VOL 2023 NO 3 | FALL 2023

International Journal of Law, Ethics, and Technology



La Nouvelle Jeunesse

SUMMER 2023 The International Journal of Law, Ethics, and Technology Staff

Editor-in-Chief
Ross Martin

Senior Article Editor
Haley Sanders
L. Ben
Yan Pan

Staff and Volunteer
Xiaowen Hu
Xuan Zhang
Zixuan Wu
Juncheng Cao

Publication: The International Journal of Law, Ethics, and Technology

cite as Int'l J. L. Ethics Tech.

ISSN: 2769-7150 (Online) | 2769-7142 (Print)

Publisher: La Nouvelle Jeunesse

Address: 655 15th Street NW, Washington, DC 20005

Date: November 12, 2023

Copyright © 2023 by La Nouvelle Jeunesse, except where otherwise indicated.

THE INTERNATIONAL JOURNAL OF LAW, ETHICS, AND TECHNOLOGY is designed to understand all matters relating to law. While the Journal emphasizes theoretical work and broad issues arising from law to technology, no topic of legal interest and no approach to scholarship is excluded. The Journal welcomes outstanding original doctrinal and critical scholarship contributions to domestic and international law and comparative law legal history and legal philosophy. In addition, we welcome interdisciplinary contributions in areas of relevance to the law, particularly work that uses the techniques of the social sciences and the humanities to contribute to understanding legal studies.

RECOMMENDED CITATION FORM: Author, title of work, volume # INT'L J. L. ETHICS TECH. (specific Journal parallel cite) page on which article begins, span of specific pages cited (translator's name trans, year), URL. When the work has an editor or translator, include that information in the parenthetical. Give the editor's or translator's full name(s) followed by the abbreviation "ed." or "trans." as appropriate. Follow with a comma. All efforts should be made to cite the most stable electronic location available. For example:

> Ge Zheng, *Reconsidering Economic Development and Free Trade Taking China's Rare Earth Industry as An Example*, Winter 2021 INT'L J. L. ETHICS TECH. 1, 3-4 (Yan Pan trans., 2021), https://www.ijlet.org/w202101/.

SUBSCRIPTIONS: For now, the print version of the *International Journal of Law, Ethics and Technology* is available to individuals and institutions as per the approval by the editors. To request a place on the list, please email us at info@ijlet.org.

SUBMISSIONS: Please send articles, responses, letters to the editors, and anything else we ought to consider for publication to the *International Journal of Law, Ethics, and Technology* at submissions@ijlet.org.

CORRESPONDENCE: Please write to the *International Journal of Law, Ethics, and Technology* at info@ijlet.org.

Table of contents

Reform and Incentive Mechanism on Equity Structure of Family Enterprises 3

Dongyi Shi

Crossing Boundaries: Autonomous Weapon Systems and the Challenge of IHL Compliance.....21

Maria Gevorgyan

The Influence and Contributions of Confucianism to the Development of International Law 46

Weitao Gong

Three Moral Challenges of Surveillance Capitalism in the Metaverse.....64

You Zhang

This page intentionally left blank

REFORM AND INCENTIVE MECHANISM ON EQUITY STRUCTURE OF FAMILY ENTERPRISES

Dongyi Shi*

Abstract: From the perspective of realizing "everlasting business", how the founders of family enterprises in China can promote the reform on the ownership structure of the enterprises and establish an inclusive incentive mechanism, so as to take into account the inheritance of family wealth, to improve the technological innovation ability and the valuation (market value) of enterprises, is an economic and legal issue that is well fit for the current demand of China's economic situation, and is also related to the long-term development of China's economy. Its legal practice is worthy of in-depth study by Chinese legal scholars. This article reviews the transformation and development process of the ownership structure of family enterprises since China's "Reform and Opening-up", through interviews with the founders of family enterprises and field research in Fujian Province, a representative area of China's family business economy. The article pays special attention to the practice under the current on-going "business succession period" between the first-generation entrepreneurs and the second-generation leaders. The article collects the relevant legal and economic issues in the internal corporate governance mechanism, and explores the relationship between properly coordinating and optimizing the ownership structure of family enterprises, incentive mechanism and maintaining the allocation of wealth resources under the tradition in Chinese family society.

Keywords: Family Business; Ownership/Equity Structure; Equity Incentives; Corporate Governance

* Billion Industrial Holdings Limited, United Kingdom. (2299.HK)

Table of Contents

I. Research Background 6

A. Raised Concerns 6

B. Research Direction 7

II. Analysis on Family Enterprise 8

A. Nature of Family Enterprise 8

B. Problems in Development of Family Enterprises 9

C. Relationship between Family Enterprises and Existing Legal Mechanism and Significance of Mutual Enhancement and Co-evolution 10

III. Arrangement of Equity Structure Reform of Family Enterprises 11

A. Separation of Ownership from Operation at the Level of Enterprise 11

1. Introduction of Manager System 11

2. Optimization and Adjustment of Equity Structure Inside Families 12

3. Decreasing Family's Shareholding Proportion in the Enterprise and Introducing Equity Incentive and Employee Shareholding Plan 12

4. Introduction of External Strategic Investors 13

B. Legal Form of Equities of Founding Families - Financial Instruments (Family Trust) and Setting of Two-tier Equity Structure 13

1. Family Trusts, Family Investment Offices, Family Foundations and Other Financial Instruments 13

2. Two-tier Equity Structure (Duality Ownership Structure) Establishment 14

IV. Equity-like Incentive and Equity Incentive Mechanism under the Equity Structure Reform of Family Enterprises 15

A. Equity-like Incentive – Virtual Equity Plan 15

1. Introduction of Virtual Equity 15

2. Cases Study on Shared Plan and Huawei's Virtual Equity 16

3. Potential Issues of Virtual Equity Plan 16

B. Equity Incentives 17

1. Equity Incentive Overview 17

2. Issues in the Practice of Equity Incentive 18

3. Employee Shareholding Platform (Limited Partnership) and Compliance Check 19

Conclusion 19

Equity structure serves as the cornerstone of corporate governance, and plays a significant role in internal governance mechanism, operation and development of enterprise. Such significance is in particular highlighted in the enterprises that already evolve into bigger and stronger size depending on the equity structure, which usually determines the development direction and performance of the enterprise.

The reform and optimization of equity structure are conducted in several aspects including equity allocation, assignment, acquisition and merger, and additional issuance. Generally, the reform to corporate equity structure is critical to enterprise development and destiny. Effective incentive produced by the reform is decisive to the future of the enterprise. In modern enterprises, the setting of equity structure and incentive mechanism are evolving from time to time on the condition of developing market economy and diversified environment requirements for relevant industries. Essentially, the development history of modern market economy equals the history of corporate governance system driven by equity structure subject to continuous change.

Family enterprise is one of the enterprise forms in the modern market economy that is known for its longest history. During the history of reform and opening-up in China, the family enterprise which is an integral part of the private sector, splendidly pushes forward the economic development. In the current stage and after the ups and downs for more than four decades, the family enterprises that have already obtained achievements enter the period of criticality embarking on the power transfer from the Generation Create. Challenged with the transformation-oriented development centering on power transfer, the fulfillment of good corporate governance structure and adaption to the steady enterprise development based on the optimization design of equity structure and equity incentive will not only produce huge influence on these family enterprises, but also have significance to sustainable economic development and industry upgrading in China. Ownership can be separated from managerial authority by means of equity structure reform and incentive mechanism design performed by these family enterprises and considering the actuality of involved enterprises and families. A shift will be triggered from consanguinity management to management by professional team. Henceforth, the family enterprises will not be dominated by private capital, but grow into the public enterprises that accept market supervision. Stemming from the perspective of practice, it is more than a study on economic incentive mechanism, but a law-related research.

I. RESEARCH BACKGROUND

A. Raised Concerns

It is of great historical significance for the equity structure reform and incentive mechanism study of China's family enterprises. At present, the research on this topic is the necessity of the times. China has implemented the reform and opening-up policy for almost 45 years. In the time of China's economic development, apart from the state-owned company, foreign investment companies and IT companies which are generated by technicians after the 1990s, the family business which evolved from the original township enterprises and family workshops in the 1980s is an important force for China's economy. Especially in some regions with developed industries, the family businesses which have accumulated for decades are an important carrier for China's profession dedication, specification and innovation, as well as the potential champions in near future. There are also plenty of family businesses in other major industries, such as real estate, household appliances and consumer products. Most of them were from "Gen Generation (1st Entrepreneurial Generation) ". After 20 to 40 years' effort, they are gradually out of the stage, but facing their business handover problem which become a

common topic among them. How to achieve the optimized design of equity structure and equity incentive in the transformation and development during the business handover time is of core significance to these enterprises, and also plays an important guidance role in the sustainable development of China's economy.

The author was born in Quanzhou city, which owns the most proactive private economy and the most centralized family businesses in Fujian Province, and therefore, is more than close to realize the urgency and necessity of this study. As early as more than ten years ago, there was a case that the family enterprise Xin Hua Du Group hired Tang Jun, a professional manager, to be in charge of the company, but it was unsuccessful in the end. Some well-known local founded companies in Quanzhou such as ANTA and HENG'AN, also invite foreign professional managers and innovative teams in recent years, some are successful, while some failed. Therefore, it is a problem that Chinese family businesses tend to face sooner or later.

The equity structure reform including incentive mechanism to the Chinese family enterprises is also of far-reaching significance for promoting national development and pooling resources. After the explosive economic development from the 19th century to the early 20th century, the USA also experienced the similar period and was relatively successful when its first or second entrepreneur generation retired. Some well-known families, such as Rockefeller family, have finished the transformation successively, providing the initial capital source for various funds in the USA, enriching and improving gradually the economic system structure. In addition to the inheritance of the common values, the Rockefeller family is also good at dealing with a variety of financial instruments. For example, family trust is one of the methods that plays an important role, which could be referenced by the enterprise's equity structure reform to a large extent. That is, the family will customize a trust for each descendant member or each small family according to the actual needs, which will be handed over to the same trustee for management. From a legal point of view, the properties will be converted into trust property after the trustor transfers the properties to the family trust. Afterwards, the trustor does not have the ownership of these properties, while the beneficiaries are entitled to benefits produced by the trust property. Through this mechanism, the modern family that founded the enterprise will avoid the fate of the natural dissolution of the land property based on the feudal lord in the previous days of modern civilization with the equal distribution of properties by multiple heirs, and promote the ultimate separation of management from ownership of the enterprise at the level of family wealth, so as to ensure that the family wealth can be centrally managed and used. This model fulfilling the ultimate separation of management from ownership (founder group) has been also practiced in the wealth management of new technology companies and their founder families (e.g. information technology companies, including HP, Microsoft and Google, which are not only listed and publicly traded at the enterprise level, but also have established various forms of legal entities such as funds and foundations at the level of founder's wealth management). Thus, this is a proposition with universal value.

