uma variedade de mecanismos de incentivos implícitos ou intrínsecos; usando gamificação para incentivar a participação de pessoas como no jogo ESP, ou altruísmo, como observado na campanha de *crowdsourcing* para encontrar o voo da Malásia desaparecido. Como alternativa, a participação também pode ser estimulada através de incentivos materiais explícitos. O paradigma de *crowdsourcing* pago surge então como resultado da grande necessidade e alto potencial de aquisição de *input* humano para resolver diversos problemas que computadores possuem dificuldades ou mesmo não são capazes de solucionar. A Figura 3 apresenta algumas plataformas de *crowdsourcing* pagas que podem ser utilizadas para solucionar tarefas através de *input* humano. Uma comparação entre algumas das principais ferramentas descritas na Figura 3 em termos de qualidade e confiança é apresentada por Peer et al. (2017). As plataformas Upwork³ e Freelancer⁴ permitem que usuários postem tarefas online e encontrem outros usuários capazes de completá-la em troca de um pagamento monetário correspondente a tarefa a ser executada. Da mesma forma, 99designs⁵ e DesignCrowd⁶ são plataformas que se utilizam do paradigma do *crowdsourcing* pago para realizar tarefas de desenho gráfico online.



Figura 3. Alguns exemplos de plataformas de *crowdsourcing* que atendem a diversos mercados e tarefas distintas oferecendo incentivos monetários para a participação de indivíduos na Web.

Ao longo dos anos, testemunhamos um aumento na adoção do paradigma de *crowdsourcing* pago para resolver problemas que exigem inteligência humana em grande escala. Este aumento foi largamente impulsionado por plataformas de *crowdsourcing* de microtarefas⁷, como o Mechanical Turk⁸ da Amazon (AMT) e o Figure Eight⁹. As

³ UpWork, Disponível em: <http://www.upwork.com/> Acesso em: 03/03/2018

⁴ Freelancer, Disponível em: <http://www.freelancer.com/> Acesso em: 03/03/2018

⁵ 99Designs, Disponível em: <http://99designs.com/> Acesso em: 03/03/2018

⁶ Design Crowd, Disponível em: <http://www.designcrowd.com/> Acesso em: 03/03/2018

⁷ Fragmentos de tarefas que podem ser realizadas por indivíduos distintos e paralelamente e, conseqüentemente, terminadas mais rapidamente.

⁸ Mechanical Turk, Disponível em: <http://www.mturk.com/> Acesso em: 25/03/2019

⁹ Figure Eight (antigo CrowdFlower), Disponível em: <https://www.figure-eight.com> Acesso em:

plataformas de microtarefas *crowdsourcing* reúnem solicitantes (*requesters*) e trabalhadores (*crowd workers*) de todo o mundo. Os *solicitantes* são aqueles que disponibilizam as microtarefas em plataformas de *crowdsourcing* e descrevem suas necessidades e requisitos específicos para reunir e fazer uso das respostas dos trabalhadores. Já os *trabalhadores* são os participantes dispostos a completar as microtarefas disponibilizadas nas plataformas de *crowdsourcing* cumprindo os requisitos previamente definidos pelos solicitantes em troca de recompensas monetárias. O fluxo padrão de trabalho em plataformas de *crowdsourcing* de microtarefas é ilustrado na Figura 4.

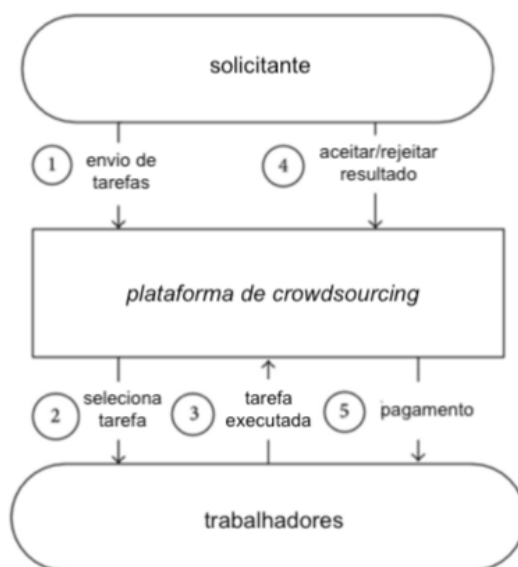


Figura 4. Fluxo padrão de trabalho em uma plataforma de crowdsourcing de microtarefas pagas. Fonte: (Samimi e Ravana 2014).

Como funciona o fluxo de trabalho

Os solicitantes postam tarefas conforme suas necessidades na plataforma *crowdsourcing*, com uma recompensa monetária explícita para incentivar a conclusão bem-sucedida da tarefa requerida. Já os trabalhadores escolhem as tarefas em que desejam trabalhar a partir da lista de tarefas disponíveis no sistema e realizam a tarefa de acordo com as descrições dadas pelos solicitantes (LITTLE et al., 2009). Com base na qualidade das respostas fornecidas pelos trabalhadores, os solicitantes aceitam ou rejeitam as respostas fornecidas. No caso de uma conclusão bem-sucedida das tarefas e aceitação por parte dos solicitantes, os trabalhadores recebem a recompensa correspondente através da plataforma. Assim, as plataformas de *crowdsourcing* desempenham um papel vital na união de solicitantes que estão à procura de *input* humano, e milhares de indivíduos dispostos a contribuir e completar tarefas em troca de pequenas quantias de dinheiro.

3 Desafios e Oportunidades no *Crowdsourcing*

Nesta seção exploraremos as características inerentes do *crowdsourcing* que precisam ser consideradas para executar experimentos válidos centrados no ser humano e na multidão. Apresentaremos também uma série de desafios e oportunidades para realização de sua pesquisa.

