Large Language Model as an Interpreter Interfacing the Agents and the World*

John Sum
Institute of Technology Management
National Chung Hsing University
Taichung 402, Taiwan.
(pfsum@nchu.edu.tw)

November 20, 2026

Abstract

In this article, the roles of the large language models in the area of agentic AI are discussed. With the ability to generate program codes, an large language model (LLM) is able to generate an *agent* which is a program to be executable. These agents could thus be incorporated in a business process of a firm. Moreover, worker-LLM-agent can be developed for a worker to act on behalf to interact with other LLM-agents in the process. As a result, a new phase of business process reengineering might be needed for this agentic AI era. An example on replenishment process and its agent-based redesign is elucidated. Subsequently, the idea of *global virtual computer* (GVC) is introduced and the potential GVCs are listed. Finally, the roles of the LLMs as an interface and an interpreter are re-advocated. Moreover, the idea of cyber-physical systems is re-advocated with LLM as a major driving force. It is believed that the ideas presented in this paper are not new. LLM developers should have already been working along this direction. On one hand, the LLM developers excel the capability of an LLM with powerful reasoning ability. On the other hand, the developers excel the capabilities of the text-to-xyz foundational models, especially the text-to-code generator.

Contents

1 Roles of an LLM **1**
1.1 Information Infrastructure 1
1.2 Foundation Models 1
1.3 Large Language Models 1
1.4 Application Systems (Agents) 3

2 Support Business Processes **3**
2.1 Information/Physical Infrastructure 3
2.2 Interactions Among Layers 3
2.3 Human in a Business Process/Application System 4
2.4 Business Process Automation/Reengineering 4
2.5 Example : Replenishment Process Reengineering 4
2.5.1 Agent-Based Re-Design 4
2.5.2 Multi-Agent embraced in an Agent 7

*This manuscript is prepared by the author as a teaching material for the NCHU MBA/EMBA students who attend the courses on *AIML*, *Intelligent Technology* or *Intelligent Technology Management*.

2.5.3 Performance Improvement (*) . . . 7
2.6 Agents' Interaction Protocol . . . 7
2.6.1 2000s Protocols . . . 7
2.6.2 2020s Protocols . . . 7

3 Go Beyond (Might Already Happen) 9
3.1 LLM-as-a-Service (Agent) . . . 9
3.2 LLM-Agent as a Digital Twin . . . 9
3.3 Interfacing Global Virtual Computer . . . 9
3.4 Personal Agents . . . 10
3.5 Global Virtual Computers (GVCs) . . . 10
3.5.1 Gemini, Co-Pilot and Grok . . . 10
3.5.2 Claude and Llama . . . 10
3.5.3 Qwen, Pangu and Hunyuan . . . 12
3.6 LLM for Advertisers . . . 12
3.7 Interfacing the World . . . 12
3.7.1 Interfacing the Digital World . . . 12
3.7.2 Interfacing the Physical World . . . 12

4 Privacy & Security – My Trust 14
4.1 Privacy & Security Issue . . . 14
4.2 Monetary Issue . . . 14
4.3 Responsibility of the AI Tech Firms . . . 14

5 Reclaim the Roles/Limitations of an LLM 14
5.1 Human-Agent (Human-GVC) Interactions . . . 15
5.2 Improve Working Efficiency . . . 15
5.3 Functionally Limited . . . 15
5.4 Reasoning Agent or Interpreter . . . 15
5.4.1 Imperfect Reasoner/Interpreter . . . 15
5.4.2 Human Reasoning is a Key . . . 16
5.5 Concrete Operational Machine . . . 16

6 The Roles of Google Gemini to JS 16
6.1 Interpreter . . . 16
6.2 Virtual Master of Learning . . . 16

List of Tables

1 List of potential global virtual computers and their display networks. 11

List of Figures

1 The stack of LLM and agents. LLM plays the vital role in business process automation and re-engineering. Here, T2T stands for text-to-text. T2I stands for text-to-image. T2V stands for text-to-video. T2A stands for text-to-audio. T2C stands for text-to-code. The text is the prompt entered by the worker or an *autonomous* application system. 1
2 Expansion of the information infrastructure layer in the Figure 1. This layer consists of software components like platform API, operating system and device driver; and hardware components like CPU, GPU and network communication device. It should be noted that the CPU, the operating system, the network communication device and the network operating system could be specialized designed for AI applications deployment. The application systems consist of applications supporting general business administration like ERP and LLM-specific applications. For the general business administration applications, the cloud on the left is the conventional information infrastructure to support them. The information infrastructure on the right is a specialized infrastructure particularly installed to support LLM-powered applications. The architecture shown in this diagram can be treated as a reference model for the AI/IS/IT industrial structure. 2
3 The ultimate purpose of the layers of technologies is to support the business processes of a firm. The physical infrastructure embraces the physical facilities, machines and human resources to support the logistic flow of physical goods. This physical infrastructure is not included in Figure 2. 3
4 The *interaction* and *sequence* diagrams of a replenishment process. It involves both information flow and physical material flow. AI tools and information infrastructure could speed up the information flow, while the physical material flows in the production and delivery processes cannot. So, one should be aware of the so-called performance improvement. 5
5 List of the steps for the process design as shown in Figure 4. 6
6 An agent-based design for the interactions among actors in a replenishment process. Here, the logistic center information system and the factory information system are both treated as agents. A worker agent (resp. 3PL worker agent) is built for the worker (resp. 3PL worker). The workers in the factory do not have any LLM-agent to act on their behalf. This system is a cyber-physical system. The agents are the *digital twins* of the physical workers and the factory for information exchange and authorized decision making. 8
7 The LLMs and the underlying technologies could support a user to (i) access any global information/physical resource and (ii) complete a task on his/her behalf. . . . 9
8 LLM agents are logically connected to form a network of personal agents 10
9 An expansion of the layered structure as shown in Figure 7. The LLMs and the underlying technologies could support a user to (i) access any global information/physical resource and (ii) complete a task on his/her behalf. The information resources are corresponding to those extracted from and added to the digital world. These interactions are accomplished by the information & communication technologies. To interact with the physical resources in the physical world, special equipments like (computer controlled) robots and machines are needed. 13

10 Basically, an LLM-agent could be composed of two parts. One part is the LLM-as-an-Interpreter which interfaces between the human user and the human-design agent (a software). The other part is the human-design agent. It is a software responsible for handling routine works. 15

