

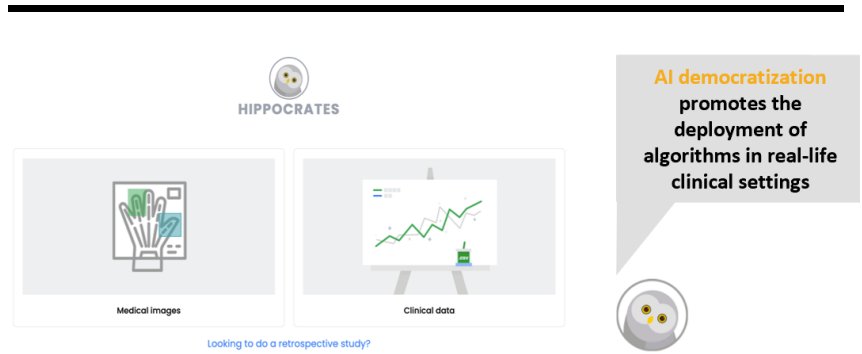
# Hippocrates AutoML: Democratizing AI in favor of medicine

---

Clinical decision support tools from Arkangel AI



Arkangel AI  
1250 Guy St Suite #600,  
Montreal, Quebec H3H 2L3  
hola@arkangel.ai  
www.arkangel.ai



## Introduction

Since its first documented appearance in 1956, the term “Artificial Intelligence” has been used to exemplify the application of complex algorithms in the process automation of different industries<sup>[1]</sup>. From farming, education, and criminal law to climate change mitigation and healthcare industry applications, Artificial Intelligence (AI) is dramatically benefiting society on a daily basis<sup>[2]</sup>. In particular, for the healthcare industry, AI-powered medical technologies are rapidly evolving into practical solutions for all aspects of clinical practice, allowing, among others, the automated diagnosis/prognosis of diseases, the tailoring of specific treatments, and the optimization of the use of medical resources<sup>[3]</sup>.

Nevertheless, the pervasive role of AI has also spurred some worries about how much these “new algorithms” can be trusted. In fact, the most recent Global CEO survey of PWC found that, although 85% of CEOs consider AI to be bigger than the internet revolution, 84% agree that AI-based decisions need to be explainable to be trusted<sup>[4]</sup>. In that sense, researchers and policymakers have proposed “AI democratization” as a means to create a level of trustability since it brings users closer to the algorithms and ensures that the benefits of this technology are not limited to a small group of people.<sup>[5]</sup>

In general terms, “to democratize something” implies to make it virtually accessible to every citizen of the world regardless of their language, color, race or creed, place of origin, or gender<sup>[6]</sup>. Specifically, AI democratization implies building, accessing, and using AI models and algorithms without requiring advanced mathematical and computing science skills. This democratization, in the words of Mark Riedl, associate professor in the College of Computing and Associate Director of the Machine Learning Center at GeorgiaTech, “means that more people are able to conduct research and build AI-driven products and services.”<sup>[7]</sup>

However, although AI democratization can be a game-changer for all industries, it is significantly relevant in healthcare as it promotes the final

goal of AI in the industry: the deployment of algorithms in real-life clinical settings.

In that sense, to contribute to the AI democratization, ArkangelAI created Hippocrates AutoML, an AI-based platform specialized in medical purposes. Hippocrates erases the barriers of access to AI in healthcare settings by allowing the creation of algorithms without writing a single line of code.

The rest of this white paper is structured as follows: The first section will present the importance of AI democratization in healthcare. The second one will describe how to democratize AI for medical purposes by using the Hippocrates AutoML tool, highlighting how easy it is to build AI solutions using the platform and how it helps overcome AI general challenges. Finally, the third section will show a glance into Arkangel AI's perspective of the future of AI Democratization using Hippocrates.