3.1 Tarefas de inteligência humana através de *crowdsourcing*

As tarefas de *crowdsourcing* podem servir a uma variedade de objetivos diferentes, que vão desde a geração e coleta de dados até a construção de *groundtruths* para avaliação e validação de pesquisas. Embora o *crowdsourcing* venha sendo usado como uma ferramenta para resolver problemas que requerem inteligência humana ou *input* humano em larga escala, ao longo dos últimos anos, pesquisadores começaram a considerar o paradigma de *crowdsourcing* pago como uma ferramenta em potencial para executar experimentos científicos que antes eram conduzidos e restritos a laboratórios (CRUMP et al., 2013; HORTON et al., 2011; MASON; SURI, 2012; PAOLACCI et al., 2010). A facilidade com que uma população representativa e diversificada pode ser obtida é uma grande vantagem quando se trata da validade de um experimento científico.

Uma classificação interessante dos tipos de tarefas mais comumente realizadas através de *crowdsourcing* pago foi proposta em (GADIRAJU et al., 2014), são elas:

- **Busca por informação:** onde os trabalhadores são encarregados de pesquisar na Web ou em outros meios por informações específicas.
- **Verificação e validação:** onde os trabalhadores são encarregados de verificar certos aspectos conforme as instruções fornecidas ou de confirmar a validade de um dado conteúdo.
- **Criação de conteúdo:** onde os trabalhadores têm a tarefa de criar novos conteúdos.
- **Interpretação e análise:** onde a sabedoria e as habilidades de interpretação da multidão são necessárias para conclusão de uma determinada tarefa.
- **Surveys:** onde é necessário obter informações, opiniões, de uma multidão.
- **Acesso ao conteúdo:** onde os trabalhadores devem simplesmente acessar algum conteúdo para atender ao propósito da tarefa solicitada.

3.2 Colaboração entre trabalhadores

Em um cenário padrão de *crowdsourcing*, cada trabalhador normalmente contribui de forma independente para o resultado final da tarefa requerida. No entanto, se um experimento precisar da colaboração entre trabalhadores, o cenário de *crowdsourcing* deve ser adaptado. Os “Jogos com um propósito” (GWAPs) são um bom exemplo dessa colaboração, onde indivíduos colaboram para resolver diferentes problemas, que vão desde a marcação de imagens (VON AHN; DABBISH, 2004) até a identificação de estruturas de proteínas (KHATIB et al., 2011). Trabalhos recentes mostram que competições podem ser efetivas para melhorar o rendimento de tarefas em larga escala (MORRIS et al., 2012; DONTCHEVA et al., 2014; DERGOUSOFF; MANDRYK, 2015).

Por outro lado, nenhuma das principais plataformas de *crowdsourcing* (como o Mechanical Turk (AMT) da Amazon ou Figure Eight) facilitam a colaboração direta entre os trabalhadores, de modo que a coordenação entre indivíduos deve ser implementada e facilitada externamente à plataforma utilizada, *i.e.*, por quem solicita a tarefa. Embora o *crowdsourcing* possa ser útil e razoavelmente barato, adaptar ou criar novos ambientes pode ser bastante custoso e inviabilizar a realização da tarefa idealizada pelo solicitante.

Além disso, há de se considerar que *colaborar* exige normalmente a disponibilidade dos indivíduos em um horário específico, o que pode prejudicar a execução da tarefa, além de possibilitar o abandono daqueles que não podem contribuir dentro do prazo estabelecido. Por exemplo, quando meios de colaboração adequados não são empregados, um trabalhador pode ter que esperar por longos períodos de tempo até que outro trabalhador seja encontrado ou, no pior dos casos, um trabalhador com ótima reputação pode ser relacionado para trabalhar com um outro trabalhador de baixa reputação, afetando a qualidade do resultado da tarefa a ser executada e, possivelmente, inutilizando os resultados e desperdiçando a participação do trabalhador de boa reputação.

3.3 Tarefas baseadas em etapas

No caso de tarefas compostas por diferentes etapas, onde um requisito fundamental é envolver o mesmo conjunto de participantes em cada uma dessas etapas, o anonimato dos indivíduos que caracterizam o ambiente de *crowdsourcing* torna a execução de tais experimentos muito desafiadora. Portanto, se uma plataforma de *crowdsourcing* não divulgar dados de contato de seus trabalhadores ou não facilitar o contato direto a determinados indivíduos (por exemplo, via *e-mail* ou *chat*), uma solução possível, mas custosa, é redirecionar os trabalhadores para uma plataforma externa e personalizada, onde a informação necessária pode ser coletada.

Embora as plataformas de *crowdsourcing* facilitem a colaboração entre o solicitante e o trabalhador, algumas plataformas são mais apropriadas para coleta de dados em massa e outras para realização de trabalhos complexos e especializados. Antes de começar uma tarefa de *crowdsourcing*, identifique as plataformas disponíveis e aquelas que melhor atendem aos requisitos da tarefa pretendida.

3.4 Custos em tarefas *crowdsourced*

Um ponto importante a ser considerado no *crowdsourcing* é que cada trabalhador usa seu próprio dispositivo para realizar a tarefa disponibilizada. Assim, neste cenário, o solicitante da tarefa não tem controle direto sobre a configuração de *hardware* e *software* do trabalhador (GADIRAJU et al., 2017a) e, muito menos, dos ambientes em que os mesmos executam as tarefas. Isso pode ser particularmente prejudicial se o experimento precisar de *hardwares* especiais ou configurações de *software* específicas para garantir a validade dos resultados.