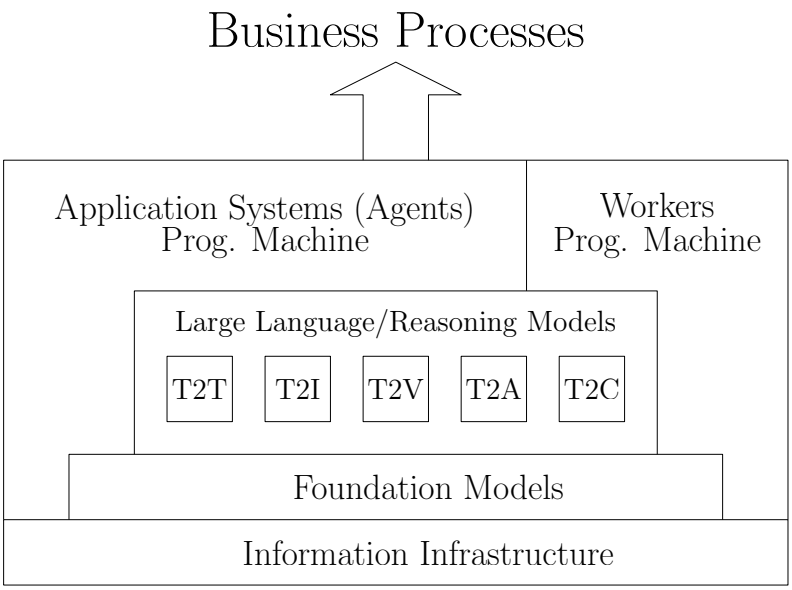


Figure 1: The stack of LLM and agents. LLM plays the vital role in business process automation and re-engineering. Here, T2T stands for text-to-text. T2I stands for text-to-image. T2V stands for text-to-video. T2A stands for text-to-audio. T2C stands for text-to-code. The text is the prompt entered by the worker or an *autonomous* application system.

1 Roles of an LLM

Due to the advancement of large language models, like ChatGPT and Gemini, and their associated foundation models, various application systems could be easily developed in the form of agents to facilitate business process automation. Figure 1 and Figure 2 show the stack of these modules.

1.1 Information Infrastructure

The layer *information infrastructure* includes the Internet, the telecommunication networks, the database systems, cloud platforms, computers and others. This layer provides services to support the upper layers in terms of data/information storage, data/information processing and numerical computation. As shown in Figure 2, this information infrastructure consists of a cloud platform supporting the general business administration applications and an AI-oriented platform supporting LLM-powered applications.

1.2 Foundation Models

The layer *foundation models* covers a number of generative AI models, including LLMs, LRMs, text-to-image generators, text-to-video generators, text-to-audio generators, translators. These foundation models provide services mainly to support the *LLM layer* and the agents.

1.3 Large Language Models

Each LLM in the layer *large language models* acts as an interface[1] in between the application systems (resp. workers) and the foundation models (resp. information infrastructure). Workers

[1] An LLM acts as a programming language interpreter. For a program its commands are coded by an interpreted language, the interpreter converts each command to a machine code sub-routine and then executes the sub-routine right the way. Afterward, the interpreter converts and executes the next command. The process repeats until no more command left.

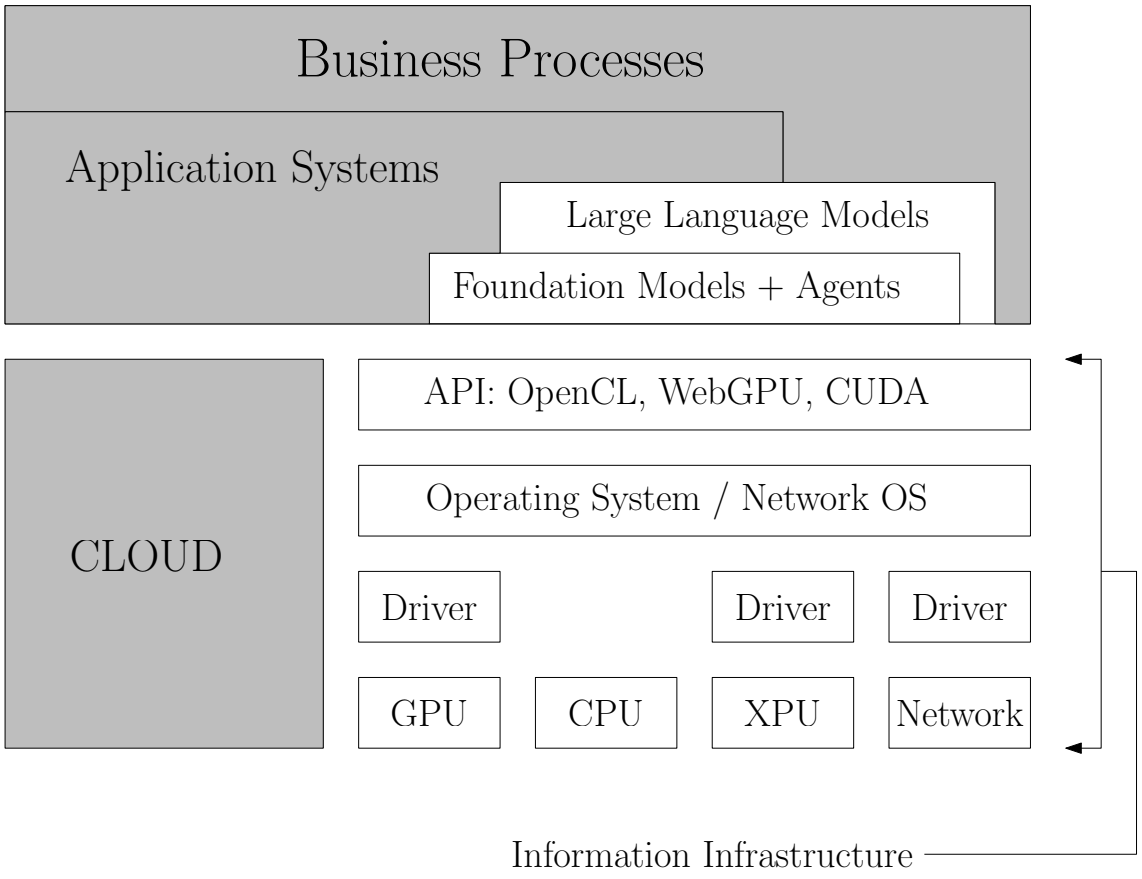


Figure 2: Expansion of the information infrastructure layer in the Figure 1. This layer consists of software components like platform API, operating system and device driver; and hardware components like CPU, GPU and network communication device. It should be noted that the CPU, the operating system, the network communication device and the network operating system could be specialized designed for AI applications deployment. The application systems consist of applications supporting general business administration like ERP and LLM-specific applications. For the general business administration applications, the cloud on the left is the conventional information infrastructure to support them. The information infrastructure on the right is a specialized infrastructure particularly installed to support LLM-powered applications. The architecture shown in this diagram can be treated as a reference model for the AI/IS/IT industrial structure.