## 2. Why democratize AI for medical purposes?

As it does in other industries, artificial intelligence has the potential to make medical labor more efficient and effective while lowering barriers to entry for new work by reducing dependence on domain-specific expertise<sup>[8,9]</sup>. However, although medical AI-based technology seems to be well received by patients<sup>[10]</sup>, there still exists a specific resistance by healthcare professionals to its implementation. This resistance has been based mainly on the “black box” that Artificial Intelligence represents for physicians, nurses, bacteriologists, and other health workers. In particular, three main reasons have been identified. First, healthcare students are not being prepared for mastering existing and future technologies like Artificial Intelligence.<sup>[11]</sup> Second, although the current literature consensus is that AI is designed to complement physician intelligence<sup>[12,13]</sup>, there is an increasing fear that AI will replace physicians and other health workers.<sup>[14]</sup> And Third, currently, a well-defined legal framework that establishes the use cases for the adoption of algorithms recommendations is still lacking, exposing physicians to potential legal outcomes when using AI.<sup>[15]</sup>

That said, it is clear that there is a need to bring health workers closer to AI technologies, which is precisely what AI democratization seeks to offer. In general terms, one can identify three potential benefits of AI democratization to any industry<sup>[16]</sup>. First, it can reduce the entry barriers for people and organizations to start working with AI, seeking to exemplify the benefits of its application in diverse daily situations. Second, once the benefits are understood, democratization can reduce the costs of experimenting with AI solutions, improving the performance of the algorithms, and boosting the creation of helpful frameworks for specific

developing tasks. Finally and foremost, it can speed up the adoption of AI solutions which is the main goal in real life.

Now, in particular for the healthcare industry, apart from bringing healthcare workers closer to the technology, AI democratization must:

- Promote the integration of courses into the health sciences curriculum.
- Showcase the benefits of implementing AI algorithms as helpers for daily healthcare activities,
- Invite healthcare workers to be at the forefront in defining the legal framework of the use of algorithms.

But, how to democratize AI for medical purposes?

### 3. How to democratize AI for medical purposes

Although it was previously said that democratization requires giving universal access to knowledge, effective AI democratization doesn't imply just giving healthcare workers all the AI-related information expecting them to turn it into gold<sup>[17]</sup>. To properly accomplish all the goals stated above, the democratization process must be correctly structured, seeking to maximize the use given to the available resources.

In order to structure the AI democratization process, it is imperative to understand what Artificial Intelligence is. In simple terms, one can define AI as the ability of machines to replicate the human brain's intelligence. Nonetheless, in real life, AI is a broad research activity that ultimately aims to give a computer program the ability to learn or solve specific problems independently, simulating natural human behavior<sup>[18]</sup>. To provide computers with intelligence, a person needs to know how to interact with the computer, i.e., programming, and need to master all the technical aspects related to the problem the machine needs to solve or optimize.

For the healthcare industry, it is undeniable that physicians, specialists, nurses, microscopists, bacteriologists, and other health workers master all the knowledge related to a particular health challenge. Therefore, the technical aspects of the problem the algorithm needs to solve are already covered. On the contrary, while the digital transformation of the health care system is advancing, the minimum knowledge in programming seems to be lacking in most health staff,<sup>[11]</sup> making "writing code" the main barrier for AI democratization.

Consequently, Arkangel AI has identified that the cornerstone to democratize AI for medical purposes is providing healthcare workers the technical tools to structure, develop and deploy AI models without learning how to program. For that, we have created Hippocrates AutoML, an AI-based platform that accompanies healthcare workers in all the steps required to build algorithms that are usable in real life.

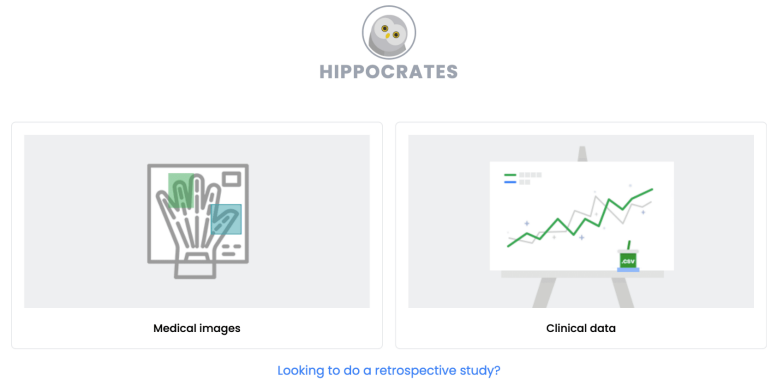


Figure x. Hippocrates AutoML platform

Once the problem is identified, by defining the relevant data and the expected output of the algorithm, four well-defined steps remain to build the algorithm: Data preparation and labeling, training, testing, and deployment. To go through each step without writing a single line of code, Hippocrates is designed as follows:

### Data Preparation and labeling

In the AI world, data is the mainstay of algorithms. As such, it requires to be collected, clean, and augmented to be useful for the training process.<sup>[19]</sup> Data collection is particularly challenging in the healthcare ecosystem since most medical data is confidential and requires specific security. Nonetheless, as Hippocrates is a cloud-based platform, all the data received is secured under international standards like HIPAA compliance, provided by all cloud support providers.<sup>[20,21]</sup>

Furthermore, the medical data collected by the experts must be correctly labeled so that Hippocrates can automatize the cleaning and augmentation process. Labeling implies defining the “gold standard” the algorithm must achieve in order to show the spect performance in a clinical setting. In general terms, Hippocrates receives the previously labeled data as a CSV file containing all the relevant information to analyze the images, videos, or clinical data. However, it will depend on the specific clinical application.

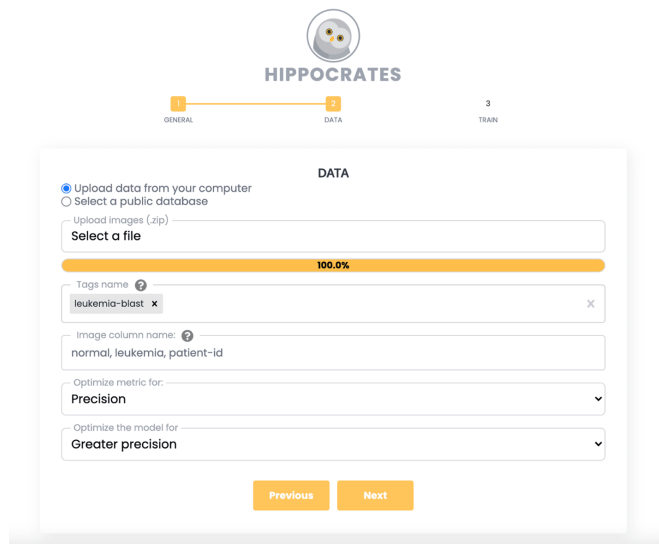


Figure x. How to upload the data to Hippocrates AutoML

Hippocrates can filter noisy elements (cleaning) in the labeled data so downstream model training is not compromised and defines the correct augmentation functions. Here it is essential to mention that the augmentation process is also critical in healthcare as data must correctly represent the studied disease or clinical situation<sup>[22]</sup>.

## Training

The training process starts with clean and augmented data. The first step is to split into training, validation, and testing sets as standard in AI for medical applications. This process, which avoids overfitting, is also automatic in Hippocrates so that the user does not need to write code to achieve it. Nevertheless, users must be aware that Hippocrates choose each of the cohorts following specific considerations in medicine<sup>[23]</sup>, like using the patient ID and not each of the entries in the splitting.

Hippocrates AUTOML carries out a hyperparameter optimization with the training set by varying architectures such as DenseNet121, Xception, MobileNet, InceptionV3, and InceptionResNetV2. Additionally, parameters such as learning rate or specific optimizer parameters are tuned by running a bayesian driven search algorithm for each architecture. For each of these configurations, Hippocrates trains a basic model seeking to identify the optimal model for the medical situation. Once the best setup is found for each architecture, models with the lowest validation loss are picked for being tested with the held-out testing cohort from the previous step.

## Testing

Testing AI models implies showing never-seen images to the model to calculate performance using well-defined metrics. Once Hippocrates performs the testing for all the best algorithms, the top 5 models that yielded the best performance are displayed as an output for the user. Nevertheless, although ArkangelAI's tech team recognizes that testing is crucial in all AI applications, we believe that it is of particular interest in healthcare since it defines the real potential of the AI-based solution. In that sense, not only can Hippocrates test the algorithm with the testing cohort (retrospective study), but it also is designed to receive prospective data (using the same preparation mentioned above) to calculate the real-life performance metrics of the algorithm. This aspect is vital since only the prospective metrics can show the actual utility of AI systems, as performance is likely to be worse when encountering real-world data that differ from that experienced in algorithm training<sup>[24]</sup>.