É importante considerar que em tarefas *crowdsourced* é bastante complicado impor qualquer tipo de controle sobre *softwares* ou *hardwares* com o objetivo de criar uma configuração uniforme entre os trabalhadores, ou torná-los mais parecidos com o cenário de uso da vida real. No entanto, é possível verificar a confiabilidade do *hardware* e *software* do trabalhador usando *scripts* que funcionem no dispositivo do trabalhador informando a configuração e outros parâmetros em termos de versão do navegador, sistema operacional, configurações de *hardware* entre outras informações consideradas relevantes. Assim, é possível pré-visualizar os trabalhadores que não satisfaçam aos requisitos necessários para realização da tarefa em questão e descartá-los do grupo de trabalhadores que atendam aos requisitos mínimos.

Por outro lado, o custo de configurar a tarefa de *crowdsourcing*, em termos de equipamentos, é praticamente zero. O único custo que precisamos levar em conta compreende o esforço da concepção da tarefa de modo que satisfaça as necessidades do experimento em questão. Mas lembre-se, conforme já mencionado, o custo pode aumentar substancialmente se um recurso específico precisar ser completamente desenvolvido devido à falta de suporte na plataforma de *crowdsourcing* escolhida. Vale considerar também que implementar um *software* compatível com vários navegadores da Web, suportando vários dispositivos, etc., pode ser bastante oneroso. Além disso, o processamento e análise dos resultados podem ocasionar custos extras caso os solicitantes não sejam capazes de realizá-los (GADIRAJU et al., 2017d). Por fim, se o paradigma de *crowdsourcing* pago for utilizado, também é necessário mensurar os gastos com os trabalhadores da tarefa.

Embora as tarefas realizadas via plataformas *crowdsourcing* sejam relativamente baratas, uma série de questões devem ser consideradas antes de iniciá-las, caso contrário, acarretará em gastos inúteis e sem validade nos resultados obtidos.

3.5 Trabalhadores

Vejam agora algumas das principais características inerentes do multidimensionamento de seres humanos decorridas do paradigma do *crowdsourcing*:

- **Quantidade:** um solicitante pode atingir uma população extremamente grande através de várias plataformas de *crowdsourcing*. Assim, experiências que anteriormente eram restritas à ordem de dezenas ou

centenas de participantes podem aumentar até a ordem de milhares de trabalhadores sem grandes alterações nos custos para realização da tarefa.

- **Disponibilidade:** experimentos de laboratório são tipicamente limitados pela disponibilidade de seus recursos. Através do *crowdsourcing* os participantes estariam disponíveis 24 horas por dia, (considerando participantes ao redor do mundo e em fusos horários diferentes) e o solicitante da tarefa não estaria necessariamente restrito por pelo menos esta limitação temporal.
- **Diversidade e alcance:** os trabalhadores que podem ser alcançados através de plataformas de *crowdsourcing* constituem uma população altamente diversificada, cobrindo uma ampla gama de atributos demográficos (idade, gênero, etnia, localização, etc.). Assim, tarefas ou experimentos centrados no ser humano podem se beneficiar dessa diversidade e, conseqüentemente, alcançar resultados mais representativos.
- **Qualidade e confiabilidade:** um dos maiores desafios do *crowdsourcing* é o controle de qualidade e a confiabilidade dos trabalhadores. Ao longo dos últimos anos, vários pesquisadores têm desenvolvido uma série de mecanismos de controle de qualidade em *crowdsourcing*, que vão desde métodos de desenvolvimento de tarefas (GADIRAJU et al., 2015b), até a pré-seleção de trabalhadores (GADIRAJU et al., 2017b), utilização de técnicas anti-spam como CAPTCHAs ou mesmo análises *post-hoc* através de testes, onde a cada grupo de tarefas realizadas, o trabalhador deve responder tarefas cujas respostas são conhecidas podendo ajudar a identificar usuários não-confiáveis (EICKHOFF; DE VRIES, 2013; MARSHALL; SHIPMAN, 2013). Portanto, embora existam custos adicionais para manter a confiabilidade dos participantes em experimentos centrados em seres humanos, é certamente possível alcançá-la.

3.6 Aspectos de desenvolvimento e controle de tarefas

Uma série de aspectos precisam ser considerados para exercer controle sobre experimentos centrados no ser humano ao usar plataformas de *crowdsourcing*. Vejamos alguns deles:

- **Design:** é necessário um esforço para projetar e especificar uma tarefa que seja adequada para a participação de profissionais em larga escala.
- **Incentivos:** uma variedade de recompensas tem sido utilizada para incentivar a participação de indivíduos em experimentos laboratoriais como, por exemplo, créditos em cursos, compensações monetárias, intenções altruístas, etc. Quando as plataformas de *crowdsourcing* são utilizadas com a finalidade de coletar dados a partir da experiência de humanos, a aquisição de participantes é realizada tipicamente através de

incentivos financeiros. Os custos envolvidos neste tipo de tarefa dependem da complexidade do experimento, do esforço exigido dos participantes e do tempo necessário para a conclusão da tarefa.