B. Research Direction

At present, due to the reason that the China's academic arena copies the current study themes of the USA[1], for family businesses in different industries, based on its internal driving

[1] For how to promote innovation on Intellectual Property laws through bonus and tax incentive, and relevant studies on how Company Law and Security Law support US capital market, refers to Michael Abramowicz, Perfecting Patent Prizes, 56 Vanderbilt Law Review 115 (2003); Nick Bloom et al., Do R&D Tax Credits

force, it is also worthy of studying whether to focus on discussing technological breakthroughs or exploring equity reform. In addition, we need to consider lessons learned from the dismissed founding team due to improper balance on equity incentives. Compared with foreign countries, in regards of balancing the interests of new and senior employees inside and outside the family business, the market environment and relevant systems are not fully matured, thus it is necessary to improve the system construction, strengthen market supervision. The article aims to do relevant research on promoting the equity incentive mechanism for achieving the best result, protecting small and medium-sized investors, and balancing the needs of different interest with the precondition of promoting a long-term development of family business.

However, the necessity and common needs of the equity structure reform of family enterprises receive less attention in both academia and industrial fields. And the systematic studies on the internal profit-driven motivation behind the equity structure reform in the long-term development of family enterprises are performed by members of family enterprises in few cases. What's more, Generation Create and its family may be expelled from the enterprise due to the improper design of the equity structure reform mechanism, and it is taken as the instructive precedent. From then on, most of the enterprises are on the decline. Therefore, it is the direction that we need to think about. From the perspective of practical operation, and compared with foreign countries, there are no matured market environment and legal system to balance the interests between the new and the old employees inside and outside the family enterprises, which necessitate the reinforcement of market supervision and the system construction improvement. It is necessary to promote the reform of relevant equity structure and the setting of incentive mechanism to achieve the best incentive effect on the premise of promoting the enterprise development in the long run, as well as protecting small and medium-sized investors, and balancing the needs of different stakeholders.

II. ANALYSIS ON FAMILY ENTERPRISE

A. Nature of Family Enterprise

Both at home and abroad, most of the enterprises are spontaneously incorporated by a large family or a combination of individual families (e.g. two partners from different families). Up to this date, family business remains a common form of business organization and a mainstay of the world economy, which provides jobs and maintains communities and environment.[2] In fact, the globalization of the last 30 years contributed to a large number of emerging enterprises in developing countries. In addition to the technology-based companies relying on venture capital, which are the major concern of media, there are a variety of family businesses in the real economy. As revealed by research and expert prediction of McKinsey & Company, the share of family businesses in the global Fortune Top 500 has increased from 15% in 2005 to 19% in 2013 owing to the rapid growth of the emerging new businesses. By 2025, family businesses will account for 40% of the world's large enterprises, higher than the figure

Work? Evidence from a Panel of Countries 1979-1997, 85 Journal of Public Economics 1 (2002); Brian Broughman & Jesse M. Fried, Carrots and Sticks: How VCs Induce Entrepreneurial Teams to Sell Startups, 98 Cornell Law Review 1319 (2013).

[2] Tharawat Magazine, The Economic Impact Of Family Business, Issue 22, May 2014, source：https://www.tharawat-magazine.com/magazine/issue-22-economic-impact-family-business/, visit on May 2, 2022.

of about 15% in 2010.[3]

The perspective of "family business" connotation that, the family retains control over the enterprise, while their generations are involved in the business operation,[4] from American scholars Astrachan and Shanker, shows that, the character of family is a high concentration of control and equity over the enterprise by family members and their generations. This is in line with the rules of many enterprises in the early development of USA, or most American's manufactures, furthermore it shows high consistency with the transformation of Chinese township enterprises and family business developed based on family workshops. This is a transformation into a modern enterprise, it is a comprehensive topic, which involves not only related (enterprise side) management and incentive mechanism, but also family inter-generational inheritance, role positioning and how to preserve, inherit and withdraw from the family equity for wealth management. According to the Corporate Governance Gap of Prof. Kastiel and Prof. Nili, the proportion of family enterprises that sticks to good governance disappoints the assumption of the public.[5] Apart from the unevenness in distribution, the development of these family enterprises also encounters various influences and challenges, including the periodicity in economy and industry, induction from external investors and other business agents, as well as problems related to society, academia, environment and social ethics. Under this circumstance, most of the family enterprises will undoubtedly transform to be modern enterprises for the purpose of long-term existence and development. At the same time, it is a proposition comprehensively and systematic work practically, which not only involves the management mechanism and incentives on the level of enterprises, but also affects the inter-generational transfer on the level of families, and the role definitions of Generation Create families and enterprises. Even the keeping and inheritance of family wealth are also included in this case. It further determines the separation of the families from the enterprises in the long run (e.g. separation of Ford families from Ford Auto). Family wealth management and evolution are also considered.

B. Problems in Development of Family Enterprises

Generally, the founder of the enterprise is entitled to the ownership and control over the enterprise in the early stage of the family enterprise, which make sure that the interests of the enterprise are consistent with those of the founder. And the corporate governance is not taken as a major concern since the company itself is not subject to any conflicts and contradictions. Business expansion receives most of the attentions.[6] However, the demands for scientific management are increasing during the expansion of the enterprise. The centralization of power existing in the internal structure of family enterprise prevents the enterprise from further development. More and more family enterprises employ professional managers for the purpose of professional management which has become the trend of corporate governance. Governance and reform of the family enterprises center on the adjustment of equity structure. In addition to the reform of excessive centralization of equities and control, and to improve the stability of the human resources on the higher level, it is also suggested to invite investors, import employee equity incentive mechanism and be listed abroad so as to fulfill further management

[3] Åsa Björnberg, Ana Karina Dias & Heinz-Peter Elstrodt: Fine-tuning family businesses for a new era, source：https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/fine-tuning-family-businesses-for-a-new-era, visit on May 2, 2022.
[4] Joseph H. Astrachan & Melissa Carey Shanker, Family Businesses' Contribution to the U.S. Economy: A Closer Look, Family Business Review 16.3 (2003), p.211-212.
[5] Kobi Kastiel & Yaron Nili, The Corporate Governance Gap, 131 Yale Law Journal, p.787.
[6] Elizabeth Pollman, Startup Governance,University of Pennsylvania Law Review 2019, Vol. 168: 1, p.196-197.

for the equity structure. Regarding the practices carried out in China over previous years, the reform of equity structure is beset with difficulties in the developed regions densely located with family enterprises such as Shandong, Zhejiang, Jiangsu, Fujian and Guangdong due to the characteristics of Chinese family enterprises and insurmountable problems existing in the resources distribution inside the family enterprises, including the benefits transfer caused by nepotism during the expansion of the family enterprises.

During the research on family businesses, one of the most important existing topics is, how to balance the three aspects of interests from the entrepreneur's perspective, including family wealth inheritance, improving the enterprise's technological innovation, and further to improve the enterprise's valuation (market value). The author has experienced many specifics and challenging questions. For example, how can the first family generation leaders improve the corporate governance structure and ensure its control power when setting up funds for future generations? Whether the regulatory rules and institutions guarantee the establishment of a double equity structure to achieve the organized purpose and institutional arrangements are perfect? Is equity incentive an effective mechanism for all industries and interest groups? It is not surprising of these findings, because the balance and trade-off between the positive and negative benefits of the equity structure reform of family businesses have not yet been deeply studied.

These are crucial topics, because the answers not only relate to the scope of equity structure reform, benefit distribution and the internal logic of corporate governance, but also allow us to understand better on the evolution of equity structure reform in a transitional economy with a booming private economy such as China. Since the reform and opening-up, the private economy has had 360 degree changes. However, with the promotion of economic reform, the rapid transformation of the ownership structure and incentive methods of family business has also showed problems. In order to resolve these open topics, the Chinese people urgently need a well-designed system that keeps pace with the times.

C. Relationship between Family Enterprises and Existing Legal Mechanism and Significance of Mutual Enhancement and Co-evolution

This is not only an economic incentive mechanism study, but also related to law research. *The Company Law* is one of the most important laws for the economy. It has a potential impact on promoting the equity structure reform being built by family businesses. First of all, *The Company Law* influences the corporate governance structure and mechanism, the responsibilities of Board members and high management, affiliated transactions and the protection of the interests of shareholders, especially minority shareholders. Secondly, the ownership structure plan may affect the economic and political stability by affecting the creation of wealth and the distribution of interests. Thirdly, under the protection of the legal system, a reasonable equity incentive system can better balance the needs from various interests via the control rights' distribution. Therefore, it is not hard to understand why private enterprises, especially family business, need to establish a clear equity structure to optimize resource allocation under the institutional guarantee of full *Company Law*. The research on the equity structure reform and equity incentive mechanism that provides consideration to the family wealth inheritance will certainly stimulate the establishment of legal framework in China as well as relevant legal research.

It is recommended to stem from the benefits of the stakeholders and break down the equity structure reform of the family enterprises, which could be highlighted through the following aspects: to define the benefits or demands that need protection; to determine the

incentive measures for personnel who are potentially trained to be in charge of enterprise operation, to encourage or prevent the actions of relevant personnel. These measures will contribute to the basic formation of the complete equity structure reform plan for family enterprises.

In view of these objectives, the optimization of the corresponding legal framework and the application of appropriate legal tools are crucial for the development of family enterprises in the long run. In current stage, China's relevant legal system usually takes ambiguous and tentative positions on and in respect of issues related to regulating the equity structure of family enterprises, leading to confusion and uncertainty in practice and also producing various disputes. This situation must be changed. Meanwhile, the tradition of Chinese society still needs to be considered in the relevant legal practice of Chinese family enterprises. This tradition as mentioned here is more prominent in the field of family enterprise management, which inevitably involves people having power. China often needs to compromise the reality in the process of reform and opening-up due to the pragmatism tradition prevailing in the Chinese society. In the historical operation of Chinese family enterprises, various principles which are socially recognized are inevitably adopted. For example, many major issues related to the development of enterprises and the distribution of interests are handled morally or under the table, resulting in the formation of the situation in the current stage that cause conflicts with laws in realistic and practical cases, or the situation that is adverse to assuring the benefits and interests of the core founder of the family enterprise. It is worthy of in-depth study by legal professionals under the topic of family enterprise reform on how to properly coordinate the relationship between the current situation and the legal system, how to not only optimize the equity structure, ensure family inheritance, maintain the allocation of wealth resources in China, but also encourage the reform of modern enterprise system, cultivate new enterprise talents (especially the talent groups outside the family) generation by generation, and transform the enterprise development to the technological innovation led by talents, so as to improve and share the enterprise value with the whole society. This research will definitely be beneficial to the systematic optimization and reconstruction of China's legal system which still awaits improvement in many aspects.