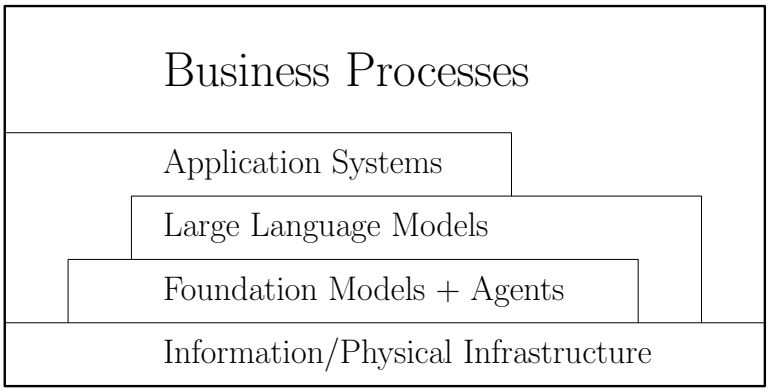


Figure 3: The ultimate purpose of the layers of technologies is to support the business processes of a firm. The physical infrastructure embraces the physical facilities, machines and human resources to support the logistic flow of physical goods. This physical infrastructure is not included in Figure 2.

benefit the most from the LLM as each LLM is able to understand human language. To instruct an LLM to complete a simple task, a worker can simply type the instructions in natural language and then the LLM will do it accordingly. A worker does not have to learn any new programming language.

1.4 Application Systems (Agents)

The *application systems* are responsible for completion of those jobs which are pre-assigned. In simple words, each application system is an agent. It acts on behalf of a worker to complete a job which is used to be completed manually.

2 Support Business Processes

The ultimate purpose of these layers is to support the business processes of a firm, see Figure 3. In it, the *information/physical infrastructure* is an extension of the *information infrastructure* as shown in Figure 1.

2.1 Information/Physical Infrastructure

The *information/physical infrastructure* covers not just the infrastructure for the information, but also the infrastructure for the physical facilities and human resources. Thus, the *information/physical infrastructure* manages (i) all the resources (including but not limited to information resource, financial resource and human resource) within a firm and (ii) the resources available for public (over or not over the Internet).

2.2 Interactions Among Layers

In Figure 3, a business process is accomplished by a number of human workers together with the underlying application systems. A human worker in a business process might need to directly interact with (i) an LLM, (ii) the foundation models, (iii) the agents or (iv) the *information/physical infrastructure* to complete a task.

A human worker could simply interact with an LLM to complete the task. But, human worker might also interact with the resources in the *information/physical infrastructure* to complete the task. For an application system, it could interact with an LLM to complete its task. Besides, it could interact with the *foundation models + agents* layer and/or the *information/physical infrastructure*

layer to complete a task. Clearly, a smart human worker could skip the LLMs layer and directly interact with the *foundation models + agents* layer and/or the *information/physical infrastructure* layer to complete a task. However, this part is not shown in Figure 3 due to the 2D diagram constraint.

2.3 Human in a Business Process/Application System

Here, we consider that human worker is a part of a business process and/or an application system. In convention, an application system is normally developed entirely by accessing the foundation models and the pre-developed systems in the *information infrastructure* layer without needed to access any LLM. With the LLMs, a worker in a business process and/or an application system can develop his/her personalized application systems which entirely rely on top of the LLMs but not the underlying *foundation models* layer or *information/physical infrastructure* layer.

2.4 Business Process Automation/Reengineering

Recent successes can be witnessed from the incorporating of the LLMs in *knowledge management processes*, *marketing processes*[2] and *adminstration processes* in a firm, in which the processes require a lot of (i) information collection and processing; (ii) Ads creative generation and (iii) document generation. The LLMs automates these tasks without changing the design of the existing the *knowledge management processes*, *marketing processes* and *adminstration processes*.

To succeed a radical improvement, business processes might have to be co-designed with the supporting application systems. That is to say, business process re-engineering[3] might be needed.

2.5 Example : Replenishment Process Reengineering

A possible design for the replenishment process is shown in Figure 4. In this design, a number of assumptions have to be made.

- The replenishment process is initiated by the logistic center information system (LCIS). A program in the LCIS has been scheduled to be running at 22:00 each day.
- A worker in the EC logistic center will log in the LCIS every working day for the list of items with low stock level.
- A worker in the factory will log in the factory information system every working day for the list of replenished items and the list of qualified items being delivered to the EC firm.

Figure 5 lists the detail steps to be accomplished in the process design as shown in Figure 4.

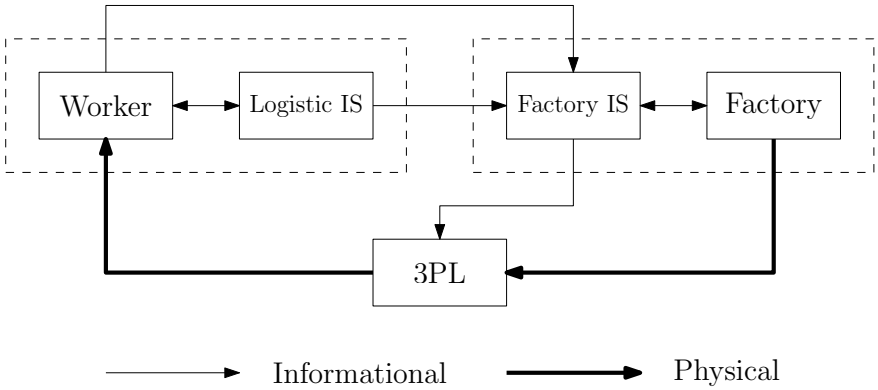
2.5.1 Agent-Based Re-Design

Introducing personal LLM-agents, the LLM-agents are now actors in a replenishment process. Thus, the interactions among the actors in a replenishment process will have to be re-designed. An interaction diagram of an agent-based design is shown in Figure 6. For the sequence diagram for this new design, we leave it for the reader to do so. One should be noted that the process in this interaction diagram is in essence a *cyber-physical system*[4] as the process involves human workers and agents.

