



State of AI Survey

AI Model Development by Industry,
GPU Adoption, and More

2021

Summary —

and Contents

In February of 2021, WEKA did a survey on the state of AI and Analytics. The study comprised 534 individuals globally across a wide range of industries and job functions related to IT, data science, and management. This update structures the information from the original survey into a shorter format that highlights key findings and the implications on the AI and analytics market, specifically around Use cases by market, Data source usage, AI Modeling software, and GPU Vs. CPU adoption rates.

Contents

01 Key Findings	3
02 Artificial Intelligence Use Cases by Market Vertical	4
03 Data Sources for AI Workloads	5
04 AI Model Development by Market	6
05 GPU Adoption In AI Workloads	7
06 Conclusion	8

Key Findings —

We asked the 534 participants in the survey questions about their IT architectures, frameworks, databases, spending, challenges, and more.



01 86% of respondents **have at least one AI initiative**. Most companies have 2-3 initiatives, while a few had as many as 5 discrete initiatives



02 **Use cases vary significantly by vertical**, but the most cited initiatives were recommender engines, scientific visualization, and image recognition.



03 The **approaches to AI vary**, but research and government tend to build their own models that are custom and specific to their needs, while commercial enterprises tend to use ready-to-consume models to gain a time-to-market advantage.



04 In general, **50% of the data is self-generated**, but other external sources contribute



05 While no respondents stated that they plan to be cloud only, **most expect a significant and increasing proportion of their workloads to be in the cloud**. Privacy concerns and the complexity of extending workloads to the cloud are the main headwinds toward additional cloud adoption.



06 While it was no surprise to see that **AWS dominates the market**, some industries favored other cloud providers.



07 Over **half of the respondents are already using GPUs** and adoption of GPUs is especially high in manufacturing and automotive segments.



08 The **largest headwind to the adoption of AI initiatives is the lack of data scientists**, followed closely by lack of budget and infrastructure to build the AI environment.

AI Uses —

By Vertical Market

Let's drill down to the specific use cases by industry. Depending on your specific industry, the use cases vary. For example compliance is #1 in Healthcare and Banking but not a relevant use case in Education, Research, and less regulated industries. Compliance

in heavily regulated industries is a multi-million dollar opportunity given the brand impact of fraud activity that is undetected (consider the case of the Target breach in December 2013) and the potential fines by regulators. (Target was fined \$18.5M.)

Artificial Intelligence Uses Cases By Vertical

Reported 2021

 Healthcare	 Education	 Banking & Finance	 Government & Research
01 Compliance 13%	01 Scientific Visualization 19%	01 Compliance 17%	01 Scientific Visualization 24%
02 Image Recognition 13%	02 Image Recognition 13%	02 Recommender systems 17%	02 Image Recognition 16%
03 Scientific Visualization 13%	03 Computational Chemistry 11%	03 Cyber Security 13%	03 Computational Chemistry 14%
04 Computational Chemistry 12%	04 Conversational AI 11%	04 Conversational AI 13%	04 Recommender systems 10%
05 Cyber Security 11%	05 Recommender systems 9%	05 Sentiment Analysis 13%	05 Cyber Security 9%

 Media & Entertainment	 Telecom	 Cloud & MSP
01 Sales / Marketing optimization 18%	01 Recommender systems 16%	01 Conversational AI 17%
02 Conversational AI 12%	02 Compliance 13%	02 Recommender systems 17%
03 Recommender systems 10%	03 Cyber Security 13%	03 Compliance 13%

KEY INSIGHT

“If you are not investing in at least one of these initiatives, you are most likely losing the competitive edge that others are gaining.”

Image recognition is very common in Research, given that Research tends to be dominated by High-Performance Computing (HPC) environments. In recent years we've actually seen the convergence of artificial intelligence and HPC given the vast amounts of data, mainly images, that research institutions

tend to store. Research institutions use AI to run multiple iterations to test models and to gain insights. The adoption of AI in these institutions is growing significantly.

The top use cases in each vertical are those that give companies a competitive edge.

Data Sources —

For AI Workloads

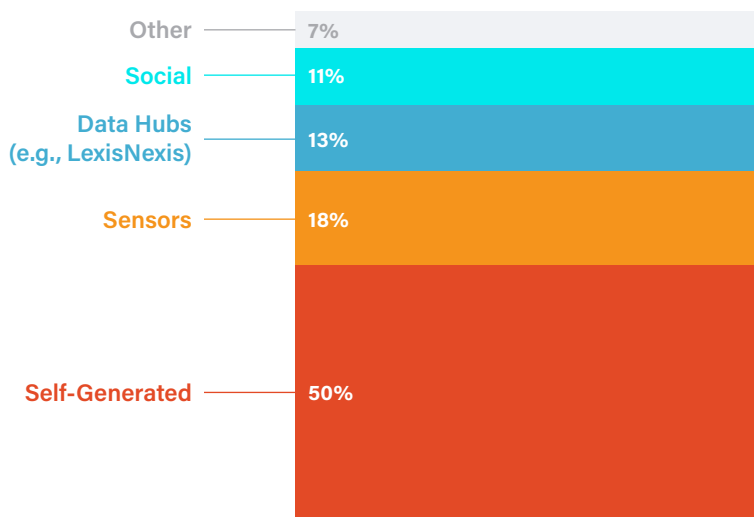
Across all market verticals, aggregated data revealed a number of interesting things.

