



# Improvement of Online Job Markets through Machine Learning Techniques

**Elinamar Santiago<sup>1</sup>-Álamo, Angel Ojeda-Castro<sup>2</sup>**

<sup>1</sup>Doctorate Student School of Business and Entrepreneurship, Universidad del Turabo, Puerto Rico

<sup>2</sup>Associate Professor, Management Information Systems, Universidad del Turabo, Puerto Rico

---

**Abstract:** This article proposes a methodological framework based on process virtualization theory and review of recent literature to improve the problem of information asymmetry that occurs when processes are carried out without physical interaction. The findings indicate that Crowdsourcing platforms are tools to people searching for jobs and generate revenues in a fast and effective way, and has unique advantages that allows the client and the worker to benefit from each other from any part of the world. On the other hand, there are still many gaps and regulations that need to be revised in this type of work, such as legal aspects. Virtualization processes must be improved to achieve the success and satisfaction of clients and workers during the implementation of projects.

**Keywords:** Crowdsourcing, eLancing, Machine Learning, Future of Work, Information Asymmetry, Process Virtualization

---

## I. INTRODUCTION

Crowdsourcing has been increasing its popularity in recent years, and is expected to be the most common working method over the traditional work model (work in an office from 8 am to 5 pm, vacations and benefits paid by the employer) in a few years. The current workforce and the one that is coming up, needs a work model that meets the expectations of the traditional work model. Therefore, any gap in crowdsourcing should be improved continuously. The following article is based on a review of recent literature, includes the proposal for a theoretical framework based on process virtualization theory, and suggests how to improve the quality of work on the marketplace through machine learning techniques. Low-quality jobs remain a very common problem in online job markets, and negatively affect the relationship and experience between clients and workers. This could culminate in demands; have to repeat a job, and the dissatisfaction of one or both parties in the contract. According to Overby (2008), more processes are carried out every day in a virtual way; without any physical interaction, and the method of working and generate an income is not the exception. The Process virtualization theory of Overby (2008) contains four variables that measure how susceptible can be a process that does not require any physical interaction. This is the reason why this theory is the basis for developing a theoretical model that helps to specifically improve the problem of information asymmetry between the client and the worker.

### 1.1 Objective

The objective of this article is to propose a theoretical framework that seeks to demonstrate or suggest how information asymmetry can be eliminated from online labor markets.

### 1.2 Justification

Many authors have addressed the importance of research on crowdsourcing, as it offers the opportunity to produce an investigation that has the potential to be valuable to researchers and society in general. Making a contribution to society and literature is one of the greatest challenges and desires of researchers, and the future of work is a topic that should be further investigated, as it can be of great help and contribute positively to society; such is the case of Puerto Rico, which is going through one of its most challenging economic moments. According to the Bureau of Labor Statistics of the United States, the unemployment rate in the island is 11.5% to March 2017 and the government recently approved a labor reform that generated significant changes in the traditional work model. It is important not only to improve processes, but also to educate the population about crowdsourcing, to become familiar with it, and to really see in this model a real and effective alternative to combat unemployment.

## II. LITERATURE REVIEW

Thomas Malone and Robert Laubacher are the creators of the concept of eLancing, being the first to declare the term in 1998 in his article "The Dawn of E-lance economy". According to the Knowledge for



decision making report (2014), the leading platforms in online labor markets were founded in 1998 and began doing business in 1999. Forbes magazine calls this type of market “an escape for unemployed professionals”, particularly amid the global economic crisis. According to Kittur et al. (2013), crowdsourcing is a socio-technical work system constituted through a set of relationships that connect organizations, individuals, technologies and work activities. The most common task in crowdsourcing is micro-work, but every day new and larger tasks are added. Some of the tasks that have been added are legal services, accounting services, distance learning, graphic design and scheduling, administrative assistants, among many other tasks. There is still some caution with the implementation of large projects, so it is necessary to innovate and constantly improve the services of the global online market, since in a few years this will be the most common method of work, above the traditional work model.