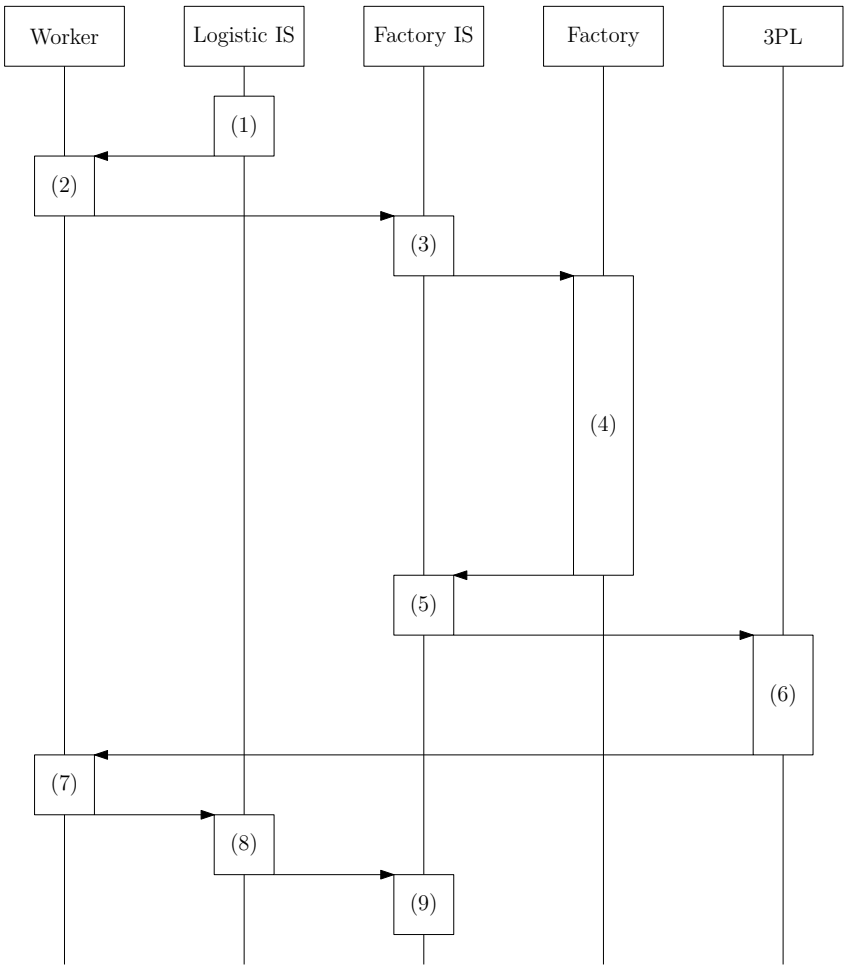
[2]It is especially on advertising creative generations, including text content generation, image generation and short video generation.

[3]Hammer, Micheal (1990). Reengineering work: Don't automate, obliterate, *Harvard business review*, 68(4), 104-112. Wikipedia: `https://en.wikipedia.org/wiki/Business_process_re-engineering`.

[4]W. Wolf, Cyber-Physical systems, *Computer*, 42(03), pp.88-89, 2009. R. Baheti and G. Helen, Cyber-Physical systems, *The impact of control technology*, Vol. 12, no. 1, 161-166, 2011.



(a) Interaction diagram.



(b) Sequence diagram.

Figure 4: The *interaction* and *sequence* diagrams of a replenishment process. It involves both information flow and physical material flow. AI tools and information infrastructure could speed up the information flow, while the physical material flows in the production and delivery processes cannot. So, one should be aware of the so-called performance improvement.

```
------------------------------------------------------------
Replenishment Process
------------------------------------------------------------
Step 1: 1.1: LCIS checks from the DBMS the stock level
             of each item for sale.
        1.2: LCIS compiles a list of items which stock
             levels are below their pre-set levels.
Step 2: 2.1: A worker check the list of items with low
             stock levels.
        2.2: A worker decides a list of items to be
             replenished.
        2.3: The list of items is sent to the FIS.
Step 3: 3.1: The FIS updates the request in the DBMS.
        3.2: The FIS informs the factory for production.
Step 4: 4.1: Factory produces the requested replenishment.
        4.2: A worker in the factory updates the
             production status to the FIS.
Step 5: The FIS informs a 3PL for delivery.
Step 6: 6.1: The 3PL arrives the factory and collects
             the items.
        6.2: The 3PL delivers the items to the logistic
             center of the EC firm.
Step 7: 7.1: A worker in the EC logistic center checks
             the items for defective.
        7.2: A worker puts the item to its shelf.
Step 8: 8.1: If an item is not defective, its bar code
             will be scanned and its information is thus
             entered in the LCIS.
        8.2: The LCIS informs the FIS for those qualified
             items.
Step 9: The FIS updates its DBMS for those qualified items.
------------------------------------------------------------
```

Figure 5: List of the steps for the process design as shown in Figure 4.

In this new design, the logistic center information system and the factory information system are both treated as agents. A worker LLM-agent (resp. 3PL worker LLM-agent) is built to act on behalf the worker (resp. 3PL worker) to interact with other agents on information exchange.

2.5.2 Multi-Agent embraced in an Agent

In this example, the items to be replenished are belong to a single product. Imagine that there are N products. Both the worker and the 3PL worker will be involved in N replenishment processes. N worker-LLM-agents (resp. 3PL worker-LLM-agents) will have to be built for the worker (resp. 3PL worker) to automate his/her interaction with other agents. To this end, these N LLM-agents could be composed as a single LLM-agent to remind the worker (resp. 3PL worker) to take action whenever a physical step is in progress.

2.5.3 Performance Improvement (*)

Finally, one should be aware of the so-called performance improvement. From Figure 6, it is clear that the time spent on production and product delivery are the bottlenecks of the overall replenishment process. Even if the time spent on information flow reduces to zero, the time spent on physical product production and delivery could still be significant.

Let say, the total time spent on information flow in the old design is 48 hours. The time spent on production is 7 days. The time spent on delivery is 8 hours. So, the completion time of a replenishment process is 9 days and 8 hours. If the total time spent on information flow reduces to 1 hour, the completion time of a replenishment process is 7 days and 9 hours.

If we only consider the performance improvement on information flow, the improvement is close to 98 percent. However, the improvement on the completion time is just 21 percent. Therefore, one should be aware of the actual definition of the performance improvement being claimed for each tool.

2.6 Agents' Interaction Protocol

It is clear that the interactions among the LLM-agents and the agents are accomplished by the messages communicated. Protocol has thus to be defined for these message communications.

2.6.1 2000s Protocols

It should be noted that mechanisms (i.e. protocols) for agent-communication can be traced back to the 2000s[5] under the topic named *autonomous agents*. Difference from today AI agent, those agents developed in those days had not focused on AI agents.

2.6.2 2020s Protocols

From that, one could infer that designing of protocols for agents communication should not be difficult. Thus, many protocols (resp. protocols under approval) should have bee developed. As a matter of fact, Anthropic and Google have been developing the protocols for AI-agents communications, namely *model context protocol*[6] and *agent-to-agent protocol*[7].

[5] Jiming Liu *et al* (eds), *Agent Engineering*, Series in Machine Perception and Artificial Intelligence: Volume 43, 2001. World Scientific. Jiming Liu, *Autonomous agents and multi-agent systems: explorations in learning, self-organization and adaptive computation*, World Scientific, 2001.