We were somewhat surprised that 50% of the data is self-generated including synthesized data from other data sets. On the other hand, there is little surprise that 18% of the data is sourced from sensors given the rise of IoT and the use of sensors in legacy manufacturing operations. Over time, we expect that this number will increase even more given the adoption of consumer IoT data into commercial datasets.

As we dig into the data a bit more, we find that healthcare and Education depended significantly on self-generated data (53% each), while Energy / Oil & Gas and Semiconductor areas were highly reliant on data sourced from sensors used for monitoring (32% and 56% respectively).

Data Sources for AI Workloads

Reported 2021



KEY INSIGHT

Know **in advance** where your data will come from, the data format, and plan for how you will transform it. Different data sources have different requirements to manage them.



Data coming from external data sources usually needs to follow a Extract, Transform, Load (ETL) process to optimize it for analytics.



Self-generated data requires cleaning and tagging, which should be a planned investment for the project.

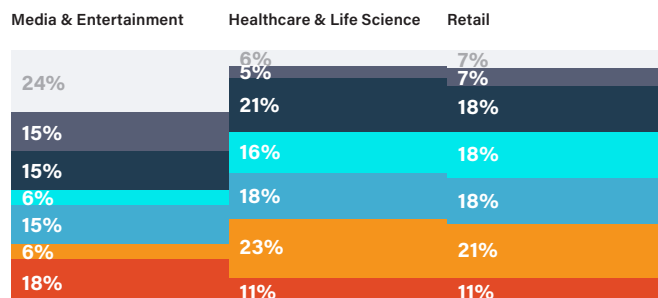
AI Model Development —

In AI Workloads

86% of responders are already adopting Artificial Intelligence and Machine Learning.

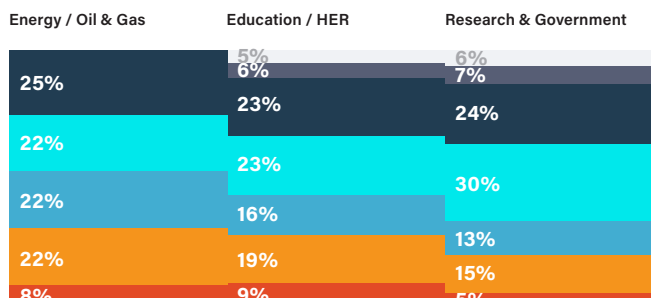
Traditional Enterprises

The approach to developing AI models changes depending on the vertical. Overall, we can identify two approaches, one for traditional enterprises, and one for research and education. Traditional enterprises have a tendency to buy pre-built models. Their focus is on the product and deliverables.



Research Institutions & Education

Research Institutions and Education focus on building their own solutions and are less likely to buy pre-built models. They invest more data science resources into solving the problems. The outcome is that the models are more relevant to the specific needs of the organization. Oil and gas is a highly specialized market and tends to reflect more of the HPC environment than enterprise.



LEGEND

- We have an AI-as-a service platform
- We use high-level data science modeling tools
- We use automatic machine learning tools
- We work directly in AI frameworks like TensorFlow or PyTorch
- We work directly in Python using Pandas or Scikit-learn type libraries
- Not applicable / none of the above
- We are planning to deploy AI in the next 6-12 months

KEY INSIGHT

Model development will impact the data pipeline. The more custom the model and workflow, the more the supporting infrastructure will need to be attuned to the specific workload for the best outcomes.

GPU Adoption —

In AI Workloads

Half of respondents already have more than 10 GPU/accelerators, with over a third of respondents using 20 or more GPUs.