According to Machado et al. (2016), there has been a growing interest in applying artificial intelligence (AI) techniques in traditional systems to solve common problems. Kamar (2016) suggests that hybrid intelligence systems offer new fields of application for micro-tasks such as those currently being seen in the online labor market and the demand for human intelligence tasks can increase with hybrid intelligence models. The main purpose of combining human intelligence tasks with artificial intelligence can be of great help in reducing the high unemployment rates currently available worldwide. According to Kamar (2016), “there is an opportunity to create crowdsourcing platforms that generate value for AI systems and humans, but this depends on overcoming the problems that continue to arise, such as the identity of the worker, the benefits and the rights of workers”. Den Besten et al. (2014), mentions that in Mechanical Turk (AMT), the workers select the tasks to be done between the ones available. Often, the employer requires workers to take a test to make sure the worker can perform the task. Most transactions in the online labor market are based on trust and reputation. That's why having a good behavior is one of the most important factors in the market. According to Wang et al. (2012b), workers who pass the qualification test to obtain a contract, deliver better quality jobs, compared to contracts that are signed without aptitude testing first.

According to Lehdonvirta & Mezier (2013), one of the advantages that the online working model has on the traditional work model is that there are no physical work sites, hours of work or other behavioral controls. The entry mode for many of the eLancing markets is by taking a commission directly from the employee. The percentage of the commission varies, but fluctuates between 5% and 15%. Micro-task markets are the only type of market that requires employers to pay for the service directly. Gareis (2014) mentions some important components of the eLancing:

1. Freelancers are free agents whose working life is organized in a sequence of projects that can last only a few hours to several months or even more than a year.
2. Freelancers market their services, according to their preferences and skills.
3. To perform complex tasks, freelancers create virtual organizations. These organizations are dissolved once the task has been fulfilled.
4. For marketing and collaboration, freelancers rely on internet-based e-markets.
5. Freelancers are free from the limitations of traditional employees.

According to Kittur et al. (2013), the crowdwork has the potential to support a flexible workforce and mitigate challenges such as shortages of experts in specific areas or geographic locations. However, some disadvantages in the history of the work of the crowdwork are abuses by the client, extremely low pay for the work that is carried out; without benefits or protections for the worker, among other things. Several authors have addressed as one of the advantages the potential of the market to break the barrier of time and space, allowing people from all over the world to work and collaborate among them. However, this same advantage causes information asymmetry problems such as language barriers, difference in culture and time zone, among others. For example: a client from the United States who hires the services of a worker living in the same country, it will communicate more quickly and effectively, than hiring the services of a person who lives in a different time zone and does not dominate the same language the client (or the crowd work) does. Avoiding geographic separation could lead to a project to be completed in or before the established time.

## **2.1 Forecast**

In 2010, the Intuit Report makes several predictions of demographic, social, economic and technological tendencies that will occur during the period from 2010 to 2020. Many of the predictions are made in technology and telecommunications. Some of the predictions are the following:

1. Cultural fusion brings global tastes to local markets – The adoption and adaptation of global traditions in local habits will emerge as a growing trend due to the use of the Web. This trend shows how quickly people can get products and services, breaking language and culture barriers over the Internet. This prediction made in 2010 is currently reflected in the online job market. Therefore, the importance



of improving the way of acquiring services through the Internet is vital, so that the experience for both the client and the worker is satisfactory.

2. Social networks fuel the participatory economy – Forums such as TurkOpticon, serve as a community for workers in the online global marketplace to exchange and share ideas, experiences and knowledge in general. Workers often use these forums to warn other workers about clients who have bad habits such as not paying well, refusing to pay, not properly describing the tasks of a project, among others.
3. Localism creates a new way of life – This is one of the reasons why many workers have abandoned their current jobs and have chosen the autonomy air that provides the power to work independently, allowing them to devote more time to the family, and to themselves.
4. Customers control the relationship – This trend can be applied to the online job market, as the report predicts the influence of social networks and other online information sources on purchase decision information for clients and business to business transactions. Maintaining a good reputation will be the key to success in the marketplace.
5. Work changes from full-time to free employment of agents – It is important to improve recruitment and selection methods by reducing hiring time. A good way to achieve this would be through machine learning techniques and algorithms that show the best profiles of the workers to the client (reputation, experience, quality of work, among others) according to the requirements of the task to be performed.
6. Market niches flourish in the new economy – This could be one of the solutions to the problem of information asymmetry in online labor markets. Large platforms like UpWork, Fiverr, MTurk, among others, are a huge breakthrough and have a lot to offer, but these platforms are an immense sea where millions of people enter daily in search of a contract. This leads too many new talents that try to join the gig economy lose interest and leave the market. However, market niches focus on a specific area and dramatically reduce language, culture, and geographic dispersal barriers.
7. Working in the cloud – Work in the cloud will move more and more towards a regime where people work in their own place, and at their own time. The way people work and do business will change and allow a growing variety of web-based technologies. This is something we see now in the transition process from the traditional work model to the online work model.