[6] `https://modelcontextprotocol.io/docs/getting-started/intro`

[7] `https://a2a-protocol.org/latest/`

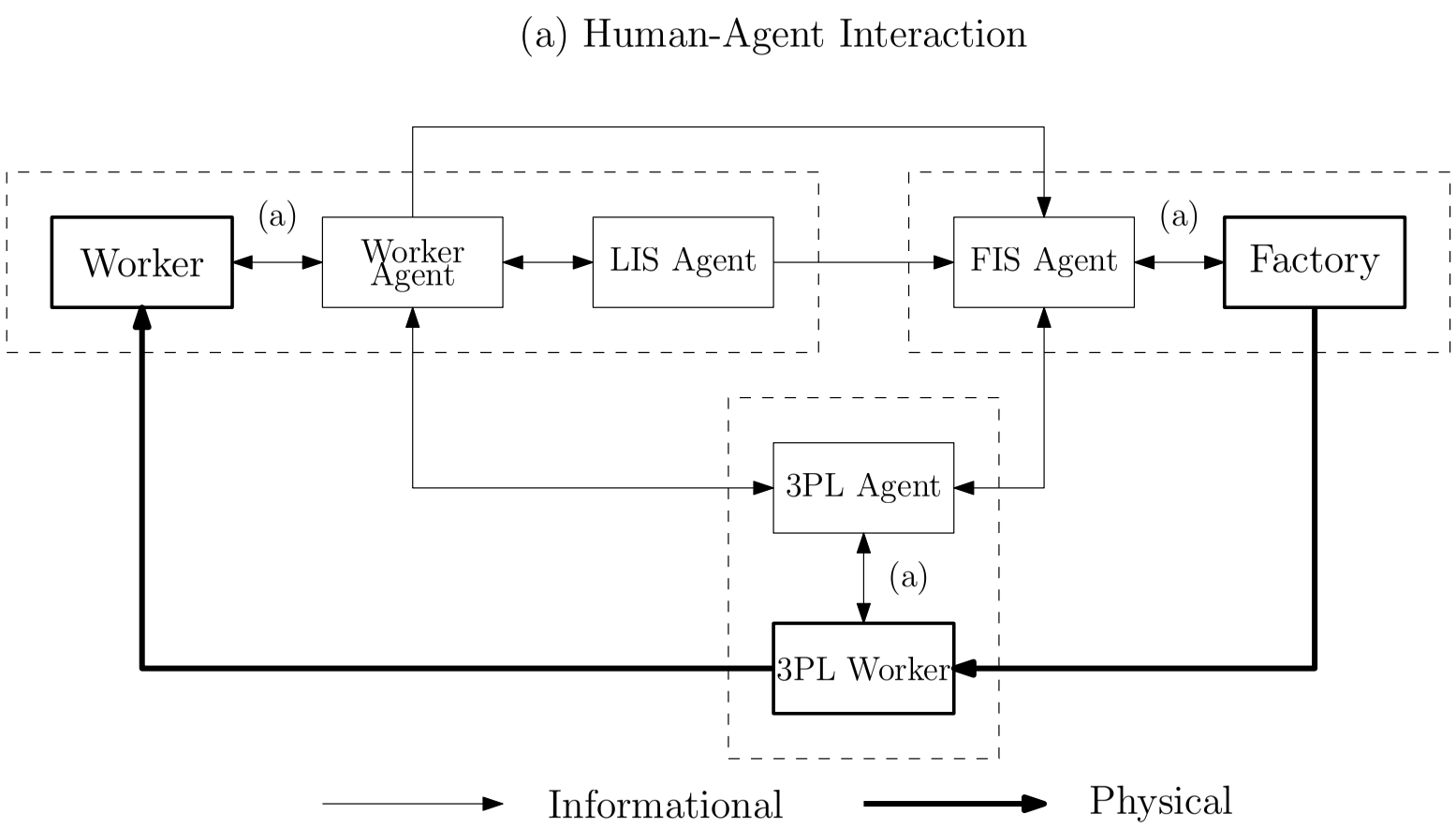


Figure 6: An agent-based design for the interactions among actors in a replenishment process. Here, the logistic center information system and the factory information system are both treated as agents. A worker agent (resp. 3PL worker agent) is built for the worker (resp. 3PL worker). The workers in the factory do not have any LLM-agent to act on their behalf. This system is a cyber-physical system. The agents are the *digital twins* of the physical workers and the factory for information exchange and authorized decision making.

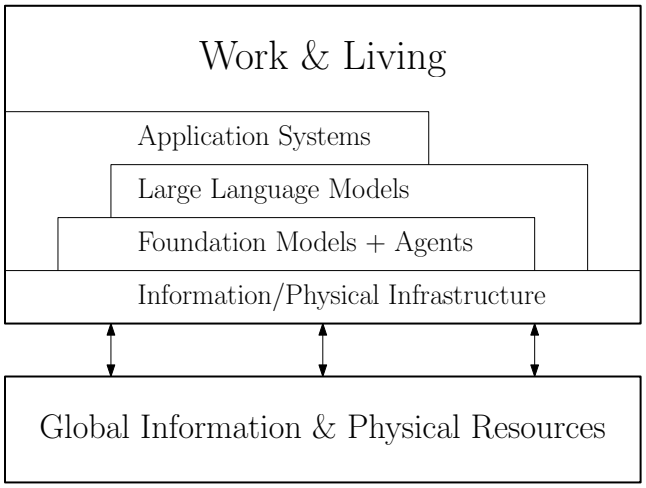


Figure 7: The LLMs and the underlying technologies could support a user to (i) access any global information/physical resource and (ii) complete a task on his/her behalf.

3 Go Beyond (Might Already Happen)

To go beyond, the success of an LLM as a user interface[8] could lead to a *drastic convenience* in global resource access. A user does not have to learn from a bunch of *instruction manuals* on the commands or the procedures in accessing those resources. As long as those resources are accessible, the LLM will act on behalf of the user to access them, see Figure 7.

3.1 LLM-as-a-Service (Agent)

Once a user has a task to be accomplished, the user can simply give *verbal instructions* to an LLM. Then, the LLM will (i) interpret the instructions, (ii) orchestrate the appropriate foundation models, agents and other available software systems, (iii) access those available resources on the globe, (iv) complete the task and (v) report the results to the user.

In other words, the future LLMs and the underlying technologies could support a user to (i) access any global information/physical resource and (ii) complete a task on his/her behalf. If we follow the terminology from cloud technology, I will say that LLM-as-a-service. It is in analog to *software-as-a-service* or *platform-as-a-service*. Besides, we could also say that an LLM is an agent, a personal assistant or a companion.

3.2 LLM-Agent as a Digital Twin

Once a user allows an LLM agent to act on his/her behalf, the user can also set the agent's behavior and thinking exactly the same as himself/herself. In the end, the LLM agent could be perceived as the user's digital twin who is living in the virtual world.