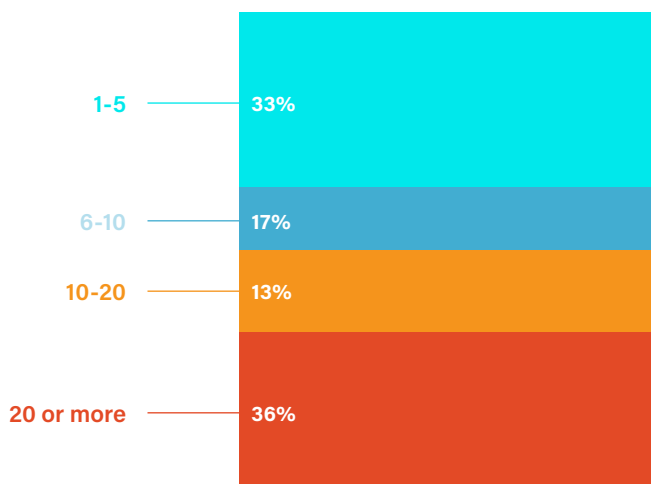
From a drill down standpoint, various industries have a much higher GPU usage rate than others. Energy / Oil & Gas, Manufacturing, and Government research institutions are leaders in this space with over 50% of the respondents indicating 20+ GPUs deployed, while Media and Entertainment, Banking (non-algorithmic usage), and Retail are significantly lower with less than 20% having 20+ GPUs on average.

Within the survey, over half of the respondents (52%) mention using GPUs in production or pilot programs. 38% mentioned that they do not have immediate plans to deploy GPUs. Most of these respondents came from the Middle East. Adoption of GPUs is especially high in Automotive, where image recognition is used to power autonomous transportation. Only 14% of Automotive respondents do not have plans to use GPUs. Within the next 6-12 months, the verticals most likely to grow their usage of GPUs are Automotive and Manufacturing, most likely for preventive maintenance and quality assurance purposes. Respondents in Telecom indicate that they are the least likely to adopt GPUs, with over 60% responding against growth in that area.

Barriers to adoption of AI and GPU infrastructures vary significantly across all verticals. Of the respondents, there was no dominant factor on its own. One very revealing number is very low lack of support from executive teams. This indicates a strong executive level desire to take advantage of AI and GPU to accelerate the business.

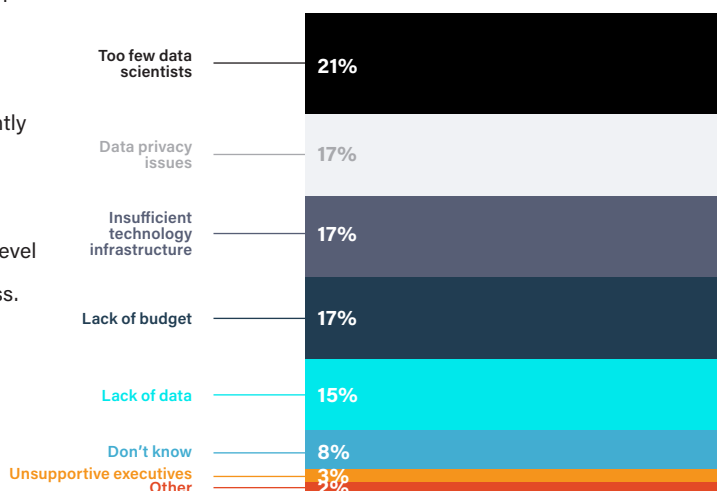
GPU Adoption in AI Workloads

Reported 2021



Barriers to Adoption

Reported 2021



Conclusion —

Historically, enterprises across all verticals have relied on old business models to support their operations.

In the era of high-performance GPUs that can be brought to bear against larger analytical datasets, the promise of AI for recommender systems, compliance, and image recognition and computer vision applications to help with decision support has become real. As the amount of data coming to feed these new analytical models and AI gets larger and larger, there will be an acceleration of GPU adoption as well as high speed networking, and storage-as-a-data-platform to help feed these new GPU's performance levels. WEKA intends to re-do the full survey on a regular basis to provide comparative insights on market trends.

WEKA is recognized as a leader in providing a Data platform for AI applications.

*As a visionary in the Gartner Magic Quadrant for Distributed File Systems and Object Storage, we've taken the best technology for accelerating your data pipeline for GPU based workloads and made it a reality. **By providing the highest performance levels and common data access across all protocols as well as data-lake levels of capacity scaling, Weka is uniquely positioned to help make your AI and analytics projects successful.***

To find out more, please contact us at info@weka.io.



910 E Hamilton Avenue, Suite 430, Campbell, CA 95008
408.335.0085 | info@weka.io

© 2019-2022 All rights reserved. WekaIO, WekaFS, WIN, Weka Innovation Network, the Weka brand mark, the Weka logo, and Radically Simple Storage are trademarks of WekaIO, Inc. and its affiliates in the United States and/or other countries. Other trademarks are the property of their respective companies. References in this publication to WekaIO's products, programs, or services do not imply that WekaIO intends to make these available in all countries in which it operates. Product specifications provided are sample specifications and do not constitute a warranty. Information is true as of the date of publication and is subject to change. Actual specifications for unique part numbers may vary.