III. ARRANGEMENT OF EQUITY STRUCTURE REFORM OF FAMILY ENTERPRISES

A. Separation of Ownership from Operation at the Level of Enterprise

1. Introduction of Manager System

As the operation scope and size expand gradually, the operation capital and involved transactions are increasing. It is a new type of corporate governance that the managerial right will be assigned to any other individuals or entities, and the ownership enjoyed by the owner of production goods remains unchanged when the company systems have developed to a certain stage and extent. The business and operation size of the joint-stock firms necessitate the professional knowledge and capabilities in scientific planning and operation obtained by the operators. For instance, the land lord may lease his or her lands to peasants for farming, or the BOD may entrust CEO to be responsible for business operation.

This reform is seemingly justified by theories. However, it is hard to be implemented in practice. The reason is that major changes in the power of enterprises are involved, especially for Chinese family enterprises that strongly rely on operation by people having power. The assignment or the transfer of important positions, is of vital importance for two consecutive

generations of Chinese family enterprises, even for brothers and sisters of the same age, let alone "outsiders". Especially in the environment of Chinese family enterprises in the private sector, a realistic problem shall not be neglected with a great many of managers having the background of training by foreign enterprises "to be introduced at a fixed price" in the market. Therefore, days since the beginning of 21st century have witnessed many classic cases that the majority of Chinese family enterprises have failed in the introduction of managers, and successful cases are rarely found. Even some of the state-owned enterprises cannot escape from the curse when they recruit talents and professionals out of the government or state-owned systems. The curse that the rise and fall of Chinese enterprises depend on one generation still awaits a systematic resolution.

2. Optimization and Adjustment of Equity Structure Inside Families

Establishment of a "family-holding platform entity" is a proposition raised here, which serves as a carrier for the survival and development of family wealth, and further realizes control of the enterprise through the investment platform. Under this entity, the equity inheritance or equity division is merely limited to the family members, which could prevent the family enterprises from equity confusion at the enterprise level (e.g. it directly specified that each family member straightly holds a certain proportion of the shares of the operating company in the early stage). The setting of the family platform entity is conducive to the internal interests' distribution among family members, e.g. by redeeming the shares of family members who held shares of business operation in the early stage, but withdrew from business operation already.

Here is another solution that is adopted by Chinese family enterprises. In the context of family enterprise inheritance, the founder may take into account dividing the business of the family business into multiple sub-businesses (Chinese family enterprises usually involved in multiple industries, e.g. the common combination in some economically developed cities such as being engaged in both the manufacturing industry that could produce wealth and local real estate industry, and also making minor investments). Each enterprise needs to also consider the cross-shareholding among multiple family members to balance the interests of family members without prejudice to the family wealth inheritance. For example, the family business is divided into several different types of sub-businesses, which are respectively handed over to the successor who is interested in such sub-business. However, the presiding successor of each sub-business only holds part of the equity of the business company, while the remained is shared equally by other successors. The aforesaid measures will enable the successor to be in charge of each sub-business to have absolute control over the enterprises under his or her operation, and have a high degree of independence and self-motivation, so as to avoid costs and internal conflicts aroused from negotiation with other brothers and sisters.

3. Decreasing Family's Shareholding Proportion in the Enterprise and Introducing Equity Incentive and Employee Shareholding Plan

For a long-term perspective, the family enterprises need to bring in a lot of new members and share the development achievements and created values with the employees (including senior officers and critical staff). Since the excessively high percentage of the family shareholders among the entire shareholders is improper, it is recommended to introduce employees that are allowed to hold shares to a certain extent so that the employees and the enterprise will have plenty of skin in the game by means of diversified equities and new distribution mechanism. According to *No.6 Guidance on Supervision of Non-listed Public Company - Supervision Requirements for Equity Incentive and Employee Stock Ownership*

Plan (Trial), the participants of the Employees Stock Ownership Plan consist of the employees that have entered into the labor contracts, including those from the leadership. Employees participating in the Stock Ownership Plan are entitled to have equal equities with any other investors and independently liable for their profits and loss, and risks.[7] It is permitted to provide the employees to be in charge of operation and management by means of incentive systems, including issuing shares to the managers that perform excellently in the performance assessment. Take the family enterprise groups in Quanzhou for example. Local listed companies such as ANTA and HENG'AN (these listed companies being defined as the family enterprises from their culture attributes) issue shares to thousands of grantees depending on the stock incentive plan.[8]

4. Introduction of External Strategic Investors

The introduction of strategic investors to balance the company's internal governance structure depending on the company's development needs will contribute to improving the control onto the enterprise exposed by the leadership so as to ensure that a series of enterprise development measures formulated by the leadership can be effectively implemented. Enterprises may consider those enterprises that have transactions of significance and stable business relations with themselves and being engaged in the same industry chain during selection of strategic investors. Financial investors however are also imported in combination with various purposes, e.g. Gree's introduction of Hillhouse in recent years. There is an economic reality worthy of objective consideration in China. Central enterprises, local platform state-owned enterprises and other entities have tight control over financial resources after the year of 2010. By contrast, involvement especially in the industrial field of competitive market was seldom seen at the level of industry. Surprisingly, China's national policies prefer to support the real economy since 2020, and many of the real economy in the competitive market are Chinese family enterprises, such as those incorporated in Zhejiang and Fujian Province. It is a topic worthy of study that whether central enterprises and other state-owned platform companies of all kinds can find the access to family enterprises as appropriate external strategic investors and fulfill co-existence. For example, Billion Industrial Holdings Ltd. (2299.HK) in Quanzhou City introduced the state-owned China Energy Conservation Group with a large proportion of additional shares in 2012, and established a harmonious relationship, which result in the rapid growth of Billion Holdings in the era of 2010.

B. Legal Form of Equities of Founding Families - Financial Instruments (Family Trust) and Setting of Two-tier Equity Structure

1. Family Trusts, Family Investment Offices, Family Foundations and Other Financial Instruments

Compared with the aforesaid enterprise-level business focusing more on the restructuring of corporate equity, the flexibility of wealth management corresponding to equities controlled by the family will be more diversified and realistic.

Firstly, equity may be diluted or acquired maliciously during inheritance. Therefore, the establishment of legal and financial instruments such as credit, insurance and fund for

[7] No.6 Guidance on Supervision of Non-listed Public Company - Supervision Requirements for Equity Incentive and Employee Stock Ownership Plan (Trial), p.9.
[8] Announcement on Granting of Incentive Shares by ANTA Based on Stocks Incentive Plan (November 5, 2019); Announcement on Granting of Incentive Shares Based on Stocks Incentive Plan and Issuing of New Shares Based on General Authorization Made in 2021 by ANTA (April, 4, 2022).

family members and future generations will support to prevent equity from being diversification. The enterprise can stipulate in the relevant articles of association that the majority shareholders of the family in the credit should be designated as the subject with actual control power, and implement the independent voting rights so as to prevent the family heirs from losing control power of the family wealth in case of any disputes due to the separation of ownership and management rights.

Secondly, the equity wealth controlled by the family can be transformed into various legal terms. In recent years, there are some foundations and other organizations, which are usually applicable to enterprises with large scale and good reputation, and which also serve as one of the ways for the country to achieve the 3rd wealth distribution and fulfill common prosperity. For example, Mr. Cao Dewang, who founded Fuyao Glass, implemented this method when he initiated Heren Charitable Foundation, and the founding team of HNA Group also launched the Cihang Charity Foundation. These foundations are generally established in the way of cash investment, and then the controller donates the equity to the foundation to achieve asset separation and equity structure optimization.[9] Before setting up the foundation, the sponsors need to sort out their total assets, determine the share of assets donated to the foundation on the basis of retaining sufficient assets required by family members, and ensure the working capital required for the subsequent operation of the enterprise.

The existence of these foundations is of great significance. For example, Mr. Cao Dewang has initiated Fuyao University of Science and Technology, and Heren Charitable Foundation played a critical role during the process (a contributor in legal sense). Over recent years, the new emerging private universities in China are all initiated by some renowned enterprise celebrities, relying on the practice of foundations in most cases. It highlights the incentive effects produced from enterprise creating and sharing wealth with the society by means of talents education and training. This effort will play a more important role in China in the future.

2. Two-tier Equity Structure (Duality Ownership Structure) Establishment

From the perspective of initiating social funds, some family enterprises tend to perform "de-familiarization" in share proportion, and gradually move towards professional management and modernization. In addition to the fund establishment, the establishment of "Two-tier equity structure" (Duality equity structure, or equity structure with different voting rights) also has the impact on optimizing the governance of family enterprises. The two-tier equity structure is equipped with the institutional advantages of "meeting the different needs of shareholders, promoting the company's long-term goals, avoiding hostile acquisitions, and enhancing the efficiency of decision-making".[10] The *Company Law* of China applicable for the time treats the effect of two-tier equity structure differently based on the company strategies. In regards of the limited liability companies, the companies are allowed to modify the voting rights allocation through the Articles of Association pursuant to the *Company Law*.[11] From literality of the provisions, the two-tier equity structure should be permitted. As for the stock corporations, the *Company Law* prohibits any other form of voting rights' allocation depending on different equity ratios.[12] "As revealed by the difference, the *Company Law* tries to reach a

[9] Problems to be Concerned before Initiating a Foundation, source: https://www.sohu.com/a/405321016_100244361, visit on May 4, 2022.
[10] Guo Li, Peng Yuchen: Reflections and References on International Regulatory Experience of Double-Level Equity Structure, published in Journal of Peking University (Philosophy and Social Sciences Edition), 2019 (2).
[11] Article 42 of the Company Law of the People's Republic of China.
[12] Article 103 of Company Law of the People's Republic of China.

balance between the ideas of shareholder autonomy and the ideas of investor protection. Of course, this balance is subject to change."[13]

IV. EQUITY-LIKE INCENTIVE AND EQUITY INCENTIVE MECHANISM UNDER THE EQUITY STRUCTURE REFORM OF FAMILY ENTERPRISES

In the context of the reform on the equity structure of family enterprises, the core of incentive mechanism is mainly designed around the equity or benefits thereof. From the realization perspective, there are mainly equity-like incentive (informal equity, such as virtual equity plan) and equity incentive.

A. Equity-like Incentive – Virtual Equity Plan

1. Introduction of Virtual Equity

When discussing virtual equity, it is firstly to ensure that the company's implementation plan targets to "specific group"[14] people and will not involve any illegal fund-raising. As an ancient and common crime, illegal fund-raising, includes the promise of a high rate of return on capital, defrauding the investment of any social public. And it is a deceptive behavior of using the consequent investment to repay the previous investors.[15] Even if the incentive investors are required to make capital investment, it cannot be handled according to the actual capital from shareholders, but only to be credited based on the external loans of enterprises. As per the advantages of this measure, the company can get some cash flow to alleviate the financial pressure of the project. Meanwhile, the disadvantages are also obvious that part of the incentive from the company is the usage funds, which will greatly reduce the incentive motivation. In order to keep the core employees, the virtual equity held by employees can be transferred into actual equity via certain proportion under certain conditions so that employees can become real shareholders through capital increase and share expansion or equity transfer.