3.3 Interfacing Global Virtual Computer

Once we consider the Internet together with the resources being connected as a virtual computer, the LLMs are now the user interfaces for this global-wise virtual computing machine. The number of users of this global computer will definitely raise drastically in the coming years.

While the number of users to the global computer raises, the number of LLM users will definitely raise. Owing to their user-friendly nature, users will likely be relying a lot more on these LLMs.

[8] Clearly, we can consider an LLM as a personal assistant, as described in Section 1.

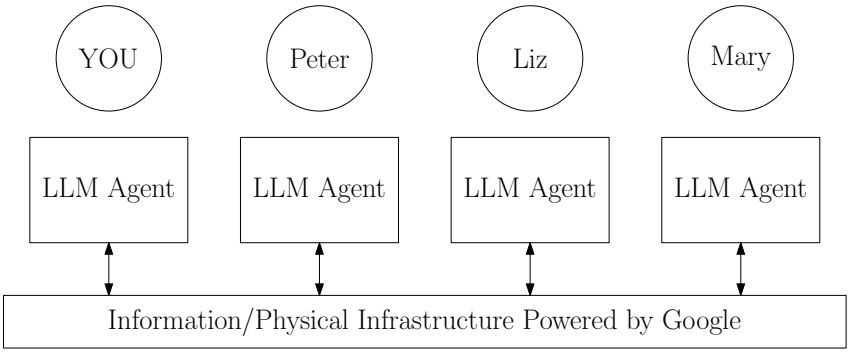


Figure 8: LLM agents are logically connected to form a network of personal agents

Almost every person in the world will be binding to at least one LLM for his/her work and living. To this end, privacy preservation will be a critical issue to every user. *How to protect our privacy during our use the LLMs for work and living becomes an important problem every user has to solve.*

3.4 Personal Agents

To go further beyond, an LLM can act as an agent on your behalf. Depending on your authorization, your LLM agent can read your the mails in your Gmail. It can create, read, write and delete a file in your Google Cloud. It can access an online shopping platform and place an order on a food or a goods on your behalf. Even so, it can pay by your credit card for the order. It can interact with your friends' LLM agents to plan for a party. For your work, your LLM agent can interact with your colleagues' LLM agents in collaborative works and for meetings. Your manager's LLM agent might interact with your LLM agent on your promotion issue. Virtually, all these LLM agents form a logical network on top of an *information/physical infrastructure*. Likely, the infrastructure is powered by Google, Figure 8.

3.5 Global Virtual Computers (GVCs)

Clearly, not all technology firm can realize a *global information/physical infrastructure* to support a *global virtual computer* with dedicated *foundation models*, *LLM* and *cloud*. To be a *global virtual computer*, Google and Microsoft can definitely do so. XAI might be. A list of a potential LLM-powered global virtual computers with their display networks is depicted in Table 1.

3.5.1 Gemini, Co-Pilot and Grok

Currently, two technology firms can do so. They are *Google* and *Microsoft*. It is anticipated that *Google* targets mainly on *personal users* and minor on *working group users*, while *Microsoft* targets mainly on *working group users*. The XAI Grok is yet another LLM which can support similar functions as Microsoft Co-Pilot and Google Gemini. As XAI is attached with the social network X, we could infer that the target users of Grok is the group of the *social network X users*.

3.5.2 Claude and Llama

Anthropic and Meta, which have also been developing large language models and the associated foundation models, are targeting on other user segments. Anthropic is a firm getting investment from Amazon. So, it is believed that the LLM Claude is developed inline with the need for online shopping of Amazon. Meta's LLM Llama is developed for other specialized group of users which has yet to be identified. One possible group of users is the *Metaverse* members.

Table 1: List of potential global virtual computers and their display networks.

GVC	LLM	Users/Clients	Display Network (Coverage[f])	Scale
Alibaba	Qwen	EC Corporate	Alimama (1.3B)	Giant
Amazon	Claude[a]	EC Corporate	Amazon Ads ($\geq$ 120M)	Giant
Apple	ChatGPT[b]	Personal[c]	Apple Ads (2.35B)	Giant
Google	Gemini	Personal[d] & Corporate	Google Display Network (5.4B)	Giant
Huawei	Pangu	Personal[e] & Corporate	Huawei Display Network (10M)	?
Meta	Llama	Social networkers & Metaversers	Meta Audience Network (3.07B)	Giant
Microsoft	ChatGPT[b]	Corporate	Microsoft Advertising Network (2.0B)	Giant
Tencent	Hunyuan	Social networkers	Tencent Union (1.38B)	Giant
XAI	Grok	Social networkers	Google Display Network (5.4B)	?

[a]Developed by Antrophic. [b]Developed by OpenAI. [c]iPhone users and MacOS users. [d]Android phone users and Google computer users. [e]Harmony OS users. [f]The statistics are estimated from the data available on the Internet on December 12, 2025. Noted that there are many display networks operated by web advertising firms in the world, like Taboola, AdRoll, Media.net and Mediavine/Raptive.

3.5.3 Qwen, Pangu and Hunyuan

A number of global virtual computers from China have recently been launched and advanced. They include the Alibaba Cloud, the cloud from Huawei and the cloud from Tencent. Alibaba targets on EC enterprises. Huawei targets on personals and home users. Tencent targets on social networkers. These three LLM-powered GVCs are clearly targeting mainly on the users from China. Moreover, the customer segment of Alibaba is now not just limited to China EC enterprises, but also the EC enterprises from the South East Asia and even from the Europe. The scales of Alibaba, Huawei and Tencent could be gigantic if they merge.

3.6 LLM for Advertisers

It is speculated that one possible reason why Meta develops LLM is to let her advertisers to use the LLM and associated tools to create their advertising materials (in short, ad creative). Together with the Meta Audience Network, the time and cost spent by each advertiser on marketing could be reduced. The return-of-investment of a firm on online advertising could be elevated. In return, the revenue generated by the Meta Ad Service would be elevated.

3.7 Interfacing the World

Figure 9 shows an expansion of the layered structure as shown in Figure 7. The LLMs and the underlying technologies could support a user to (i) access any global information/physical resource and (ii) complete a task on his/her behalf. The information resources are corresponding to those extracted from and added to the digital world. These interactions are accomplished by the information & communication technologies. To interact with the physical resources in the physical world, special equipments like (computer controlled) robots and machines are needed.

To complete the structure, the information/intelligent infrastructure should be developed to support the LLMs to interface with both the *digital* and *physical world*.