- **Interação e canais de *feedback*:** um dos fatores limitantes no *crowdsourcing* é a falta de interação entre o solicitante e os trabalhadores. As plataformas de *crowdsourcing* normalmente fornecem canais de *feedback* com flexibilidade limitada (por exemplo, através de *chats* ou *e-mails corporativos*). Assim, são necessários esforços adicionais do solicitante para garantir que os trabalhadores tenham entendido os objetivos da tarefa de forma satisfatória para realizá-la (MASON; SURI, 2012). Ao contrário de ambientes em laboratório, os trabalhadores não podem ser monitorados facilmente e há menor controle sobre o protocolo experimental.
- **Configurações específicas:** experimentos que requerem equipamentos específicos, dispositivos especiais ou que exigem que os participantes usem as mesmas configurações (por exemplo, resolução de tela, distância à tela, contextos etnográficos, *software/hardware* específicos, etc.), são menos adequados ao paradigma do *crowdsourcing*. Embora existam formas de pré-selecionar os profissionais da multidão para satisfazer os requisitos, isso requer um esforço adicional.
- **Otimização:** uma grande vantagem de executar experimentos usando *crowdsourcing* é o potencial para otimizar necessidades específicas como, por exemplo, a quantidade de tempo dentro do qual as respostas devem ser reunidas. Já se o critério mais importante é garantir respostas mais confiáveis do que em volume então a tarefa pode ser configurada na plataforma de forma que esses requisitos sejam atendidos (GADIRAJU et al., 2015b). Dependendo da configuração realizada, a tarefa pode levar mais ou menos tempo para ser finalizada. Uma forma de obter resultados rápidos e confiáveis é realizar uma filtragem através de análises *post-hoc*. A escalabilidade do *crowdsourcing* permite tais otimizações conforme os requisitos de cada tarefa.

4 Discussão

É importante saber que as plataformas de *crowdsourcing* existentes não são completamente adequadas para realizar experimentos centrados no ser humano, mas sim em resolver tarefas que computadores apresentem dificuldades. Plataformas que seguem os paradigmas de *crowdsourcing* estão em constante desenvolvimento e algumas plataformas voltadas para pesquisa acadêmica estão em ascensão. Tais plataformas têm contribuído de forma substancial na resolução de tarefas que necessitam da inteligência humana.

A falta de controle direto e supervisão sobre os trabalhadores em tarefas em larga

escala faz com que o controle de qualidade seja obrigatório para obtenção de resultados válidos. Uma gama de mecanismos e técnicas foram propostas para validar resultados e incluem a criação de perguntas para testar a qualidade do trabalhador (*gold standards* (DIFALLAH et al., 2012; EICKHOFF; DE VRIES, 2013; OLESON et al., 2011)), outras para verificar a atenção do mesmo além de verificações de consistência e métodos psicométricos (MARSHALL; SHIPMAN, 2013), métricas comportamentais do trabalhador e de *design* da tarefa ideal (GADIRAJU et al., 2015b), *feedback* (DOW et al., 2012; GADIRAJU et al., 2015a) além da otimização de tarefas (CHENG et al., 2015; GADIRAJU et al., 2015c; MASON; SURI, 2012).

Várias técnicas foram desenvolvidas a fim de aumentar a eficiência e eficácia dos trabalhadores e maximizar a relação custo-benefício da implantação de tarefas multidimensionadas (ROKICKI et al., 2014). A gamificação, por exemplo, é um método que melhora a retenção de trabalhadores e a produção de tarefas (FEYISETAN et al., 2015). Desta forma, há de se considerar a adequação do método à tarefa a ser realizada. Sem dúvidas o *crowdsourcing* pode contribuir em diversas atividades, tanto na indústria como na academia.

No que tange o domínio educacional, a utilização de *crowdsourcing* para coleta de informações ou realização de experimentos pode ser encontrado em diferentes cenários. Jiang et al. (2018) descreve alguns desses cenários propondo uma taxonomia para classificar o *crowdsourcing* na educação de acordo com o seu uso. A taxonomia engloba tarefas para construção de conteúdos, práticas, trocas de conhecimento, além de avaliações em larga escala realizadas pela multidão. É evidente que existe uma gama de possibilidades de uso do *crowdsourcing* na Educação, principalmente no caso de pesquisas quantitativas, onde a abundância de dados é fundamental para se obter resultados significantes. Contudo, algumas precauções devem ser tomadas para garantir que o processo de ensino-aprendizagem seja conduzido com sucesso e as tarefas realizadas pelos trabalhadores da multidão sejam de qualidade. Alguns trabalhos discutem como tornar tarefas robustas para evitar trapaças (EICKHOFF; DE VRIES, 2013).

5 Cenário Ilustrativo – Analisando o ganho de conhecimento através de sessões de máquinas de busca

Máquinas de busca na Web são frequentemente utilizadas para aquisição de novos conhecimentos e satisfazer objetivos relacionados à aprendizagem. Contudo, pouco ainda se sabe sobre como o conhecimento de um usuário evolui ao longo de uma sessão de busca. Em Gadiraju et al. (2018), o *crowdsourcing* foi utilizado para simular sessões de pesquisa com o objetivo de investigar como o conhecimento de usuários na Web evoluem ao longo de uma sessão de busca. A utilização do paradigma de *crowdsourcing* foi essencial para permitir o recrutamento de 500 usuários únicos durante a pesquisa. Através de pré e pós-testes, foi possível verificar e quantificar o ganho de conhecimento dos usuários recrutados via plataforma de *crowdsourcing*.

Máquinas de busca não disponibilizam as sessões de seus usuários por uma série de motivos, incluindo privacidade. Dessa forma, investigar o processo de buscas na Web

se torna um grande desafio. Com o uso de *crowdsourcing*, foi possível identificar uma série de aspectos relativos à busca, bem como obter uma amostra representativa de usuários por um baixo custo. Como resultado do cenário de pesquisa descrito, os usuários da plataforma *crowdsourcing* utilizada, em média, obtiveram um maior ganho de conhecimento em tópicos com os quais estavam menos familiarizados.