## Deployment

Finally, even if the algorithm fulfills all the performance requirements provided by the specialist, it is still not ready to be used in real life. For taking the algorithm into a clinical setting, it must be deployed as a new service. Consequently, Hippocrates can deploy the final model using ArkangelAI's platform so that users, once again, do not need to worry about writing code.

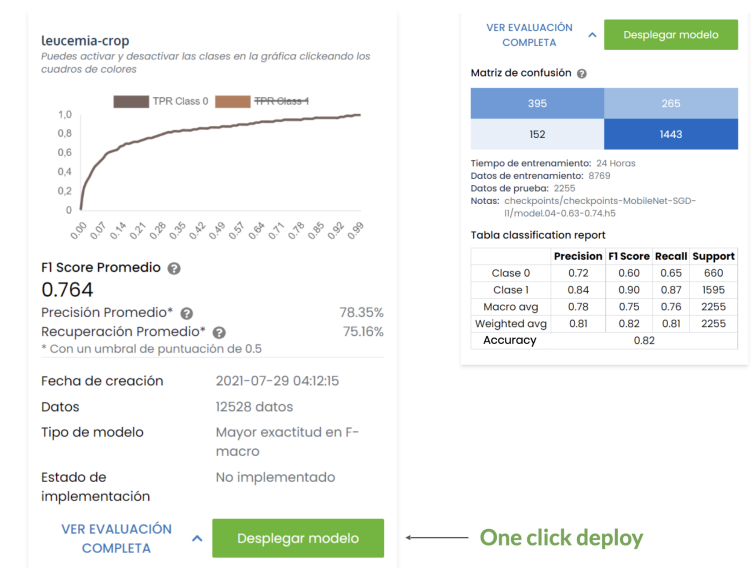


Figure x. Model deployment possibility of Hippocrates AutoML

Without the need to write code, which is a time-consuming process and the main barrier to access AI, we seek to urge healthcare workers to understand the conceptual bases of artificial intelligence inputs and outputs to get the most out of this technological tool. In fact, with the use of Hippocrates AutoML, not only is the time required to build an AI-based solution 10X less than the classical approaches, but it is also more affordable since no tech team needs to be hired and no hardware needs to be bought.

That said, Arkangel AI is aware of the existence of other great platforms for streamlining the process of custom AI model creation. Nevertheless, although those platforms have shown excellent results for different industries, unlike Hippocrates, they have not been optimized for medical purposes. As indicated in the Hippocrates description, the platform not only complies with all the AI requirements, but it also follows all the recommended steps for building an AI-based model for medical purposes, which are vital in order to make the model suitable for deployment in a clinical setting.

## 4. Final thoughts

The creation of a platform like Hippocrates AutoML can reduce the entry barrier for people (particularly healthcare workers) to start working with Artificial Intelligence as it reduces the demand for programming engineers and enables domain experts to automatically build AI applications without much requirement for statistical and AI knowledge.<sup>[19]</sup> By erasing the programming challenge, which is innate in AI, Arkangel AI is bringing the healthcare industry closer to this type of technological solution, and therefore is paving the way to full and genuine democratization of Artificial Intelligence for medical purposes.

Nevertheless, although Hippocrates is showing outstanding results using medical images,<sup>[25]</sup> some challenges remain in the AI democratization.

First, strictly speaking, Hippocrates is susceptible to model bias. As Doctors do not fully understand AI bases, the data used to train Hippocrates might be biased to a specific population. However, this problem is not only affecting Hippocrates. Hundreds of reports show that this is a significant challenge in applying AI in health since most available data correspond to white males<sup>[26]</sup>. Nonetheless, at ArkangelAI, we genuinely believe that creating platforms like Hippocrates the creation of introductory courses in artificial intelligence in health sciences curriculums is a matter of time. And with healthcare students learning the bases of AI, they will start to create much more complete and diverse datasets putting an end to this alarming challenge. Meanwhile, Hippocrates is fighting bias by encouraging users to perform prospective studies seeking to evaluate the generalizability of the algorithms.