## 2.2 Information Asymmetry

According to Brawley & Pury (2016), good communication is an essential key to the success of the virtualization processes in the market. Despite the many benefits of the online work model, there is still a lot to do. As mentioned in the Freelance Marketplace Industry Competitive Analysis (2014), the term information asymmetry is the lack of effective communication between the client and the worker. This is a problem that cannot be neglected, because the lack of communication can take various forms:

1. Front-End: The client does not define the scope of the project sufficiently well, leading to unaligned expectations.
2. Scope-Creep: The customer does not understand the depth of scope of a project sufficiently well, leading to demands, find solutions and repeat a project.

According to Allahbakhsh (2013), the customer may set some criteria in motion to ensure that only eligible individuals can perform the task or specify assessment and compensation policies. Even high-quality workers may present low-quality contributions due to misunderstandings or mistakes. Because of this, it is important that both the client and the worker have an effective communication during the project. The choice of appropriate incentives and a compensation policy can affect workers' performance as well as the quality of the results. Knowing the policies of assessment and compensation will help workers align their work based on these criteria and produce contributions with higher quality.

## 2.3 Global Frictions

According to Yili & Pavlou (2013), the language barrier is one of the biggest challenges in the global online marketplace, as effective communication rests with the dominance of a common language between customers and workers. The client-worker transaction in the same country can generate better communication; streamline processes, and better ideas. Correct use of words can eliminate ambiguity, reduce communication costs, and avoid redundancy at work. Another important factor to consider is frequent transactions. Geffen & Carmel (2008), mention that companies are more likely to select the workers they have previously hired. From the customer's perspective, if a company is satisfied with the service of a worker, it is likely to continue to use its services. Subsequent transactions tend to be more polished and fast as the vendor knows the needs of the company.



According to Yili & Pavlou (2013), as a company acquire experience in the market, learns to avoid losses by opportunistic or low-quality workers, and have more realistic expectations about the quality and reputation of workers. In the same way, it happens with the workers, who as they gain experience, learn to avoid the opportunistic companies or clients, which do not pay well in the established time. When there are many offers available, finding a worker who offers more satisfaction is a formidable task. The authors categorized the problems of geographic dispersion, complexity of information technology services, language barrier, cultural differences and time zone differences as global frictions, as it is difficult to monitor a worker for geographical reasons and effort. The worker's reputation plays an extremely important role in mitigating this series of problems caused by global frictions. Therefore, a customer not only observes worker prices, but also features such as their reputation.

Yili & Pavlou (2013), suggest that companies select a service provider to maximize their benefits, and are less likely to select a worker who speaks a language or live in a different time zone. The main factors that constitute a global friction are the following:

1. Lack of familiarity
2. Different Language
3. Different Time Zone
4. Different Culture

#### **2.4 Reputation Signals**

Reputation signals are indicators of the quality of service providers. Freelancer.com has an established reputation system, which includes received average rating valuation, experience in gold member projects and status. The results indicate that service providers are very sensitive to lack of familiarity and language difference. Beyond global friction, the best suppliers of reputation signals are more likely to be hired by the company, as well as ask more money for a project and still get the contract (Yili & Pavlou, 2013). According to Allahbakhsh (2013); quality is a subjective topic in general. Some efforts have proposed models and metrics to evaluate quantitatively and objectively the quality along different dimensions of a software system, such as reliability, accuracy, relevance, integrity and consistency. The quality of the results of a crowdsource task may be affected by the capacities and quality of the workers. The trust relationship between a client and a worker reflects the likelihood that the applicant expects to receive a quality contribution from the worker. Reputation scores are built primarily in community members activities in the system. One of the measures to improve quality is that customers announce the rules of the game clearly in the description of the task and announce penalties if deficiencies are observed. However, Franklin et al. (2011) warns that refusing to pay may provoke a backlash from workers, who qualify customers in forums (Ipeirotis, 2010b).