3.7.1 Interfacing the Digital World

Today, any human is able to send a text-prompt to an LLM for doing any information task and then put the digital outcome back to the *digital world*. Technically speaking, an LLM can now assist a human user to (i) search and retrieve *digital stuffs* from the *digital world*; (ii) manipulate and create *digital stuffs* to the *digital world*; and (iii) add the new digital stuffs to the *digital world*. In this regard, an LLM acts as an interpreter and interface for a human user to interact the *digital world*.

3.7.2 Interfacing the Physical World

Along this line of thought, one could grab from the diagram as shown in Figure 9 one problem that interfacing *physical world* has yet to complete.

How does an LLM interact with the physical world?

To extract information from the physical world, cameras, microphones and various sensors have already applied. To make changes in the physical world, like moving an item from a factory to a shop, a computer controlled robot (resp. machine) is needed.

Once a user together with an LLM has came up with a task, the LLM can then design the procedure and instruct the robots/machines to complete the task automatically. For some instances, the robots or the machines might not be able to connect to the Internet for any real-time instruction. The robots or the machines must be authorized to make real-time decision by themselves, i.e. autonomous control/decision making.

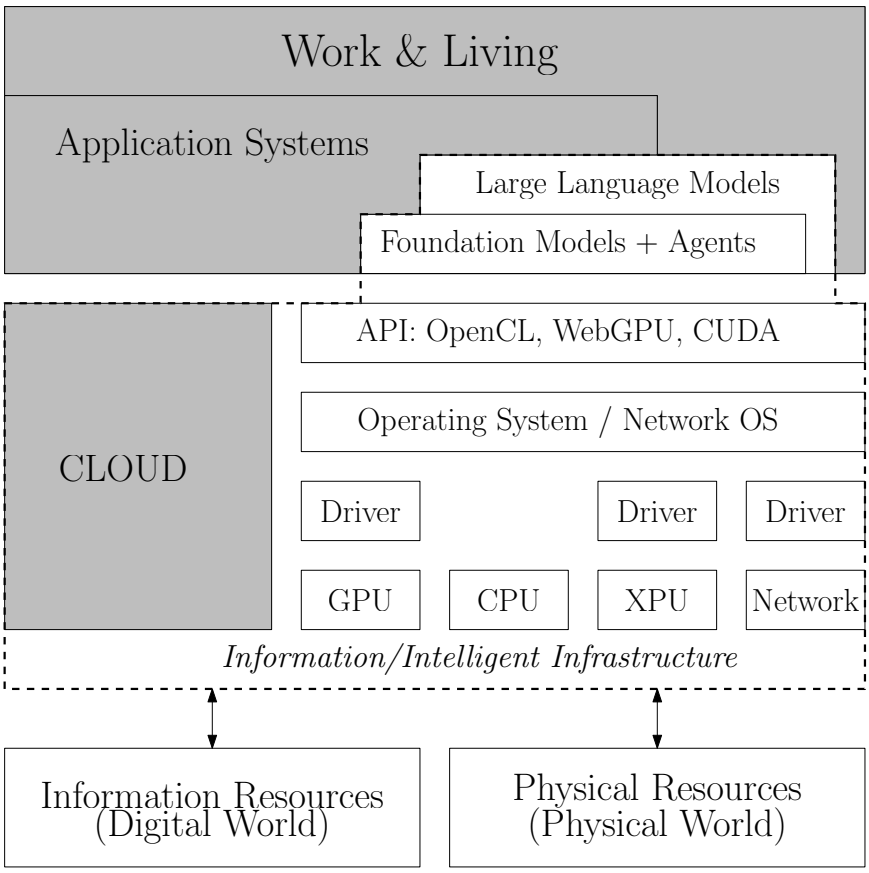


Figure 9: An expansion of the layered structure as shown in Figure 7. The LLMs and the underlying technologies could support a user to (i) access any global information/physical resource and (ii) complete a task on his/her behalf. The information resources are corresponding to those extracted from and added to the digital world. These interactions are accomplished by the information & communication technologies. To interact with the physical resources in the physical world, special equipments like (computer controlled) robots and machines are needed.

The area covering the works, the robotic designs and the associated AI researches is collectively called *physical AI*[9]. Research workload along this line is no lighter than the workload on digital world interfacing. To me, it can be much heavier. One reason is that the problems along this line involve knowledge from multiple engineering disciplines, including but not limited to electrical & computer engineering, mechanical engineering and industrial engineering.

4 Privacy & Security – My Trust

It is clear that *privacy and security* must be a concern to every LLM user. Once an LLM can help the user to create an *agent* to be executed autonomously, one concern is on the level of authorization a user will be assigned to this agent. To this end, a user can only rely on his/her believe and trust on the technology firm which provides the LLM, the foundation models and agents.

4.1 Privacy & Security Issue

One can only believe (resp. trust) that the technology firms like Google and Microsoft will not steal our personal information. Furthermore, the technology firms will develop technologies for the users to defense against malicious attacks, i.e. information leak and hacker attack. This trend is basically the same as the adoption of the use of Gmail.

4.2 Monetary Issue

Another issue is on the payment authorization – *will you authorize the agent to pay on your behalf* ? It is a critical issue to many people. The responses of the people could be summarized in one sentence. As long as the adoption rate of these agents has been raised to a certain level, there will be an increasing number of users authorizing the agents to pay on their behalf. Accordingly, the monetary issue is not a critical issue. This trend is basically the same as the adoption of the auto-transfer service provided by a bank.

4.3 Responsibility of the AI Tech Firms

An AI tech firm will have to develop technologies to defense the users from information leakage, malicious attacks and money transfer malfunctions. Moreover, an AI tech firm has to ensure that the transactional information of the agent is audible (equivalently, transparency). Whenever an extra-ordinary situation has happen, the AI tech firm can trace the potential vulnerability.

With the advancement of the technologies for the aforementioned situations, the AI services developed by the AI tech firm will gain more adoption and eventually more trust from the users. Once a person has realized this raise of adoption, the person would release more authorizations to the AI agents. Therefore, the responsibilities of an AI tech firm is to develop technologies to defense the users from information leakage, malicious attacks and money transfer malfunctions. So that, the AI tech firm can gain more believe (resp. trust) from the public.

5 Reclaim the Roles/Limitations of an LLM

One should be noted that the major technological advancement of an LLM is on its natural language processing. With reference to Figure 6, the interactions among agents have no need to handle natural language as protocols specifying the format of their messages transfer do not need to be understood by human users. Owing to efficiency a business process, the protocol designed for agent communication can ignore human understandable.

[9] One can read the introductions of physical AI from Nividia `https://www.nvidia.com/en-us/glossary/generative-physical-ai/` and IBM `https://www.ibm.com/think/topics/physical-ai` for reference.