6 Resumo

O uso de *crowdsourcing* tem aumentado ao longo dos anos para solucionar tarefas simples e do dia-a-dia a um baixo custo e alta eficiência. Na Informática na Educação este pode ser um grande aliado para condução de pesquisas que necessitem de abundância de dados, o que é normalmente um grande problema nesta área. Com a facilidade de se coletar dados, construir aplicações, validar tarefas, entre outras tarefas árduas que permeiam o processo de ensino e aprendizagem, o conhecimento de multidões pode ser bastante útil.

Neste capítulo apresentamos uma visão geral do uso do *crowdsourcing* utilizando inteligência humana em larga escala para realizar tarefas de forma rápida e confiável que computadores ainda não são capazes de solucionar. Inicialmente discutimos os conceitos de computação humana bem como o paradigma do *crowdsourcing* (pago). Em seguida, apresentamos o uso do paradigma de *crowdsourcing* e suas principais características em conjunto com uma taxonomia dos tipos de tarefas *crowdsourcing* mais comuns e utilizadas no mercado. Além disso, formas de colaboração entre trabalhadores e formas de estimar e gerenciar custos em tarefas *crowdsourcing* também foram apresentadas. Finalmente, discutimos o design de tarefas *crowdsourcing* bem como o seu controle de qualidade permitindo que as tarefas executadas sejam bem-sucedidas.

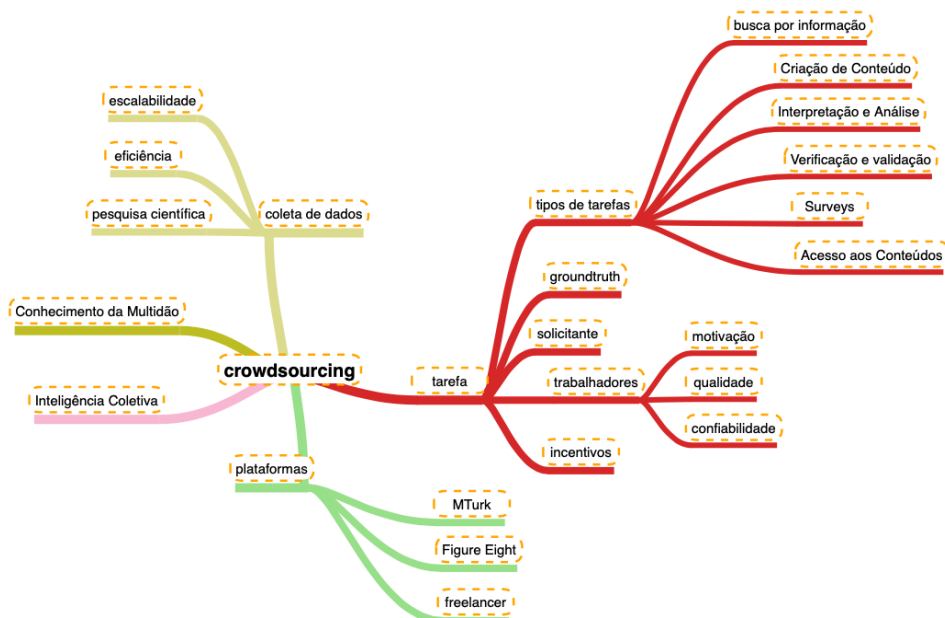


Figura 5. Mapa conceitual da noção de crowdsourcing abordada neste capítulo.

7 Leituras Recomendadas

Human computation. (LAW; AHN, 2012). Este livro define a área de pesquisa de computação humana além de fornecer uma revisão da literatura existente, criando conexões para uma ampla variedade de disciplinas, incluindo IA, Aprendizado de Máquinas, HCI, Market Design e Psicologia.

Evaluation in the Crowd (ARCHAMBAULT et al, 2017). Este livro discute experiências e considerações metodológicas ao usar plataformas de *crowdsourcing* para executar experimentos.

An introduction to hybrid human-machine information systems (DEMARTINI, 2017). Este livro apresenta uma visão geral de diferentes aplicações para os quais os sistemas de informação humano-máquina foram utilizados nos domínios de recuperação de informações, processamento de linguagem natural, Web semântica, aprendizado de máquina e multimídia.

8 Artigos Exemplos

A pilot study of using crowds in the classroom (DOW et al., 2013). Neste trabalho, os autores propuseram o uso do *crowdsourcing* em salas de aula para melhorar o processo de aprendizagem dos alunos, recebendo *feedback* sobre o material didático.

Training workers for improving performance in crowdsourcing microtasks (GADIRAJU et al., 2015). In: Design for Teaching and Learning in a Networked World, 2015, Toledo. Anais... New York: Springer, 2015. p. 100–114. Neste trabalho, os autores invertem o papel de trabalhadores *crowdsourcing* e os tratam como “aprendizes” em um novo contexto de aprendizagem. Os autores propõem uma fase de aprendizagem de curta duração e aplicação imediata de conceitos aprendidos.

A Review on Crowdsourcing for Education: State of the Art of Literature and Practice (JIANG et al., 2018). Este trabalho apresenta uma breve revisão sistemática de artigos que usam o conceito de *crowdsourcing* na educação. A leitura é recomendada aos leitores que desejam buscar ideias de uso e práticas envolvendo o conceito de *crowdsourcing* diretamente ligado a educação.

9 Checklist

Em qualquer atividade realizada via *crowdsourcing* é necessário verificar uma série de aspectos para garantir que sua tarefa seja bem-sucedida. Nesta seção sugerimos um *checklist* para ajudá-lo a realizar sua tarefa em plataformas *crowdsourcing*.