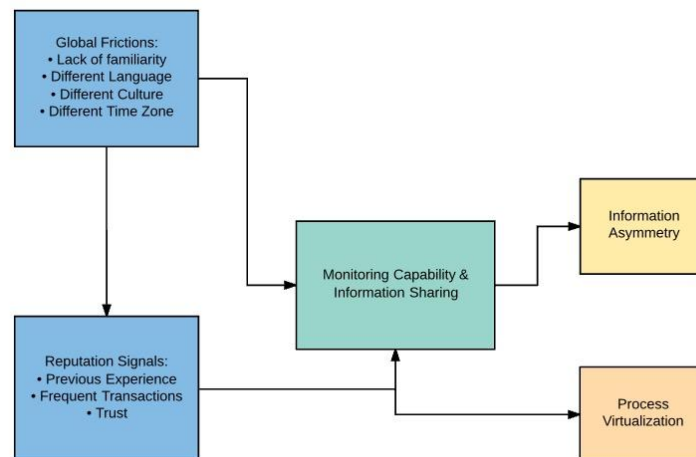
#### **2.5 Monitoring Capability**

According to Aguinis & Lawal (2013) and the Freelance Marketplace Industry Competitive Analysis, information asymmetry is one of the main factors affecting virtualization processes in the market. To worsen the situation, geographical separation makes monitoring capability virtually impossible. When a client and worker are geographically dispersed, monitoring performance becomes a difficult one, allowing the worker to circumvent its obligations, resulting in a moral and conflicting risk.

### **III. CONCEPTUAL FRAMEWORK**

Based on the recent literature review and process virtualization theory, the following theoretical framework consists of global frictions, reputation signals, monitoring capability/information sharing and process virtualization. The theoretical framework suggests the following:

1. Global frictions negatively affect monitoring capability and information flows slower when there are, for example, time zone differences, which increases the problem of information asymmetry.
2. Reputation signals positively affect monitoring capability because there are variables as trust, which is an important key to build a relationship where there is no physical interaction. However, communication could be ineffective with global frictions in between, even if there are reputation signals favoring the worker.
3. To avoid information asymmetry, it is necessary that there are no global frictions in the middle, and only reputation signals in the relationship between client and worker to a faster and more efficient virtualization process.



Source: Own Elaboration

#### IV. DISCUSSION

Rhyn & Blohm (2017) are the first to demonstrate that it is possible to reliably explain and predict the quality of crowdsourcing contributions based on the textual characteristics of the data alone. They provide empirical evidence of the relationship between contextual and representative characteristics of contributions and their quality in crowdsourcing. This indicates that well developed and accurate solutions, ideas, or suggestions are vital for companies that try to take advantage of the information presented by the workers. In addition, its results suggest that companies require textual contributions to be presented in a clear and easy to interpret way to fully benefit from them. The authors contribute with a set of models and variables to put these contextual and representative characteristics into operation. The models and variables proposed in his study have proven to work well with algorithms capable of automatically evaluating and classifying textual contributions. Provide the basis for partially automating the evaluation of textual data in crowdsourcing, which has been frequently requested by related literature. These findings can help researchers select variables for the predictive model in crowdsourcing.

Further research on the online work model can contribute much to theory and discipline as it has changes in how we see work today and how to improve online work processes. There are many gaps with the online working model that still must be filled. According to Aguinis & Lawal (2013), the eLancing is a model of work that has the potential to change the way work works around the world and is becoming a powerful economic phenomenon worldwide. Moreover, due to its singular type of work organization, it has the potential to change labor relations in the International Society of the 21st century. There are many economic, cultural, technological, and social changes that can encourage companies to adopt the online work model, as well as millions of people around the world are doing it daily. This study would also help companies implement and adapt the benefits of the traditional work model to the online work model.