In practice, this situation occurs at the level of specific project in most cases, which is, to incentivize the key people at the project level rather than company level. Although listed companies can achieve equity incentives by issuing shares, at the same time, there are top management who are only in charge of a specific project rather than the entire listed company, especially where the business is mainly based on project. It frequently occurs in real estate project, and also adopts virtual equity design method. There is a so-called "co-investment system" for real estate projects. The top management involved in the system must invest to specific projects in order to obtain the project dividend. Then types of issues come. For example, should it be regarded as a special debtor-creditor relationship if the project company is given virtual equity? Any methodology of contracting requires answers from 2 questions: what risks the parties must face and what management methods the parties can use to eliminate or reduce these risks.[16] As long as the rights and responsibilities of both parties that are involved, a proper

---

[13] Guo Li, Peng Yuchen: Reflections and References on International Regulatory Experience of Double-Level Equity Structure, published in Journal of Peking University (Philosophy and Social Sciences Edition), 2019 (2).
[14] The first paragraph of Article 2 of the Measures for the Supervision and Administration of Unlisted Public Companies stipulates that an unlisted public company refers to a joint stock limited company whose shares are not listed and traded on the stock exchange under one of the following circumstances: (I) the issuance or transfer of shares to specific objects results in a total of more than 200 shareholders; （II） Public transfer of shares.
[15] Wang Kaiyuan, Shi Dongyi and Luo Meiting: Empirical Study on Ponzi Schemes in China, Financial Law Forum, 2010 (80).
[16] Matthew Jennejohn, The Private Order of Innovation Networks, 68 Stanford Law Review 281 (2016), p.294.

legal framework is necessary to ensure the normal operation of the mechanism.

2. Cases Study on Shared Plan and Huawei's Virtual Equity

There is such practice in the family business group in Quanzhou. Some more powerful unlisted family businesses, invited professional institutions to make a sharing plan for the core top management and directors from non-family business members. The purpose of the sharing plan is to encourage the management in advance and sign the agreement with the employees in order to avoid the inconsistent interests between the core management and the company, which will further impact the listing. The agreement can only be saved by the company and no copy for employees. The employee can only assign three agents to lock the agreement in the locker. With the permission of the company, the three people can enter their own passwords and open it at the same time.

This sharing plan is a bonus incentive of "achievement and risk sharing" initiated by the shareholders as individual and face for key personnel. This plan has no relation with the legal representative of the company, but only takes the company's performance results as the incentive target and calculation basis. Participants shall receive the relevant share via actual capital investment. According to the share from the shared plan, participants can get the related bonus incentive income from shareholders when they achieved performance KPI as agreed with shareholders. At the same time, participants also have to bear if there were any financial losses. The methodology of profit distribution and loss bearing of the plan is to be decided by the shareholders. Regarding the system arrangement, the sharing plan participants shall focus on the common business objectives; they shall not only consider the short-term profit incentives, but also the long-term development of the company, support and cooperate together for the strategic direction and core business of the company. Regarding the capital investment and incentive policy, participants can enjoy the loan provided by shareholders to subscribe the shared plan shares, and receive the bonus provided by shareholders according to their share portion.

In similarity, every year, Huawei will nominate the outstanding employees and issue a contract with the specific shares that they can purchase from the company. After reviewing the contract, the employee will sign off, then submit, and the contract will be kept by the company. Although there is no copy and corresponding shareholding certificate, employees can check their shares hold via their internal account number. At the same time, the employee names will not be shown in the industrial and commercial registration certification, and the equity of the employees will be deputized by the Huawei labor union, essentially the contract proves the recognition of employees' equity. However, this recognition is partial, it mainly aligns via legal contract; this belongs to a virtual equity[17], rather than the ownership equity legally.

3. Potential Issues of Virtual Equity Plan

It should be noted that the top management purchase the virtual equity, but did not actually own the shareholder's rights and interests defined in the *Company Law*. If the company loses money, the principal of the virtual equity will also be lost. At this time, will the interests of the company or of the virtual equity investors be protected? How to define the nature of this investment, whether it belongs to shareholder's loan, company loan, or equity investment? If this money is used as a loan, whether it should be separated from the business operation? If so,

[17] Yang Jingyi: The Unique Role of Virtual Equity Incentive from Huawei's Experience, source: https://www.sohu.com/a/509078370_99901588, visit on May 6, 2022.

will it deviate from the original intention of signing a virtual equity plan to bind the interests of the company and employees? The answers were not found from the laws and regulations to above problems, which are urgently to be resolved from upper level management of China laws.

In addition, how to define the nature of the dividends from the virtual equity is yet a question. The first paragraph of Article 6 of *the Regulations for the Implementation of the Individual Income Tax Law* states that, "Income from wages and salaries refers to the wages, salaries, bonuses, year-end salary increases, labor dividends, allowances, subsidies and other income related to the employment relations."[18] Paragraph 6 states that, "Income from interest, dividends and bonuses refers to income from interest, dividends and bonuses get from individual creditor's rights and equity."[19]Therefore, the equity dividends to employees without actually holding of the company's equity are virtual equity dividends, which are not income from dividends and bonuses, and are not subject to personal income tax. At this time, if the virtual equity dividends to employees are regarded as their labor income, the individual income tax shall be calculated and paid accordingly. It is worth considering that if it is divided into bonuses, which is regarded as the reward for its labor protected by the Labor Law, is it not in line with the original intention of the company to let employees invest (become shareholders)? It is worth attention that in the case of higher income, the proportion of tax paid according to labor income is higher than that of dividend, which is obviously not a benefit to the effect of incentives. But if this dividend is defined as interest or shareholder dividend, when facing the disputes, should top management represent as creditors or equity holders?

B. Equity Incentives

1. Equity Incentive Overview

The Article 8 of *the Measures for the Administration of Equity Incentive of Listed Companies* regulates the objects of equity incentive. The objects can be BOM, directors, senior managers, core technicians or core businesspersons of listed companies, as well as other employees that the company believes to be encouraged or have a positive impact on the company's business performance and future development, but should not include independent BOM and supervisors. Foreign employees who are BOM, senior managers, core technicians or core business personnel of listed companies can become incentive targets.[20] Article 10 states that "If the incentive objects are BOM and senior managers, the listed company shall establish performance KPI as the conditions for the incentive objects to fulfill their rights and interests."[21]

Equity incentives are not only applicable to large enterprises, but also to small and medium-sized enterprises and family enterprises. When preparing the financial reports, the return on equity should be highlighted, which can help enterprises reduce debts, improve cash flow, increase production and improve shareholder returns. The salary plus equity benefits the transition and empowerment of family business power, reducing the cost of talent can alleviate the principal-agent problem and save cash flow. Let employees enjoy the return of rising stock price, and promote family enterprises development from small to large, and then to go public, can also benefit the unity of long-term and short-term incentives and the development of family

[18] Article 6 of the Regulations for the Implementation of the Individual Income Tax Law of the People's Republic of China.
[19] Article 6 of the Regulations for the Implementation of the Individual Income Tax Law of the People's Republic of China.
[20] Article 8 of the Administrative Measures for Equity Incentive of Listed Companies.
[21] Article 10 of the Administrative Measures for Equity Incentive of Listed Companies.

enterprise culture.

It is to be mentioned that, the equity incentive of family enterprises is not only a property rights system's innovation, but also a profound cultural change. The innovation has promoted the corporate culture reconstruction. For the decision maker, this is a change from family culture to corporate culture. The family culture has strong united power with its irreplaceable loyalty and team spirit together, which has enhanced the family business development. It ensures that the family enterprises adapt to the increasingly fierce market competition environment by attracting more talents for technological innovation, thus promotes the profits growth. For employees, this is a change from work culture to entrepreneurial culture. Employees are allowed to purchase shares, share dividends and return in profits. The incentive mechanism combines the interests between employees and family enterprises. The identity of employees has changed from outsiders to insiders, and from working for the company to working for themselves. They are willing to pay more attention to the long-term development of family enterprises, thus promoting the increase of relevant business income and equity income.

2. Issues in the Practice of Equity Incentive

Article 39 of the *Measures for the Administration of Equity Incentive of Listed Companies* also mentioned that, "Listed companies should hire law firms to issue legal opinions on equity incentive plans and provide professional opinions on the matters."[22] Since the major purpose of equity incentive is to link the interests of employees, enterprise development and shareholders to achieve a win-win situation, the incentive can fully work only by establishing a certain threshold of ownership conditions. If the price of the listed company shows strong deviation from the stock price, there shall be issues from the commercial rationality and interest collaboration. The regulatory authorities should further improve the equity incentive system, raise the price standard of granting equity incentive, by considering the market value of listed companies in the condition assessment, strictly review the rationality of the enterprise equity incentive plan, and well manage the interests of the responsible parties to avoid illegality.

However, in practice, it is found that the family's non-founder shareholders of listed family enterprises are difficult to deeply understand the actual business operation; The conflict will become more obvious after the founder's retirement. Within the time-limited equity incentive mechanism, as a short-term incentive plan for managers, it may potentially conflict

[22] Article 39 of the Measures for the Administration of Equity Incentive of Listed Companies regulates that, listed companies should hire law firms to issue legal opinions on the equity incentive plans, and at least give professional opinions on the following matters: (1) whether the listed company meets the conditions for the implementation of equity incentive stipulated in the Measures; （2） Whether the contents of the equity incentive plan comply with the provisions of the Measures; （3） Whether the procedures for the formulation, review and publicity of the equity incentive plan comply with the provisions of the Measures; （4） Whether the determination of equity incentive objects complies with the provisions of these Measures and relevant laws and regulations; （5） Whether the listed company has fulfilled the obligation of information disclosure in accordance with the relevant requirements of the CSRC; （6） Whether the listed company provides financial assistance to the incentive objects; （7） Whether the equity incentive plan has any situation that obviously damages the interests of the listed company and all shareholders and violates relevant laws and administrative regulations; （8） Whether the directors who are intended to be the incentive objects or the directors who are associated with them have evaded in accordance with the provisions of the Measures; （9） Other matters that should be explained.

with the shareholders' long-term expectation. For example, annual accounting can achieve the goal of reducing enterprise costs by adjusting financial statements and extending the machines' depreciation life. If the professional managers adjust the financial report data, it makes the founders and their families who have retired to the second-tier difficult to understand the real operation of the company from the data, thus leads the misjudgment on the manager's competence. Therefore, to discuss on how to take into account short-term personal interests and long-term value of enterprise development under the premise of employing professional managers to manage the business is a legal and economic topic.