Busque por referências: Antes de começar a desenvolver sua tarefa, busque por referências nas plataformas de *crowdsourcing* existentes. O fato de existirem tarefas similares a sua é um bom indicador de que a sua tarefa pode ser atendida por esse paradigma. Caso não encontre, reflita sobre o seu problema e verifique se o *crowdsourcing* é a melhor opção.

Planeje sua tarefa: A execução de experimentos, ou desenvolvimento de qualquer trabalho, exige um esforço grande no planejamento da tarefa. Um planejamento mal feito pode resultar em perdas significativas de dinheiro. Por isso, certifique-se de que tenha respostas bem-definidas para algumas perguntas básicas, como, por exemplo:

- Quem é meu público-alvo? Ele pode ser atingido através de plataformas *crowdsourcing*?

- A tarefa a ser executada requer ambientes/*hardware/software* específicos? Se sim, é possível atingi-los via *crowdsourcing*?

- A tarefa a ser executada é complexa? Se sim, é necessário utilizar uma plataforma *crowdsourcing* com usuários especialistas?

- Quantos usuários são necessários para realizar a minha tarefa ou validar meu experimento? Posso recursos suficientes para tal? Caso não obtenha, posso utilizar outras estratégias? Quais?

Teste sua tarefa: Antes de iniciar sua tarefa em uma plataforma *crowdsourcing*, faça um teste com pessoas próximas a você. Isso irá permitir que identifique erros de *design* da sua tarefa.

Algumas plataformas disponibilizam um *checklist* baseado em suas funcionalidades. Recomendamos que as leia com atenção. Um exemplo é fornecido pela plataforma Figure Eight¹⁰, bastante utilizada em pesquisas científicas.

A Figura 5 ilustra um processo simplificado para criação de tarefas baseadas no paradigma do *crowdsourcing*.

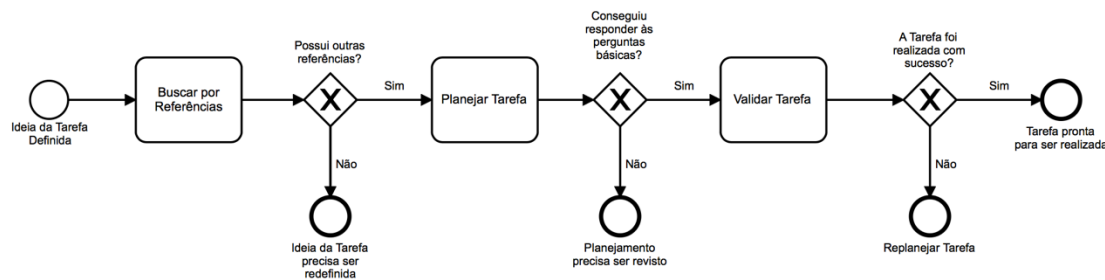


Figura 6. Processo simplificado para criar uma tarefa em plataformas *crowdsourcing*.

¹⁰ <https://success.figure-eight.com/hc/en-us/articles/202703195-Job-Launch-Checklist>

10 Exercícios

1. Selecione um trabalho qualquer na literatura de Informática na Educação e pondere se a avaliação conduzida poderia ser realizada através do paradigma do *crowdsourcing*. Levante aspectos positivos e negativos bem como as dificuldades existentes para condução do experimento usando plataformas *crowdsourcing*.
2. Liste soluções criativas usando inteligência humana e que não seriam possíveis de realizar sem uma plataforma de *crowdsourcing*. Use como referência o caso da companhia aérea *Malaysia Airlines* descrito anteriormente.
3. Quais aspectos fundamentais devem ser levados em consideração na criação de uma tarefa *crowdsourcing*?
4. Teste o poder do paradigma do *crowdsourcing* na prática. Escolha um livro qualquer de sua biblioteca pessoal, talvez este livro mesmo e, em um grupo de pelo menos 30-40 pessoas, peça que estimem o número total de páginas do livro em questão. Analise as respostas e margem de erro média da população em relação ao número de páginas real do livro. Compare com os números obtidos pelo Sir Francis Galton. Leve em consideração que a amostra dele era de 800 participantes.
5. Suponha que por algum motivo relevante você tenha que transcrever uma grande quantidade de artigos científicos pois eles estão em um formato não editável. Como são muitas páginas a serem transcritas e o trabalho bastante tedioso, a tarefa não poderá ser terminada no prazo determinado se realizada sozinha. Como você poderia usar *crowdsourcing* para solucionar o problema da transcrição de artigos de forma que os participantes se sintam motivados? Reflita sobre a relação entre o engajamento dos participantes nas tarefas e a qualidade do trabalho realizado via *crowdsourcing*.
6. O paradigma de *crowdsourcing* pago pressupõe recompensas financeiras em troca de tarefas realizadas. Como um trabalhador pode tentar burlar o processo de *crowdsourcing* pago para receber suas recompensas financeiras sem realizar adequadamente as tarefas? Como prevenir que as tarefas sejam burladas pelos trabalhadores?