Second, AI models created using Hippocrates do not provide by themselves the explanation required for the treating doctor to understand the results, which is one of the main stoppers in the implementation<sup>[27]</sup>. Notwithstanding, the deployment aspect of Hippocrates is being upgraded. The upgrade will allow users to observe in the service explainable layers to the algorithm so that heatmaps, bounding boxes, segmentation, and other observable characteristics will be presented to the users to be used as a new asset in the decision-making process. With this new deployment tool, ArkangelAI aims to showcase the benefits of implementing AI algorithms as helpers for daily healthcare activities, boosting AI solutions' implementation in real life.

Last but not least, even though progress has been made in the definition of legal frameworks for the use of algorithms in medical settings<sup>[28]</sup>, it is necessary to continue working on creating a precise regulation for the development and deployment of AI solutions for medical purposes. Fortunately, the recurrent use of platforms like Hippocrates will drive the emergence of innovative solutions that will put healthcare professionals at the forefront in defining the legal framework of the use of algorithms created by themselves.

Finally, although the future of Artificial Intelligence is uncertain, democratizing the use of AI for medical purposes using Hippocrates AutoML, promises to change the healthcare industry as we know it, putting health workers at the forefront of innovation and development using AI.

## SOURCES

- [1] Dick, S.: Artificial Intelligence. Harvard Data Science Review, 1(1), (2019). <https://doi.org/10.1162/99608f92.92fe150c>
- [2] European Commission.: WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust. [https://wayback.archive-it.org/12090/20200811222622/https://ec.europa.eu/info/sites/info/files/comm-ission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://wayback.archive-it.org/12090/20200811222622/https://ec.europa.eu/info/sites/info/files/comm-ission-white-paper-artificial-intelligence-feb2020_en.pdf). Last Accessed 2021/09/10.
- [3] Panch, T., Mattie, H. & Celi, L.A.: The “inconvenient truth” about AI in healthcare. npj Digit. Med. 2, 77 (2019). <https://doi.org/10.1038/s41746-019-0155-4>
- [4] Rao, A., Palaci, F. & Chowi, W.: A practical guide to Responsible Artificial Intelligence (AI). <https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.pdf>. Last Accessed 2021/09/10
- [5] Ahmed, N., Wahed, M.: The de-democratization of ai: Deep learning and the compute divide in artificial intelligence research.arXiv preprint (2020). arXiv:2010.15581
- [6]Ahmad, R., Rashid, K., David, P.: The social democratization of knowledge: Some critical reflections on e-learning. Multicultural Education & Technology Journal. 1. 64-74 (2007). [10.1108/17504970710745210](https://doi.org/10.1108/17504970710745210).
- [7] Riedl, M.: AI Democratization in the Era of GPT-3. The gradient 2020. <https://thegradient.pub/ai-democratization-in-the-era-of-gpt-3/>. Last Accessed 2021/09/10
- [8]Weissglass, D. E.: Contextual bias, the democratization of healthcare, and medical artificial intelligence in low- and middle-income countries. Bioethics, 1– 9, (2021). <https://doi.org/10.1111.bioe.12927>
- [9] Price II, W. N.: Medical AI and Contextual Bias. 33 Harv. J. L. & Tech. 66 (2019), U of Michigan Public Law Research Paper No. 632, Available at SSRN: <https://ssrn.com/abstract=3347890>
- [10] Chockley, K, Ezekiel, E.: The End of Radiology? Three Threats to the Future Practice of Radiology. Journal of the American College of Radiology, 13, 12 A, 1415-1420, (2016). <https://doi.org/10.1016/j.jacr.2016.07.010>
- [11] Haag, M et al.: Digital Teaching and Digital Medicine: A national initiative is needed. GMS journal for medical education vol. 35,3 Doc43, 15(2018), doi:10.3205/zma001189