3. Employee Shareholding Platform (Limited Partnership) and Compliance Check

In practice, the legal perspective of the employee stock ownership platform involved in equity incentive is basically limited partnership. This helps to further stabilize the equity structure of the target company and prevent the continuous change due to the employees' resignation. For tax payment, normally to split first and then to be taxed, and the individual tax is paid by the incentive individuals, rather than directly at the shareholding platform.

Before listing, few people paid attention to the compliance of equity incentives. However, as long as the enterprise starts to go public, regulators, brokers and other intermediaries, and even competitors and other parties will pay special attention to the enterprise's compliance issues, and the enterprise's information will be even disclosure. Therefore, as a long-term strategy, equity incentive must be strictly considered by the company before implementation. In order to judge whether the equity incentive is compliant, the aspects that can be considered include but not limited to whether the arrangement of the equity incentive system itself is legal or compliant; whether the implementation of each equity incentive plan is carried out in accordance with relevant laws and regulations; and whether the incentive individual are held on behalf of others. Only after thoughtful consideration of compliance issues (incl. tax), can family enterprises be officially listed.

CONCLUSION

The equity structure and corporate governance system by family enterprises in the early stage can only guarantee the development of the company at the time of establishment. However, while as long as the enterprises are growing, the development has been restricted by initially framework. From the development needs, the management system of enterprises should adapt to its development. The initial ownership structure and corresponding management system have led to the inevitable transformation of many enterprises, also the necessity for equity structure reform. Meanwhile practice has also proved that the long-term effective equity structure is rare to have, and the relevant reform of the system needs to keep updated with the times.

Regardless of whether the family enterprise decides to go public or go public successfully, the ownership structure will have a great impact on its business operation, and the ownership structure problem is also the core issue that cannot be avoided in developing. In order to go public, enterprise managers should adjust their equity structure according to the relevant policies issued by the CSRC, and should optimize the equity structure based on the development objectives of enterprises.

As China's economic development steps into a new era, either from the industrial upgrading perspective, which is the most important thing for the Chinese government, and

constantly cultivating new talents from generation to generation (cultivating non-family business management teams and successors), or from the perspective of common prosperity, corporate wealth invest in social construction constructively, which benefits the people, and creates new sources of social productivity (such as private education), the reform on Chinese family enterprises' ownership structure and related derivative topics will become the highlights of China's economic, social life and regional development. To take initiative, and create institutional Chinese characteristics, to positively face the development issues which are widely concerned by public in the current drastic changes in the domestic and international situation, to resolve social contradictions, to eliminate confusion, to stimulate the vitality of Chinese civil society again, to constantly increase productive competence in the industry, to achieve common prosperity, so as to support the great rejuvenation of the Chinese nation in politics, are of great significance.

CROSSING BOUNDARIES: AUTONOMOUS WEAPON SYSTEMS AND THE CHALLENGE OF IHL COMPLIANCE

Maria Gevorgyan*

Abstract: The introduction of autonomous weapon systems (AWS) marks a transformative juncture in the modern landscape of warfare. Promising operational efficiency, enhanced soldier safety, cost reduction, and workforce minimization, these systems have ignited a global discourse concerning their compliance with international law and the necessity of comprehensive regulation. This study delves into the multifaceted challenges associated with the deployment of AWS within the context of International Humanitarian Law (IHL) while evaluating their alignment with the principles of IHL, especially in terms of direct participation in hostilities (DPH). The research begins by establishing a fundamental understanding of autonomy, delineating the criteria that define AWS, and addressing their legal categorization—whether they constitute a method of warfare or serve as a replacement for combatants. By employing a diverse range of research methodologies encompassing system analysis, comparative legal analysis, synthesis, comparison, analogy, deduction, classification, interviews, and case studies, this study provides a comprehensive examination of the intricate AWS-IHL relationship. Further depth is added to the theoretical analysis through real-world case studies, including the STM Kargu-2 and the United Nations (UN) expert group's involvement in Libya, offering practical insights into the challenges posed by AWS in armed conflicts. Additionally, consideration is given to the SGR-A1, an autonomous system employed for border safeguarding, further illuminating the complexities of AWS in practice. This research aims to provide a nuanced and insightful understanding of the pressing regulatory and challenges arising from the utilization of AWS within the contemporary framework of IHL.

Keywords: Autonomous Weapon Systems; International Humanitarian Law; IHL Principles; Modern Warfare

* Center for Truth and Justice, United Kingdom.

Table of Contents

Introduction 23

I. Methodology 24

II. Discussion 24

A. Understanding an Autonomy 24

B. Defining AWS and Their Classification 25

C. Legal Categorisation of AWS 27

D. Challenges in International Regulation of AWS 29

E. Examining Autonomous Weapons Systems AWS Through the Lens of IHL Principles 30

1. Principle of Distinction 30

2. Direct Participation in Hostilities in the Context of AWS 33

3. The Principle of Proportionality 36

4. The Precautionary Principle 38

5. The Principle of Unnecessary Suffering 39

F. Case study: Current Challenges in the Deployment of AWS in Armed Conflicts (Autonomy in Existing Weapon Systems) 40

1. United Nations Expert Group on Libya: Existing Lethal Autonomous Weapon Systems - A Line Crossed? 40

2. GR-A1 Autonomous Security Robot: Border Safeguard or Severe IHL Violation on the Korean Peninsula 42

III. Results 45

Conclusion 45

INTRODUCTION

The development and deployment of autonomous weapons systems (AWS) represent a significant paradigm shift in the landscape of modern warfare. As militarily developed states continue to invest in the creation of increasingly autonomous weapons, the implications for both the means and methods of warfare, as well as compliance with the law of armed conflict, have become subjects of intense debate and scrutiny. While it is true that certain existing weapons systems have exhibited limited autonomous capabilities, the current trajectory of development suggests a substantial expansion of these capabilities in the near future. It is increasingly unlikely that the deployment of AWS can be halted, and this reality has ignited a profound global conversation.

This prospect has ignited a passionate debate. Human rights organizations are calling for a preemptive ban on the use of autonomous weapons systems, while numerous states are voicing their concerns within international forums, such as the UN. Despite these debates, there exists no specific international legal framework governing AWS. Nevertheless, the motivation for increasing the level of autonomy of weapon systems is compelling, driven by the promise of greater operational efficiency, enhanced safety for one's own soldiers, reduced personnel requirements, and significant cost savings. In June 2022, representatives of the US Defense Ministry underscored the pivotal role of digital transformation and artificial intelligence (AI) in maintaining a competitive edge on the battlefield.[1] Similar efforts in the field of AI and AWS are underway in the armed forces of other nations, including Israel[2] and China.[3] Russia, likewise, has not remained passive, and approximately a year before its another invasion of Ukraine in 2022, it became evident that Russia was expanding its arsenal of weapons equipped with AI capabilities.[4] However, these developments occur without internationally agreed guarantees, including legal ones. Despite advocating for the use of AI in armed conflicts as early as 2021, as reflected in the 2019 Report on AI, the ICRC has expressed[5] profound concerns regarding AWS. Furthermore, the UN Secretary-General articulated in March 2019 that AWS are politically unacceptable, morally repugnant, and should be prohibited by

---

[1] Dave Vergun, *'Digital Transformation, AI Important in Keeping Battlefield Edge, Leaders Say'* (2022) US DEPARTMENT OF DEFENSE'S NEWS, https://www.defense.gov/News/News-Stories/Article/Article/3058028/digital-transformation-ai-important-in-keeping-battlefield-edge-leaders-say/ (last visited 15 September 2023).

[2] S. Biddle, *'Documents reveal advanced AI tools Google is selling to Israel'* (2022) THE INTERCEPT https://theintercept.com/2022/07/24/google-israel-artificial-intelligence-project-nimbus/ (last visited 15 April 2023).

[3] PEOPLE'S REPUBLIC OF CHINA, 'POSITION PAPER ON REGULATING MILITARY APPLICATIONS OF ARTIFICIAL INTELLIGENCE (AI)' (2021) https://www.fmprc.gov.cn/mfa_eng/wjdt_665385/wjzcs/202112/t20211214_10469512.html (last visited 15 September 2023).

[4] Lieutenant General Michael Groen, the Director of the Pentagon's Joint Artificial Intelligence Center, who has been involved in implementing artificial intelligence within the US Department of Defense since 2018, stated, 'The Russian Armed Forces are striving to become a leader in artificial intelligence technologies.' Additionally, CNA, a research organization based in Arlington, Virginia, was commissioned to examine the Russian market. In a report titled 'Artificial Intelligence and Autonomy in Russia,' it is noted that there are over 150 military systems with artificial intelligence in various stages of development. Groen explained that the country aims to utilize AI for electronic warfare, intelligence, surveillance, reconnaissance, and strategic decision-making processes as its leaders seek information dominance on the battlefield. Yasmin Tajde, '*Algorithmic Warfare: Russia Expands its Fleet of Weapons with Artificial Intelligence Support*', NATIONAL DEFENSE (2021) https://www.nationaldefensemagazine.org/articles/2021/7/20/russia-expanding-fleet-of-ai-enabled-weapons (last visited 15 September 2023).

[5] 'ICRC POSITION ON AWS' (2021) https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems (last visited 15 April 2023).

international law.[6] This divergence in viewpoints underscores the urgency of addressing the regulatory and ethical challenges posed by AWS.

Member States of the Convention on Certain Conventional Weapons (CCW) have been engaged in discussions related to these systems for over seven years, and while these deliberations hold the potential for future regulation, concrete steps in that direction have yet to be taken. In the context of deploying AWS, a highly significant yet intricate matter emerges. The core concern revolves around the potential fallibility of machine-based decision-making, which may arise from the inability to account for intricate nuances specific to a particular situation. Human agents are presumed to possess superior capabilities in this regard. Consider, for instance, a scenario in which a drone, armed with explosive ordnance, is dispatched to eliminate an enemy military target positioned in close proximity to a residential building housing civilians. In such a situation, the act of bombing the military facility carries a substantial risk of collateral damage to civilians. When AI governs the drone's operations, numerous pressing questions arise. How effectively can AI navigate the complexities of decision-making? Can the drone ensure adherence to the norms of International Humanitarian Law (IHL) in a manner that strikes a delicate balance between the principles of humanity and military necessity? These represent merely a fraction of the weighty issues inherent in the deployment of AWS.

The primary objective of this study is to explore the multifaceted issues that emerge when deploying AWS within the context of IHL. Moreover, this research will assess the compliance of AWS with the foundational principles of IHL. With a foundational understanding of autonomy, legal classification of AWS according to IHL, and meticulous case studies involving modern AWS, this study aims to provide a comprehensive analysis of the complex relationship between AWS and IHL.

I. METHODOLOGY

To address these complex and multifaceted issues, this research relies on a diverse array of methodological approaches. These include system analysis, comparative legal analysis, synthesis, comparison, analogy, deduction, classification, interviews, monitoring, and case studies. Through this comprehensive approach, we aim to provide a nuanced and insightful understanding of the challenges posed by AWS within the framework of IHL.