Referências

- ARCHAMBAULT, D.; PURCHASE, H.; E HOSSFELD, T. **Evaluation in the Crowd**. Dagstuhl: Springer, 2017.
- CHENG, J.; TEEVAN, J.; AND BERNSTEIN, M. S. **Measuring crowdsourcing effort with error-time curves**. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, 33, 2015, Seoul. Anais... New York: ACM, 2015. p. 1365–1374.
- COOPER, S. et al. **Predicting protein structures with a multiplayer online game**. *Nature*, v. 466, n. 7307, p. 756–760, 2010.
- CRUMP, M. J.; MCDONNELL, J. V.; GURECKIS, T. M. **Evaluating amazon’s mechanical turk as a tool for experimental behavioral research**. *PloS one*, v. 8, n. 3, 2013.
- DERGOUSOFF, K.; MANDRYK, R. L. **Mobile Gamification for Crowdsourcing Data Collection: Leveraging the Freemium Model**. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, 33, 2015, Seoul. Anais... New York: ACM, 2015. p. 1065–1074.
- DEMARTINI, G.; DIFALLAH, D. E.; GADIRAJU, U.; CATASTA, M. **An introduction to hybrid human-machine information systems**. In: Foundations and Trends in Web Science, v. 7, n. 1, 2017. p 1-87.
- DIFALLAH, D. E. et al. **Scaling-up the crowd: Micro-task pricing schemes for worker retention and latency improvement**. In: Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, 2, 2014, Pittsburgh. Anais... Palo Alto: AAAI Press, 2014. p. 50 – 58.
- DIFALLAH, D. E.; DEMARTINI, G.; CUDRE-MAUROUX, P. **Mechanical cheat: Spamming schemes and adversarial techniques on crowdsourcing platforms**. In: Proceedings of the International Workshop on Crowdsourcing Web Search, 1, 2012, Lyon. Anais... CEUR-WS.org, 2012. p. 26–30.
- DONTCHEVA, M. et al. **Combining crowdsourcing and learning to improve engagement and performance**. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, 32, 2014, Toronto. Anais... New York: ACM. p. 3379 – 3388.
- DOW, S. et al. **Shepherding the crowd yields better work**. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, 15, 2012, Seattle. Anais... New York: ACM. p. 1013–1022.
- DOW, S.; GERBER, E.; E WONG, A. **A pilot study of using crowds in the classroom**. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2013, Paris. Anais... New York: ACM, 2013. p. 227–236.
- EICKHOFF, C.; DE VRIES, A. P. **Increasing cheat robustness of crowdsourcing tasks**. *Information retrieval*, v. 16, n. 2, p. 121–137, 2013.
- FEYISETAN, O. et al. **Improving paid microtasks through gamification and adaptive**

- furtherance incentives.** In: Proceedings of the International Conference on World Wide Web, 24, 2015, Florence. Anais... New York: ACM. p. 333–343.
- FISHWICK, C. **Tomnod – the online search party looking for malaysian airlines flight MH370.** *The Guardian*, 2014. Disponível em: <<https://www.theguardian.com/world/2014/mar/14/tomnod-online-search-malaysian-airlines-flight-mh370>>. Acesso em: 12 abril 2019.
- GADIRAJU, U. et al. **Modus operandi of crowd workers: The invisible role of microtask work environments.** *ACM Interact. Mob. Wearable Ubiquitous Technol.*, v. 1, n. 3, p. 49:1 – 49:29, 2017a.
- GADIRAJU, U.; DIETZE, S. **Improving learning through achievement priming in crowdsourced information finding microtasks.** In: Proceedings of the International Learning Analytics & Knowledge Conference, 7, 2017, Vancouver. Anais... New York: ACM. p. 105–114.
- GADIRAJU, U.; FETAHU, B.; KAWASE, R. **Training workers for improving performance in crowdsourcing microtasks.** In: Design for Teaching and Learning in a Networked World, 2015a, Toledo. Anais... Cham: Springer, 2015. p. 100–114.
- GADIRAJU, U. et al. **Using worker self-assessments for competence-based pre-selection in crowdsourcing microtasks.** *ACM Transactions on Computer-Human Interaction*, v. 24, n. 4, p. 30:1 – 30:26, 2017b.
- GADIRAJU, U.; KAWASE, R.; DIETZE, S. **A taxonomy of microtasks on the web.** In: Proceedings of the ACM conference on Hypertext and social media, 25, 2014, Santiago. Anais... New York: ACM. p. 218–223.
- GADIRAJU, U. et al. **Understanding malicious behavior in crowdsourcing platforms: The case of online surveys.** In: Proceedings of ACM Conference on Human Factors in Computing Systems, 33, 2015b, Seoul. Anais... New York: ACM. p. 1631–1640.
- GADIRAJU, U. et al. **Crowdsourcing versus the laboratory: Towards human-centered experiments using the crowd.** In: Proceedings of the Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments, 2017c, Dagstuhl. Anais... Cham: Springer, p. 6–26.
- GADIRAJU, U. et al. **Breaking bad: Understanding behavior of crowd workers in categorization microtasks.** In: Proceedings of the ACM Conference on Hypertext & Social Media, 26, 2015c, Cyprus. Anais... New York: ACM. p. 33–38.
- GADIRAJU, U.; YANG, J.; BOZZON, A. **Clarity is a worthwhile quality—on the role of task clarity in microtask crowdsourcing.** In: Proceedings of the ACM Conference on Hypertext & Social Media, 28, 2017d, Prague. Anais... New York: ACM. p. 5–14.
- GADIRAJU, U. et al. **Analyzing knowledge gain of users in informational search sessions on the web.** In: Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval, 3, 2018, New Brunswick. Anais... New York: ACM. p. 2–11.
- GALTON, F. **Vox populi (the wisdom of crowds).** *Nature*, v. 75, n. 7, p. 450 – 451,

1907.