- [12]Topol, E.J.: High-performance medicine: the convergence of human and artificial intelligence. Nat Med 25, 44–56 (2019). <https://doi.org/10.1038/s41591-018-0300-7>
- [13]Vergheze A, Shah NH, Harrington RA. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. JAMA. 2018;319(1):19–20. doi:10.1001/jama.2017.19198
- [14] Shah, N.: Health Care in 2030: Will Artificial Intelligence Replace Physicians?. American College of Physicians Journal,(2019). <https://doi.org/10.7326/M19-0344>
- [15]Price WN 2nd, Gerke S, Cohen IG.: Potential Liability for Physicians Using Artificial Intelligence. JAMA, Nov 12;322(18):1765-1766, ( 2019), doi: 10.1001/jama.2019.15064. PMID: 31584609.
- [16]Rao, A.: Democratization of AI A double-edged sword. Towards Data Science. <https://towardsdatascience.com/democratization-of-ai-de-155f0616b5>. Last Accessed 2021/09/13
- [17] HICX.: What Are The Pros And Cons Of Data Democratisation?. HICX blogs. <https://www.hicx.com/blog/data-democratisation-what-are-the-pros-and-cons/>
- [18] Gonzalez, A., Munera, N., Zea, J.: AI in Medical Imaging: Chest X-Ray case study. <https://app.hubspot.com/documents/8676854/view/234345503?accessId=df46ac>
- [19]He, X., Zhao, K., Chu, X.:AutoML: A survey of the state-of-the-art. Knowledge-Based Systems, Volume 212 (2021). <https://doi.org/10.1016/j.knosys.2020.106622>.
- [20] Google Cloud Platform, Cumplimiento de la HIPAA en Google Cloud Platform. <https://cloud.google.com/security/compliance/hipaa>
- [21]Amazon Web Services. Informacion general-HIPAA. <https://aws.amazon.com/es/compliance/hipaa-compliance/>
- [22]Hussain, Z., Gimenez, F., Y, D., Rubin, D.: Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. AMIA Annu Symp Proc. 2018 Apr 16;2017:979-984. PMID: 29854165; PMCID: PMC5977656.
- [23] Draelos, R.: Best Use of Train/Val/Test Splits, with Tips for Medical Data. <https://glassboxmedicine.com/2019/09/15/best-use-of-train-val-test-splits-with-tips-for-medical-data/>
- [24]Kelly, C.J., Karthikesalingam, A., Suleyman, M. et al.: Key challenges for delivering clinical impact with artificial intelligence. BMC Med 17, 195 (2019). <https://doi.org/10.1186/s12916-019-1426-2>
- [25] Gonzalez, A., Munera, N., Zea, J.: How to build an open-source AI tool for early detection of Childhood Leukaemia without code. <https://blog.arkangel.ai/leukemia/>
- [26] Cirillo, D., Catuara-Solarz, S., Morey, C. et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. npj Digit. Med. 3, 81 (2020). <https://doi.org/10.1038/s41746-020-0288-5>
- [27]Aniek F. Markus, Jan A. Kors, Peter R. Rijnbeek.:The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. Journal of Biomedical Informatics,Volume 113(2021),103655, ISSN 1532-0464, <https://doi.org/10.1016/j.jbi.2020.103655>.

## About Arkangel AI:



Arkangel AI is a software company based in Montreal, Canada that specializes in early disease detection using artificial intelligence. The company's mission is to enable people to live free of preventable diseases through early disease detection. Arkangel AI's products are optimized for diseases from the global south and medical equipment available in the region, translating into fewer entry barriers to urban and rural settings. Arkangel AI has operations in Canada, Colombia, Uruguay, and Mexico. For further information on this research or strategic partnerships:

### **Arkangel AI**

1250 Guy St Suite #600, Montreal, Quebec H3H 2L3

[hola@arkangel.ai](mailto:hola@arkangel.ai)

[www.arkangel.ai](http://www.arkangel.ai)

### **Contributors:**

#### **Álvaro Jose Gonzalez Grajales**

Biomedical Data Scientist

Contact: [alvaro.gonzalez@arkangel.ai](mailto:alvaro.gonzalez@arkangel.ai)

#### **Nicolas Munera**

Machine Learning Engineer - CTO

Contact: [nicolas.munera@arkangel.ai](mailto:nicolas.munera@arkangel.ai)

#### **Jose Zea**

CEO and Founder

Contact: [jose@arkangel.ai](mailto:jose@arkangel.ai)