II. DISCUSSION

A. Understanding an Autonomy

The initial step in evaluating the legal implications of heightened autonomy in weapon systems is to establish a clear understanding of the technological characteristics underpinning these changes. Only then will the legal significance of these developments become apparent. The technical discussions in this chapter will serve as the foundation for the legal analysis throughout the entire study. In the context of legal analysis, there is no need to delve excessively into technical details. Instead, the primary focus should be on the general possibilities for limiting and utilizing AWS.

[6] ANTÓNIO GUTERRES, 'REMARKS AT WEB SUMMIT' (2018) https://www.un.org/sg/en/content/sg/speeches/2018-11-05/remarks-web-summit (last visited 13 May 2023).

The concept of "autonomy" varies depending on the areas of study and the debates surrounding the use of AWS, which poses a significant challenge. The absence of a unified definition hampers rational discussions of their legal significance. It is crucial to understand that a weapon system, whether in its simplest form, or in a complex form, can be misleading, especially from a legal perspective. The lack of interaction between the machine and the operator during operation does not necessarily mean that the machine's behavior is not determined by a human. Rather, it indicates that the intended behavior was predetermined before the machine's activation and is executed by some part of it, typically through computerized control systems. This control system monitors the weapon system's operation and issues commands as necessary to achieve the desired programmed behavior.

The critical point to emphasize at this stage is that when a "manual" system is replaced by a system capable of a certain degree of autonomy, it may appear as though the control system is effectively supplanting the human operator. This is partially true because the operator's understanding of how to operate the machine is programmed into the control system: the physical means through which the operator manipulates the machine are transformed into a set of actuators that can be activated by the control system itself. Moreover, additional sensors can be incorporated, allowing the machine to process available information based on relatively accurate and dependable environmental perception, in order to develop a meaningful plan and utilize its actuators to set that plan into motion.[7]

Indeed, to an observer, these may seem like highly intricate programs; however, they are merely a collection of predefined instructions, and the machine executes instructions that were pre-written, rather than acting independently. For greater clarity, let's turn to the following practical example. A hypothetical Unmanned Aerial Vehicle (UAV) used for counterterrorism purposes may be equipped with cameras and image recognition software that matches images captured by the cameras to images of known terrorist locations. An instruction such as "if the camera image matches those of known terrorists, behave as if terrorists are present" or some other set of rules by which the UAV can compare the images captured by its cameras does not mean the UAV is making a determination in any real sense as to whether a person is a terrorist; it is still inaccurate to portray the UAV as making independent judgments. Instead, the UAV is simply following instructions provided to it in advance.

Such a lengthy description of technical capabilities aims to eliminate misleading formulations that exist in legal debates, suggesting that AWS have the ability to make real choices during operations. At this stage of development, no computer is capable of independently choosing to execute or not execute a specific instruction in a program.[8] Any such appearance of choice may result from other encoded instructions in the software. Thus, autonomy is the ability of a system to behave as desired and achieve the goals *provided by its operator* without the need for constant external instructions.

B. Defining AWS and Their Classification

In the context of the research, an interview was conducted with Dr. Alex Leveringhaus, a Ph.D. holder in the field of public administration, a research fellow at the Institute of Ethics,

[7] OFFICE OF THE SECRETARY OF DEFENSE, 'UNMANNED AIRCRAFT SYSTEMS ROADMAP, 2005– 2030', US DEPARTMENT OF DEFENSE (2005), https://fas.org/irp/program/collect/uav_roadmap2005.pdf (last visited: 10 October, 2023).

[8] T. MCFARLAND, AWS AND THE LAW OF ARMED CONFLICT, IN AWS AND THE LAW OF ARMED CONFLICT: COMPATIBILITY WITH INTERNATIONAL HUMANITARIAN LAW, 36, (CAMBRIDGE UNIVERSITY PRESS, 2020).

Law, and Armed Conflict at the University of Oxford, and the coordinator of the Special Group on Ethics and AI. Through this interview, it was possible to establish that one of the most challenging questions in nearly all debates concerning AWS is the issue of defining these systems.[9] Various definitions of AWS exist; however, there is no consensus on this issue.[10] A more technical approach to autonomy considers the actual ability of a system to control its behavior and deal with uncertainties.[11] According to this approach, an AWS is a system capable of, based on its perception of the environment, taking the necessary actions to achieve a desired goal. Machines that can adapt to changes in the environment and exercise control over their actions can be characterized as automated or autonomous. Here arises the question: what is the essential difference between these two mentioned systems? Some experts see the difference in the degree of self-governance, considering AWS as more complex, intellectually advanced forms of automated systems.[12] However, there is no definitive answer. There is also an approach that focuses on the command-administrative relationship between humans and AWS.

In the context of this research, the focus has deliberately shifted away from treating autonomy as a general attribute of AWS. This is because such a broad approach can be misleading and lead to intricate debates concerning the threshold for considering a weapon system as autonomous. Instead, the contention is that issues, particularly those with legal implications, should be addressed by considering the specific functions or tasks for which autonomy is employed. Autonomy is most effectively analyzed by categorizing it according to the functions performed at the level of an AWS.[13] A primary advantage of adopting this approach lies in its flexibility for investigating issues related to AWS. For instance, the functional approach allows for acknowledging that the extent of human interaction, including operator control, varies between different functions within AWS. Some functions may necessitate a higher degree of autonomy, while human control may be retained for others or relinquished entirely. Furthermore, the level of human operator involvement can fluctuate based on the mission at hand. Consequently, this approach implies that the concept of "AWS" is a comprehensive term encompassing a wide array of weaponry with autonomy integrated into their critical functions. This includes weapons capable of autonomously selecting (searching, identifying, tracking) and engaging (applying force to) targets without continuous human intervention.[14]

Thus, an AWS is a weapon that, once activated, can select and engage targets without further *constant* intervention by a human operator.[15] Furthermore, within this approach, the following types of AWS can be identified.

---

[9] Dr. Alex Leveringhaus, a Ph.D. holder in the field of public administration, a research fellow at the Institute of Ethics, Law, and Armed Conflict at the University of Oxford, the coordinator of the Special Group on Ethics and AI, personal interview 02.07.23.

[10] P. Scharre, 'Autonomous Weapons and Operational Risk', 16, Center for a New American Security, Ethical Autonomy Project, (2016).

[11] S. Thrun, 'Toward a Framework for Human-Robot Interaction', 9-24 (2004) 19(1–2) Human–Computer Interaction.

[12] D. Mindell, *Our Robots, Ourselves: Robotics and the Myths of Autonomy*, 12 (2015)Viking: New York,.

[13] United Nations Institute for Disarmament Research (UNIDIR), Framing Discussions on the Weaponization of Increasingly Autonomous Technologies, UNIDIR Resources No. 1 (UNIDIR: Geneva, 2014).

[14] 'AWS: Is it Morally Acceptable for a Machine to Make Life and Death Decisions?', 13-15, ICRC, (CCW Meeting of Experts on Lethal AWS, (Geneva, April 2015).

[15] P. Scharre, 'Where Does the Human Belong in the Loop?', 4, CCW Meeting of Experts on LAWS: Technical Issues, (May 2014).

1. Semi-autonomous system - human in the loop. In this configuration, the system is programmed to await input from a human operator before taking action.

2. Supervised autonomous system - human on the loop. In this case, the program allows for human intervention but does not require real-time and mandatory human involvement, as is the case with the aforementioned autonomous systems.

3. Fully autonomous system - human out of the loop. In this scenario, the system is programmed in a way that does not permit real-time human intervention.

C. Legal Categorisation of AWS

The subject of debates has also revolved around the question of the legal category to which AWS belong. Some participants in the discussions argue that AWS occupy a position between weaponry and combatants, which in turn raises questions about their ability to adhere to the norms of IHL. The central idea behind this approach is as follows: when a significant portion of the "targeting" process is encoded within the weapon system, the AWS assumes the responsibilities of a soldier, acting as a kind of delegate of the combatant or an artificial surrogate for the combatant. The concept of "AWS" often appears in formulations that seemingly position these systems as bearers of obligations under IHL.[16] In the context of this research, arguments of this nature, which imply the replacement of humans by AWS in operational and possibly legal terms, do not consider AWS as tools used by humans. However, within the scope of this study, autonomy implies a form of control rather than its absence, which is why the aforementioned positions do not appear sufficiently substantiated. Let us attempt to analyze the legal validity of the position presented above and address the question: is it appropriate to classify AWS as a means of warfare for legal purposes, as envisaged by IHL? The argument that the ability of AWS to perform tasks traditionally assigned to combatants justifies categorizing them into a legal category distinct from the means of warfare does not appear well-founded.

The provisions of Protocol Additional I to the Geneva Conventions 1949 (API) "Believing it necessary nevertheless to reaffirm and develop the provisions protecting the victims of armed conflicts and to supplement measures intended to reinforce their application," as well as the provisions related to weaponry, imply that terms should be interpreted expansively to ensure such protection. This is also confirmed in the Commentary of the ICRC to Article 35 of AP I, the provisions of which affirm the principle that the right of parties to the conflict to choose methods or means of warfare is not unlimited, and that means of warfare encompass weapons in the widest sense.[17] Article 36 of API refers to the scope of application of the weapons review mechanism, thereby indicating the need to adhere to a broad concept of weaponry, which is supported by state practice.

---

[16] HUMAN RIGHTS WATCH, LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS (2012) 30; C. HEYNS, REPORT OF THE SPECIAL RAPPORTEUR ON EXTRAJUDICIAL, SUMMARY OR ARBITRARY EXECUTIONS, HUMAN RIGHTS COUNCIL, 23RD SESS, 5-6 [28], AGENDA ITEM 3, UN DOC A/HRC/23/47 (9 April 2013).

[17] COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949, (1987), 130, INTERNATIONAL REVIEW OF THE RED CROSS (1961 - 1997).

In accordance with the concept of a broad understanding of weaponry, substantiated by state practices, the Working Group of the U.S. Department of Defense on Military Law defines "weaponry" as encompassing all armaments, munitions, material components, mechanisms, or devices intended to have the presumed effect of causing injuries, damage, destruction, or incapacitation of personnel or equipment. Furthermore, the same Working Group defines a "weapon system" as the weapon itself and the components required for its operation, including new advanced or emerging technologies that may lead to the development of weapons or weapon systems with significant legal and political implications.[18] In turn, the Australian Department of Defense specifies that the concept of "weaponry" encompasses weapon systems, ammunition, submunitions, guidance devices, and other destructive mechanisms.[19]

These definitions and characteristics apply to the components, functions, and effects typically possessed by AWS, thereby allowing them to be considered as means of warfare, especially given the need for a broad interpretation. It should also be acknowledged that, despite the presented technical anthropomorphic designs, the physical composition and form of AWS do not significantly differ from other types of weaponry. Indeed, the software distinguishes itself with a range of capabilities, but it remains software developed by a human programmer, with a design fundamentally similar to that used in other armaments.