- HEINZELMAN, J.; WATERS, C. **Crowdsourcing crisis information in disaster-affected Haiti**. *US Institute of Peace Washington, DC*, 2010. Disponível em: <<https://www.jstor.org/stable/resrep12220>>. Acesso em: 12 abril 2019.
- HELLERSTEIN, J. M.; TENNENHOUSE, D. L. **Searching for Jim Gray: a technical overview**. *Communications of the ACM*, v. 54, n. 7, p.77–87, 2011.
- HORTON, J. J.; RAND, D. G.; ZECKHAUSER, R. J. **The online laboratory: Conducting experiments in a real labor market**. *Experimental Economics*, v. 14, n. 3, p. 399–425, 2011.
- IPEIROTIS, P. G. **Demographics of mechanical turk**. *NYU Working Paper No.; CEDER-10-01*. Disponível em: <<https://ssrn.com/abstract=1585030>>. Acesso em: 12 de abril de 2019.
- JIANG, Y.; SCHLAGWEIN, D.; BENATALLAH, B. **A Review on Crowdsourcing for Education: State of the Art of Literature and Practice**. In: Proceedings of the Pacific Asia Conference on Information Systems, 22, 2018, Yokohama. Anais... AISEL, 2018. p. 180-193.
- KHATIB, F. et al. **Crystal structure of a monomeric retroviral protease solved by protein folding game players**. *Nature structural & molecular biology*, v. 18, n. 10, p. 1175–1177, 2011.
- LAW, E.; E AHN, L. V. **Human computation**. In: Synthesis Lectures on Artificial Intelligence and Machine Learning, v. 5, n. 3, p. 1–121, 2012.
- LITTLE, G. et al. **Turkit: tools for iterative tasks on mechanical turk**. In: Proceedings of the ACM SIGKDD workshop on human computation, 2009, Paris. Anais... New York: ACM, 2009. p. 29–30.
- MARSHALL, C. C.; SHIPMAN, F. M. **Experiences surveying the crowd: Reflections on methods, participation, and reliability**. In: Proceedings of the ACM Web Science Conference, 5, 2013, Paris. Anais... New York: ACM, 2013. p. 234–243.
- MASON, W.; SURI, S. **Conducting behavioral research on amazons mechanical turk**. *Behavior research methods*, v. 44, n. 1, p. 1–23, 2012.
- MORRIS, R. R.; DONTCHEVA, M.; GERBER, E. M. **Priming for Better Performance in Microtask Crowdsourcing Environments**. *IEEE Internet Computing*, v. 16, n. 5, p. 13-19, 2012.
- OLESON, D. et al. **Programmatic gold: Targeted and scalable quality assurance in crowdsourcing**. *Human computation*, v.11, n. 11, (2011).
- PAOLACCI, G.; CHANDLER, J.; IPEIROTIS, P. G. **Running experiments on amazon mechanical turk**, *Judgment and Decision Making*, v. 5, n. 5, 2010.
- PEER, E. et al. **Beyond the Turk: Alternative platforms for crowdsourcing behavioral research**. *Experimental Social Psychology*, v. 70, p.153-163, 2017.
- QUINN, A. J.; BEDERSON, B. B. **Human computation: a survey and taxonomy of a growing field**. In: Proceedings of the SIGCHI conference on human factors in computing systems, 2011, Vancouver. Anais... New York: ACM, 2011. p. 1403–

1412.

- RAND, D. G. **The promise of mechanical turk: How online labor markets can help theorists run behavioral experiments.** *Journal of theoretical biology*, v. 299, p. 172–179, 2012.
- ROKICKI, M. et al. **Competitive game designs for improving the cost effectiveness of crowdsourcing.** In: Proceedings of the ACM International Conference on Conference on Information and Knowledge Management, 23, 2014, Shanghai. Anais... New York: ACM, 2014. p. 1469–1478.
- SAMIMI, P.; RAVANA, S. D. **Creation of reliable relevance judgments in information retrieval systems evaluation experimentation through crowdsourcing: a review.** *The Scientific World Journal*, v. 2014, p. 1 – 13, 2014.
- VON AHN, L. **Human computation.** In: IEEE International Conference on Data Engineering, 24, Cancun. Anais... New York: IEEE, 2008. p. 1–2.
- VON AHN, L.; DABBISH, L. **Labeling images with a computer game.** In: Proceedings of the SIGCHI conference on Human factors in computing systems, 2004, Vienna. Anais... New York: ACM, 2004. p. 319–326. ACM. 2004

Sobre os autores



Ujwal Gadiraju

<http://www.l3s.de/~gadiraju/>

Ujwal Gadiraju é atualmente pesquisador no L3S Research Center, Leibniz Universität Hannover, na Alemanha, onde também completou seu doutorado em informática (2017). Ujwal possui mestrado em informática pela Delft University of Technology (2012) na Holanda. Seus interesses abrangem os campos de Computação Humana e *Crowdsourcing*, Recuperação de Informações, Computação Social, Computação Humana e Mineração de Dados na Web. Publicou mais de 45 trabalhos em conferências e periódicos de alto impacto.



Bernardo Pereira Nunes

<http://lattes.cnpq.br/1728746187630338>

Bernardo possui graduação em Engenharia da Computação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) (2006), mestrado em Informática pela PUC-Rio (2009) e doutorado em Informática pela PUC-Rio com período sanduíche na Leibniz Universität Hannover (2014). Atualmente é pesquisador e professor da PUC-Rio e Professor Colaborador da Pós-Graduação em Informática Aplicada na UNIRIO. Seus interesses abrangem as áreas: Web Semântica, Ciência da Web e Informática na Educação. Além disso, atuou como pesquisador nos laboratórios de pesquisa da L3S Research Center na Alemanha (2012-2013 e 2017), pesquisador visitante no CNR, Itália (2015) e pesquisador visitante na Universidade Nacional da Austrália (2018), ganhou prêmios nacionais e internacionais e publicou mais de 100 artigos em conferências e periódicos de alto impacto.