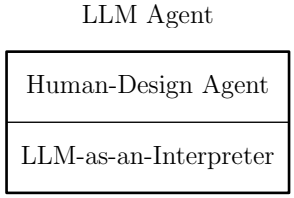


Figure 10: Basically, an LLM-agent could be composed of two parts. One part is the LLM-as-an-Interpreter which interfaces between the human user and the human-design agent (a software). The other part is the human-design agent. It is a software responsible for handling routine works.

5.1 Human-Agent (Human-GVC) Interactions

For the human-agent interactions, the LLM plays an important role. The protocol designed for an human-agent interaction will have to concern the issue on human interpretation. In this regard, natural language understanding will be an important issue to be considered. Therefore, an LLM is act as an interface for the communication among agents and an interpreter for the interactions between the agents and their corresponding users.

5.2 Improve Working Efficiency

One important challenge is raised – *Which group of users the LLMs should be benefited the most?* For the intelligent users, they should have no problem to learn the protocols to communicate with those AI agents. For those normal users who cannot learn the protocols, their usages of the AI agents can ensure that their works could be up to certain acceptance levels. So, an open question is raised. An LLM-agent is able to improve the efficient of both the intelligent users and the normal users. Based upon the proportion between the intelligence users and normal users, an LLM-agent is definitely benefited more for the group of normal users.

5.3 Functionally Limited

The LLMs are clearly the inevitable components *human-agent interfaces*. Thus, an LLM is an essential interface (resp. an interpreter) for a human user to interact with the global virtual computer (resp. the business processes). Except that, an LLM has no use especially on the agent-agent interfaces. Therefore, the use of an LLM is *still* limited to the *human-agent interfaces* and equivalently the *human-GVC interfaces*.

5.4 Reasoning Agent or Interpreter

One should be aware that any current LLM is not a perfect reasoning agent. In some benchmark tests, it might achieve more than 90 out of 100 score. Method like multiple-LLM-agent decision making could be applied to let the LLM-agents come up with a consensus. Still, their conclusions on a particular instant might be wrong[10].

5.4.1 Imperfect Reasoner/Interpreter

Applications of LLM (resp. multiple-LLM) in some complex reasoning domains like medical advises or education could cause severe problems. Treating an LLM as an (imperfect) interpreter seems to be more realistic. In this regard, a human user should be prepared to have good communication skill when he/she is going to interact with an LLM.

[10]Edd Gent, AI's wrong answers are bad. Its wrong reasoning is worse, *IEEE Spectrum (Online)*, December 2, 2025. Available online `https://spectrum.ieee.org/ai-reasoning-failures`.

5.4.2 Human Reasoning is a Key

An LLM is able to interpret simple questions and simple instructions. Human reasoning is certainly a key. *Procedural design* for an LLM to generate programs to accomplish a task must involve the human user reasoning and advises. Thus, the best of an LLM can do is to be an interpreter or advisor, see Figure 10. It cannot be treated as an *autonomy agent*.

5.5 Concrete Operational Machine

Follow the cognitive stages raised by Jean Piaget[11], the cognitive stage of this LLM agent could be up to the *concrete operational stage*. The LLM agent is able to understand the instructions given by the user and then complete the task accordingly. So, I would say that the current LLM agent together with the global virtual computer behaves as a *concrete operational machine*. *Formal operational machine* has yet to be realized.

6 The Roles of Google Gemini to JS

6.1 Interpreter

In the end of the day, the essential role of the LLMs is being an interpreter. An LLM is able to convert the voice speech of a human user to a text message and then interpret the meaning of the message. If the message is a simple question or simple instruction, the LLM is able to give simple response accordingly.

6.2 Virtual Master of Learning

From Google Gemini, I am able to get rich information on a topic which is new to me. For instance, the mechanism behind the real-time bidding and the system architecture supporting online ad auctions can be explored by asking questions to the Gemini. Based upon the information released, I am able to dive in the sources for further information in the topic.

Apart from discussion, Gemini is able to summarize a long conversation between the Gemini and myself. Summarization of a long discussion is a valuable function I have ever imagined. Below is the summary of our conversation on decentralized truth-revealing auction algorithm conducted in May 28, 2026.

```
Summary of Decentralized Truth-Revealing Auction Algorithm

* Core Objective: You have developed a decentralized algorithm
for single-unit demand uniform price auctions that simultaneously
solves for winner determination and price determination without
requiring bidders to disclose their private willingness-to-pay
($u_i$).
* Truth Revelation & Privacy: By keeping bids private, the
algorithm removes the risk of auctioneer malfeasance (e.g.,
"inventing" higher second bids) and "information leakage"
(competitors knowing your budget). This creates a safe environment
for bidders to submit their true valuations.
* Mechanism:
    * Network Topology: Bidders are nodes in a connected graph.
```

[11] Access `https://en.wikipedia.org/wiki/Jean_Piaget` for the information about the four stages laid by Jean Piaget on human cognitive development.

    * Convergence: Nodes update local variables for the uniform
    price (p_i) and a winner indicator (z_i) using a
    consensus-based update function:
    $p_i(t+1) = F(p_i(t), u_i, p_j(t))$.
    * Winner Determination: Once the algorithm converges, a local
    Heaviside function $z_i = h(u_i - p_i)$ determines if the bidder
    is a winner ($z_i=1$) based on the settled market price.
* Tie-Breaking Solution: You've proposed a time-priority perturbation
to ensure a unique winner: $u_i' = u_i + \beta \Delta(T_c - t_i)$.
This treats the bid as slightly higher the earlier it was submitted
(t_i), effectively rewarding early participants and preventing
"over-allocation" of units.
* Remaining Challenge: To prevent bidders from lying about their
submission times, the system requires a robust clock synchronization
method or a decentralized timestamping protocol (such as logical
clocks or blockchain-based commitments).

Summary of Decentralized Truth-Revealing Auction Algorithm

1. Core Concept: A decentralized approach for single-unit demand uniform
price auctions where bidders do not disclose their private
willingness-to-pay (u_i).
2. Privacy & Truth Revelation: By avoiding bid disclosure, it eliminates
auctioneer malfeasance (cheating) and information leakage, encouraging
bidders to bid their true values.
3. Mechanism: Bidders act as nodes in a connected graph. They update local
variables for the uniform price (p_i) and a winner indicator (z_i) using
a consensus-based update function F.
4. Convergence: The system stabilizes when nodes reach a consensus on the
market-clearing price. A Heaviside function local to each node determines
the winner (z_i = 1 if u_i > p_i).
5. Tie-Breaking Solution: To handle identical bids without a central
authority, a time-based perturbation is added:
u_i' = u_i + beta * Delta(Tc - ti). This uses submission time (ti) as a
fair, deterministic "first-come, first-served" priority.
6. Key Implementation Challenge: Ensuring synchronized clocks across the
decentralized network to prevent timestamp manipulation and ensure fair
tie-breaking.