The examination of the functional aspects of AWS should not influence a change in the legal categorization of AWS as a means of warfare. Regarding input data, as previously noted, their role remains to receive a command from a human source and execute it, with the nature of commands varying as the autonomy of the weapon system changes. It is crucial to understand that the enhanced capabilities of AWS may lead to significant changes in the conduct of military operations, but they do not alter the legal category. In other words, there is no causal relationship between the function performed by an AWS and its legal categorization under IHL.

Consider another argument put forth by proponents of the concept that AWS are something more than just weapons. It is argued that the ability of a system to gather data about the surrounding world and use it to formulate high-level commands, which were not explicitly given, serves as yet another example of the common human tendency towards anthropomorphism.[20]Anthropomorphic concepts of autonomy bring to the forefront one or more human-like qualities or models of behavior that autonomous systems can exhibit. Definitions related to operational independence of weapon systems are among the most common. In technical terms, "independence" implies the absence of a need for explicit instructions from an operator, as previously explained; these instructions are pre-encoded in the weapon system's control software. However, this factor is often distorted to support legal arguments and claims that it signifies the independence of actions of AWS from humans.[21] In the legal context, arguments equating the role of a soldier with that of an AWS ignore a crucial fact: weapons and soldiers are distinct legal categories, despite potential functional overlap, which the author of this study does not acknowledge. Unlike weapons, computers, software,

---

[18] DEPARTMENT OF DEFENSE DICTIONARY OF MILITARY AND ASSOCIATED TERMS (12 April 2001), https://irp.fas.org/doddir/dod/jp1_02-april2010.pdf (last visited: 10 October 2023).

[19] 'LEGAL REVIEW OF NEW WEAPONS' (DEFENCE INSTRUCTION (GENERAL) OPS 44-1, AUSTRALIAN DEPARTMENT OF DEFENCE, (2 June 2005) sub-s 3(a).

[20] S. FUSSELL, 'HOW PEOPLE ANTHROPOMORPHIZE ROBOTS', 145, (2008) 3RD ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN ROBOT INTERACTION PROCEEDINGS.

[21] *'Mind the Gap: The Lack of Accountability for Killer Robots',* HUMAN RIGHTS WATCH, (April 2015), https://www.hrw.org/sites/default/files/report_pdf/arms0415_summary_mindthegap.pdf (last visited: 10 October 2023).

'combatant' is a legal category encompassing only humans, who, in turn, possess rights and responsibilities that underlie IHL. Positioning AWS as combatants in a legal sense would grant them the role of exercising primary control over decisions and committing acts of violence. However, the analysis of autonomous weapon technology presented above reveals that any such assertion of independence from human control is an illusion. AWS, like other forms of weaponry, differ from combatants, and drawing a legal analogy between these two categories appears fundamentally untenable and irrational. Neither existing legislation nor political considerations support the assumption that legal personhood should extend to artifacts, such as AWS.[22] In legal terms, AWS should be regarded as means of warfare, with humans and the states deploying and operating them being the bearers of legal obligations.

The question also necessitates an answer: can AWS be legally regarded as a method of warfare? Given the inherently fluid nature of machine autonomy and the myriad of ways it can be employed, a sufficiently expansive interpretation of the latter concept will be required to classify AWS as methods of warfare. A more plausible proposition is that the specific behavior exhibited by a particular AWS might qualify as a method for conducting military operations. Nevertheless, even in such cases, it is imperative to acknowledge that any weapon possesses distinct behavior patterns that should be construed as integral components of the means employed in the conduct of military operations rather than constituting distinct methods. For instance, the behavior of a mine's detonation mechanism represents a distinguishing feature of that particular means for conducting military operations. The software underpinning AWS, despite its considerably enhanced complexity, fundamentally adheres to this principle. Consequently, it is paramount to recognize that the concept of a method of warfare encompasses the manner in which weaponry is deployed (i.e., the interactions between a human operator and the weapon), rather than the intrinsic behavior of the weapon itself. Therefore, even if the pre-programmed behavior embedded within an AWS emulates actions that a human could undertake, it remains an integral aspect of the means employed in the conduct of military operations. However, it is essential to note that the aforementioned does not imply that the utilization of all categories of AWS inherently constitutes a means for conducting military operations.

D. Challenges in International Regulation of AWS

The participating states of the CCW are currently engaged in a series of discussions concerning issues related to AWS. One possible outcome of this process could be the establishment of a specialized international legal framework for AWS. However, at present, no concrete steps have been taken in this direction. Due to the absence of specialized international legal regulations for AWS, their design and use are governed by general conventional and customary norms of IHL that regulate weapons not subject to separate control.

Nevertheless, the legal implications of introducing specific legal regulations for AWS, which may contain restrictive or prohibitive provisions, are of particular interest. In most cases, rules that restrict or prohibit certain types of weapons define the regulated weaponry based on its inherent nature or the effect it has on the objects against which it is used, including the equipment used in conjunction with the weapons, such as delivery mechanisms. Thus, the introduction of autonomy is unlikely to significantly impact the applicable law.

---

[22] M. SASSÒLI, 'AUTONOMOUS WEAPONS AND INTERNATIONAL HUMANITARIAN LAW: ADVANTAGES, OPEN TECHNICAL QUESTIONS AND LEGAL ISSUES TO BE CLARIFIED', 308, 323, (2014) 90 INTERNATIONAL LAW STUDIES.

In examining the approach used in The Biological Weapons Convention (BWC), participants commit to never, under any circumstances, develop, produce, stockpile, acquire, or otherwise retain microbiological or other biological agents or toxins or weapons, equipment, or means of delivery intended for the use of such agents or toxins.[23] These prohibitions are not related to any properties that could be altered through the combination of biological weapons with autonomous control. Therefore, at first glance, the utilization of autonomy as a means of control over biological weapons would similarly be prohibited by the BWC. Comparable instances can be found by referencing the Chemical Weapons Convention (CWC) and the Ottawa Convention. Thus, the applicable law concerning AWS at present and in the near future is confined to general conventional norms, as well as customary IHL relating to weaponry and its use.

E. Examining Autonomous Weapons Systems AWS Through the Lens of IHL Principles

In the absence of any international legal regulation for AWS, this study has chosen to examine the application of AWS through the lens of IHL principles. The optimal approach is to peel back the layers of laws and assess the compliance of AWS with the fundamental principles of IHL, as these principles play a foundational role and can serve as a guide in addressing complex issues. It should be noted that the principles of IHL are not fixed in a single specific source, so the first step is to identify them. The International Court lists four core principles of IHL: the principle of distinction,[24] the principle of military necessity,[25] the principle of avoiding unnecessary suffering,[26] and the principle of proportionality.[27] However, scholars propose various sets of principles. For instance, C. Droege identifies among the fundamental principles of IHL: the principle of distinction, the principle of proportionality, and the principle of taking precautionary measures.[28]

1. Principle of Distinction

The principle of distinction consists of two components: parties in an armed conflict must distinguish between civilians and combatants, as well as between civilian and military objects. The target of an attack may only be military objects and combatants.[29] This principle is codified in Article 48 of Additional Protocol I (API) For states that have not ratified API, this principle is a customary rule of IHL. Additionally, the International Court has ruled that this norm is a cardinal principle of IHL.[30]

At first glance, the principle seems to be a relatively straightforward black-and-white rule: a potential target is either military or it is not. However, difficulties arise because whether a target is qualified as civilian or military can depend on the context. Indeed, the analysis

[23] Biological Weapons Convention, opened for signature 10 April 1972, 1015 UNTS 163 (entered into force 26 March 1975) art 1.
[24] Nuclear Weapons Advisory Opinion, 1996, ICJ Rep 226 paras 78, 92 and 95.
[25] ibd, paras 30, 32, 43 and 48.
[26] ibd, paras 77, 78 and 95.
[27] ibd, paras 41, 44 and 46.
[28] C. Droege, *'Get Off My Cloud: Cyber Warfare, International Humanitarian Law, and the Protection of Civilians'*, 533-553 (2012) 94(886) INTERNATIONAL REVIEW OF THE RED CROSS.
[29] Protocol I Additional to the Geneva Conventions, Arts 48 and 51, especially 51(4). ICRC Study on Customary International Humanitarian Law, Volume II, Chapter 24, Section B.
[30] Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion), 76, ICJ Reports (1996).

typically required to adhere to the principle of distinction is highly complex and highly contextual. Granting weaponry, the ability to choose its own targets means that part of the weapon targeting process is removed from the hands of a human operator and encoded in the control system of the AWS. It should be noted that this principle implies a simple dichotomy: a weapon is either capable of being directed or not directed against a specific military target in the circumstances of a particular attack. Thus, the legal assessment of AWS requires a broader understanding of "precision," which includes the elements of the system's ability to select lawful targets in addition to factors that typically characterize the "precision" of a weapon. This approach may seem illogical, especially to proponents of an anthropomorphic approach (see above). Nonetheless, this approach aligns with the paradigm chosen in this study, where autonomy is considered a form of control. The principle requires that the combination of the operator's actions and the behavior of the weapon system result in the identification of a lawful target and the realization of conditions under which it can be legally attacked. The weapon system must then behave in a way that ensures the attack is directed at the chosen target, which can be legally attacked. When a significant portion of the target selection process is encoded in the weapon control system, the distribution of tasks between the operator and the weapon system changes.

Consider the measure of "precision" in the specific case of an anti-missile defense system. The weapon operator applies a "time and place" restriction to the set of potential targets that can be set. Once this external limit is established, control shifts to the weapon system itself, which then aims and fires. Thus, the overall precision of the weapon system has two components: the precision with which it identifies that a potential target falls within the established set of targets it is programmed for and the precision with which it can actually engage that target. The same logic can be applied to more AWS. The fact that an AWS has greater operational freedom, such as increased range or adaptability, does not preclude its actions from remaining limited and tied to the point of activation; the system will only select targets within its operational range. The commitment of persons conducting an attack to target only military objectives remains unchanged. However, the share of the AWS in the fulfillment of this commitment increases. As long as it is possible for the operator to restrict the set of targets available to the AWS in such a way that there is sufficient certainty that only lawful targets will be engaged, the AWS will satisfy the threshold requirement of Article 51(4)(b) of the API. Other questions that require answers include: How effectively can the targeting system identify lawful targets under attack conditions? How well can the targeting system distinguish combatants from civilians? How accurately can the chosen target be engaged? Given the achievement of a sufficient level of accuracy in target identification and the operation of the AWS, it will not have uncontrolled effects, and thus legal requirements will be met. The perspective that certain obligations related to the "law of targeting," such as the requirement to take all feasible precautions to verify that targets are lawful, might be absorbed by the functions of the AWS may intuitively appear attractive. However, this concept is flawed because it implies that the software should do everything that is practically possible in the circumstances of the attack.[31] These assertions are components of an anthropomorphic concept and are fundamentally incorrect, as legal obligations cannot automatically transfer from humans to weapons. Obligations related to precautionary measures remain exclusively with those who plan or make decisions about an attack. It is they who must do everything possible to ensure

[31] C. PILLOUD, J. PICTET, 'PROTOCOL I - ARTICLE 57 - PRECAUTIONS IN ATTACK' IN YVES SANDOZ, (C. Swinarski, B. Zimmermann eds), Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949 (Martinus Nijhoff, 1987) n 43 in Chapter 2.

that the targeting functions of AWS operate appropriately. Any targeting functions encoded in AWS must conform to the requirements of the principle of distinction.