#### V. RECOMMENDATIONS AND FUTURE RESEARCH

Shared information is a very important aspect of the online job market since sometimes geographical separation can be a global fiction, it is important that the worker enjoys a good reputation and a client confidence to mitigate disadvantages. As it happens on eBay or Amazon, where the most prominent and most renowned sellers are the ones who sell the most, in online labor markets the reputation is built through a good feedback between client and worker.

Brawley and Pury (2016) make the following recommendations for future research:

1. More research on job satisfaction and quality to understand the impact and results in crowdsourcing in long term.
2. Additional research on the crowdsource experience, focusing on the interaction between clients and workers.
3. To investigate the experience of MTurk workers in different countries.
4. Investigate the experience of workers on other platforms apart from MTurk.

According to Kittur et al. (2013), a vision of the future for crowdsource implies the following:

1. Considerations of workers such as motivation, feedback and good pay, as well as mechanisms to maintain reputation, provide better interaction with customers, and increase motivation. It has been



mentioned in several research articles the need to improve remuneration in this type of labor market, as well as the interaction between the employer and the worker, which is one of the main suggestions of this article.

2. Considerations on the part of the client, such as coordination, decomposition of tasks and quality control. This is something that can be achieved and needs to be improved.
3. Better theoretical models, markets, or automatic processes could drastically reduce development costs and address search friction, an important issue in the labor economy.
4. Crowdsourcing management could involve a stratified tree composed of worker leaders, clients, machine learning systems and algorithms.
5. It is possible to design algorithms and workflows to guide workers to quickly complete the synchronous tasks.
6. More and better communication options with the customer, something better than email, a kind of immediate conversation.
7. Hybrid systems could harness the best of human intelligence and artificial intelligence.
8. It may be possible to design machine learning algorithms that understand human nature more thoroughly and can more directly model the cost-performance tradeoff: For example, using a combination of active and semi supervised learning to collect the most informative labels.
9. Create a trusted network in which customers and workers validate each other as trustworthy. The creation of technical tools to share information about workers or customers should be combined with more robust systems to monitor and report abuses by the client.
10. A self-sustaining cycle could involve artificial intelligence and automate some of the simplest tasks.
11. Online Mentoring systems, perhaps augmented with human tutorials, could provide a path to education that is more scalable to the public.

According to Rhyn & Blohm (2017), automated planning has interesting mechanisms to generate plans that can contribute significantly to the description and standardization of other prerequisites that define a task and a user in crowdsourcing applications. The adoption of classical planning techniques can improve the allocation of crowdsourcing tasks in a deliberate crowd. One solution is for customers to specify the minimum cost and minimum skill required for their work. Another solution could be hybrid support for both the allocation and the programmed approaches (Machado et al., 2016). Allahbakhsh (2013) suggests that many problems and open challenges remain to define measure and manage quality in crowdsourcing systems, and these topics require further research. Another important limitation of existing quality control approaches comes from the subjective nature of quality, particularly in crowdsourcing systems. A solution to this limitation is a recommendation system that could use machine learning techniques to provide more accurate recommendations.

## VI. CONCLUSION

Crowdsourcing platforms are tools to people search for jobs and generate revenues in a fast and effective way, and has unique advantages that allow the client and the worker to benefit from each other from any part of the world. The authors who have addressed the topic have made different recommendations according to the findings in their articles. Although the online work model has advanced to an impressive level in recent years, there is no doubt that this working model has too many opportunities for further research. It is very important to continue improving the virtualization processes for the benefit of customers and workers who seek to generate an income through this method.

There is still some skepticism in people who have not entered the online work model for lack of trust in virtual platforms, such as fear that they do not get the pay after performing a task. It would be advisable that apart from the techniques of machine learning, algorithms, mining of text, among others that can improve the experience of the online work that also does not lose sight to the people who still do not make the transition of the traditional work model to the model of work crowdsource. According to Aguinis and Lawal (2013), there are still many gaps and regulations that need to be revised in this type of work such as legal aspects. Virtualization processes must be improved to achieve the success and satisfaction of clients and workers during the implementation of projects. Brawley & Pury (2016) demonstrate that some constructs, relationships and recommendations of the traditional work model can be applied to the online work model.

## REFERENCES

- [1]. Alice M. Brawley, C. L. (2016). Work Experiences on MTurk: Job Satisfaction, Turnover, and Information Sharing. *Computers in Human Behavior*, 531-546.
- [2]. Angela Zhou, I. C. (2016). Dynamic Task Allocation for Crowdsourcing Settings. *International Conference on Machine Learning*.



- [3]. Aniket Kittur, J. V. (2013). The Future of Crowd Work. CSCW, (pp. 1-16). San Antonio, Texas.
- [4]. Gareis, K. (2014, September 24). Retrieved from Research Gate: <https://www.researchgate.et>
- [5]. Harris, M. (2016, June 6). Say Goodbye to your highly skilled job. It's now a "Human Intelligence Task". Retrieved March 2, 2017, from Back Channel: <https://backchannel.com>
- [6]. Herman Aguinis, S. O. (2013). Conducting field experiments using e-Lancing's natural environment. *Journal of Business Venturing*, 493-505.
- [7]. Herman Aguinis, S. O. (2013). ELancing: A review and research agenda for bridging the science-practice gap. *Human Resource Management Review*, 6-17.
- [8]. (2010). Intuit 2020 Report. Intuit.
- [9]. Kamar, E. (2016, March 1). Microsoft Research. Retrieved April 28, 2017, from Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence. <https://www.microsoft.com>
- [10]. Kamar, E. (2016). Hybrid Intelligence and the Future of Work. Retrieved March 2, 2017, from Microsoft Research: [www.microsoft.com](http://www.microsoft.com)
- [11]. Kevin Yili Hong, P. A. (2013). The IBIT Report. Philadelphia, PA: Institute for Business and Information Technology, Temple University.
- [12]. Leticia Machado, R. P. (2016). Task Allocation for Crowdsourcing using AI Planning. SCI-SE'16, (pp. 1-5). Texas.
- [13]. Making, K. f. (2014, February 21). Freelance Marketplace Industry Competitive Analysis. Retrieved from Knowledge for Decision Making: <http://www.k4dm.com>
- [14]. Making, K. f. (2014, February 7). Introduction to Freelance Marketplace Infrastructure - A Business Reference Guide. Retrieved from Knowledge for Decision Making: <http://www.k4dm.com>
- [15]. Marcel Rhyn, I. B. (2017). A Machine Learning Approach for Classifying Textual Data in Crowdsourcing. 13th International Conference on Wirtschaft's Informatik (pp. 1171-1185). St. Gallen, Switzerland: Institute of Information Management.
- [16]. Matthijs Den Besten, C. M.-M. (2014). Human Computing via Online Labor Markets. The Perils and Promises of Crowdsourcing in Data-Rich Ecosystems. Madrid: Instituto de Políticas y Bienes Públicos (IPP).
- [17]. Mohammad Allahbakhsh, B. B. (2013). Quality Control in Crowdsourcing Systems. *IEEE Computer Society*, 76-81.
- [18]. Overby, E. (2008). Process Virtualization Theory and the Impact of Information Technology. *Organization Science*, 277-291.
- [19]. Statistics, B. o. (2017, March). United States Department of Labor. Retrieved May 16, 2017, from [www.bls.com](http://www.bls.com)
- [20]. Thomas Malone, R. L. (1998). The Dawn of e-Lance economy. . *Harvard Business Review*, 145-152.
- [21]. Vili Lehdonvirta, P. M. (2013). Dynamics of Virtual Work: Identity and Self-Organization in Unstructured Work. *European Cooperation in Science and Technology*, 1-35.
- [22]. Wald, J. (2016, December 12). Predictions for the Freelance Economy. Retrieved from Forbes: <http://www.forbes.com/sites/waldleventhal/2016/12/12/5-predictions-for-the-freelance-economy-in-2017/#3a5121c587ae>