Opponents of AWS often cite (hypothetical) concerns related to attacks on protected persons, describing such incidents as indiscriminate attacks.[32] For example, it is difficult to distinguish a farmer tilling the soil from a member of an armed group laying improvised explosive devices. Disagreements persist regarding the precise circumstances under which deadly force can lawfully be used against civilians who are in some way connected to an armed conflict. Indeed, such a problem related to the principle of distinction exists. The opposing side, in turn, may use civilian objects enjoying special protection under IHL, such as hospitals or mosques and churches, for cover. But does the firing of an AWS, for instance, at a civilian necessarily equate to a violation of the principle of distinction? There are several factors that could lead AWS to cause harm to individuals or objects enjoying protection under IHL. It should be noted that some of these factors are common to other weapon systems, while others are specific to the characteristics of AWS. The legal character of harm inflicted on persons or objects enjoying protection under IHL depends on the grounds. Possible grounds can be classified by examining the components of an attack by an AWS. Firstly, there is the human operator who activates the weapon system. Secondly, the guidance system, which is part of the control system of the AWS. Thirdly, the weapon itself. Incidents related to the actions of the operator or the behavior of the weapon have the same legal character as the use of "manual" weapons as well as AWS. Therefore, it is not necessary to consider situations in which the operator intentionally conducts an attack on an unlawful target, which would violate Article 51(4)(a) of the API, or activates the weapon system in circumstances for which it is not intended, or the AWS simply misses and fails to hit the target due to limited accuracy, which could

One of the primary concerns related to AWS is the possibility of system malfunctions. An AWS, experiencing accidental software failure, may attack unlawful targets. It is not evident that such a failure would have any legal consequences beyond those associated with an equivalent software or hardware malfunction in "manual" weapon systems. Another issue is the potential for the "intentional" opening of fire by an AWS on an unlawful target. In this case, the AWS conducts an attack on an unlawful target, but unlike the previous example, in this case, the system's targeting code executes without errors, albeit not in line with the operator's intent. Presumably, there are two possible reasons for such an outcome: intentional and unintentional misconfiguration. In the first case, the AWS may be maliciously programmed to target civilians. The legal character of such a scenario does not differ from a similar act committed using any weapon relying on a targeting system. Depending on the responsible party, these actions may be classified as acts of sabotage or indiscriminate attacks under Article 51(4)(a) of the API. In the case of unintentional misconfiguration, the developer incorrectly configured the AWS in a way that its targeting system identifies a civilian as a legitimate target. Such errors are not inherent to AWS but are human errors.[33] Failure can be attributed to an individual, the programmed weapon system, or individuals who failed to detect an error during the weapon's testing process. The executable code for determining a lawful target ultimately reflects the decision-making process of a human. Any failure leading to an AWS opening fire on unlawful targets is equivalent to a failure that can occur with any other type of weaponry, so AWS are not an exception. At a technical level, there are no failure modes that are unique to AWS or cannot be regulated by the law applicable to AWS. Malevolent actions by a weapon

[32] ibd, 684.
[33] R. McLaughlin, *'Unmanned Naval Vehicles at Sea: USVs, UUVs, and the Adequacy of the Law'*, 105, (2011) 21(2) JOURNAL OF LAW, INFORMATION AND SCIENCE.

developer, in turn, do not alter their legal character simply because they are carried out using an AWS rather than another type of weaponry.

The second category includes inherently unlawful weapons, which are those types of weaponry whose consequences cannot be controlled well enough to limit their use solely for military purposes. API describes such weapons as those that employ methods of warfare whose effects cannot be confined within the requirements of the Protocol. While the first category of unlawful weapons is prohibited due to the inability to precisely direct their effects towards legitimate targets, this type of weaponry is prohibited because it is impossible to subsequently restrict the impact of these weapons to lawful targets. The reservation 'as required by the Protocol' deserves special attention: this wording indicates that the mentioned effects do not necessarily result from the direct impact of the weapon but can also include any effects that raise concerns in accordance with other articles of Additional Protocol I. These include Articles 35(3) and 55, which prohibit 'widespread, long-term, and severe damage to the natural environment,' suggesting that it pertains not only to the direct effects of the weapon.

It is worth noting that the prohibition of weapons whose effects cannot be confined is only loosely related to autonomy and is not a subject of discussion in the context of AWS as a class of weaponry. This prohibition is more closely associated with the physical means by which the weapon component of the system inflicts harm, such as the type of ammunition.[34] As such, AWS must possess advanced skills in observation and recognition, as well as the developed ability to make judgments because people make distinctions, taking into account various factors, such as raising hands in the air or signs of an enemy becoming incapacitated.

2. Direct Participation in Hostilities in the Context of AWS

As demonstrated by the analysis presented above, the ability of AWS to adhere to the principle of distinction is highly questionable. "Attacks shall be limited strictly to military objectives. In so far as objects are concerned, military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage.[35]"

With regard to individual persons, context can be important for the following reasons. In the context of modern warfare (as mentioned above), an individual may transition from being a combatant, who is a legitimate target, to a person protected under IHL, and vice versa. For example, an individual may be dressed in camouflage attire with various distinctive markings, carrying a rifle in the midst of combat. Based solely on observation and recognition capabilities, such an individual might be identified as a participant in hostilities. However, upon further analysis of the entire situation, taking the context into account, it may become evident that this person is not a legitimate target, for instance, having been incapacitated due to injury or illness.[36] Indeed, not only clothing and distinctive markings but also contextual factors are of significant importance in determining an individual's status under IHL.

---

[34] I. HENDERSON, 'THE CONTEMPORARY LAW OF TARGETING (2009), 237, ch 1, para 14, C. PILLOUD, J. PICTET, 'PROTOCOL I – ARTICLE 57 – PRECAUTIONS IN ATTACK' IN YVES SANDOZ, C. SWINARSKI, B. ZIMMERMANN (EDS), COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949, 686, (1987) n 43 in ch 2.

[35] Additional Protocol I to the Geneva Conventions, Art 52.

[36] Additional Protocol I to the Geneva Conventions, Art 41(2).

Measures to protect individuals who are not or are no longer participating in hostilities rely on combatants' capacity to make reasoned judgments about contextual factors. For example, recognizing the raising of hands as a sign of surrender or identifying an enemy in a state of unconsciousness or disorientation, indicating their incapacity to continue fighting. Conversely, a situation can also arise where an individual transitions from being a person protected by IHL to a combatant, simultaneously becoming a legitimate target. IHL generally defines civilians negatively, meaning anyone who is not a combatant is considered a civilian.[37] Furthermore, IHL provides an additional safeguard for the protection of civilians: in cases of doubt, an individual is considered a civilian.[38] Under certain circumstances, civilians may lose the protection afforded to them by international humanitarian law, including the Geneva Conventions. For example, in cases of "levee en masse," the status of a prisoner of war, and hence a participant in the conflict, is extended to residents of unoccupied territory who voluntarily take up arms to resist invading forces without having had time to form regular armed forces. Today, the notion of "direct participation in hostilities" (DPH) is actively used.

It would be overly optimistic to think that there is a rule, however complex, that can definitively categorize every individual in terms of IHL. Let's demonstrate this with the example of determining a "civilian taking direct part in hostilities." The ICRC has developed a carefully thought-out guide defining what constitutes an act of direct participation in hostilities, based on which civilians do not receive the protection typically afforded to civilians under IHL. For a civilian to become a lawful target, the following criteria must be met: 1) the threshold of harm; 2) a causal link between the act and the harm; 3) affiliation with a party to the conflict.

To meet the threshold of harm necessary for an act to qualify as direct participation in hostilities, it must, with a high degree of probability, have a negative impact on military operations or the military capabilities of the opposing party in an armed conflict. In the absence of military harm, the harm threshold can also be reached when the act may lead to the death, destruction, or injury of individuals or objects protected against direct attack. In both cases, actions that meet the required harm threshold can be considered as direct participation in hostilities only if they additionally satisfy the criteria of a causal link between the act and the harm, as well as an affiliation with a party to the conflict (as mentioned above). For the causal link requirement to be met, it must be possible to establish that either the specific military operation or the specific action, as part of which the harm occurred, is directly causally linked to the harm inflicted, which reaches the required threshold. However, actions meeting the causality and harm thresholds can be considered as direct participation in hostilities only if they also satisfy the criterion of affiliation with a party to the conflict. To meet this requirement, the action must be specifically designed to cause the required harm in support of one of the parties to the armed conflict to the detriment of the other party. Note that certain actions, while causing harm that reaches the required threshold, may lack this specific affiliation. For example, harm caused in individual self-defense, protecting others from violence prohibited under IHL, exercising authority over individuals and territory, and addressing civil disturbances and internal violence may lack this specific affiliation.

When all these criteria are applied in combination, it allows us to distinguish between actions that are considered equivalent to direct participation in hostilities and actions that are not part of the behavior of armed forces, even though they occur within the context of armed

[37] Additional Protocol I to the Geneva Conventions, Art 50(1).
[38] Third Geneva Convention, Art 4A(6).

conflict. However, even in cases where a specific action is the equivalent of direct participation in hostilities, the type and degree of force used in response must comply with the norms and principles of IHL.[39] These guiding principles are an attempt to provide the means to determine whether an individual in specific circumstances meets each of these requirements. It requires a thorough understanding of the complex situation, including the strategic consequences of potential harm, the status of individuals facing potential harm, sociocultural and psychological indicators in which the intentions and actions of this individual are classified as military rather than, for instance, personal self-defense (see above). In other words, AWS must be able to perceive all the necessary information while taking into account the contextual element. For example, distinguishing a civilian holding a large piece of metal from an armed combatant with a rifle in civilian clothing.

The International Criminal Tribunal for the former Yugoslavia (ICTY), in its decision,[40] provided a definition of what behavior can be qualified as direct participation in hostilities. While this list raises more questions than it provides answers, it is worthy of attention. For example, examples of direct participation include not only carrying or using weapons during involvement in armed conflict or operations but also engaging in attacks on the personnel, property, or equipment of the opposing armed forces, transmitting military information for immediate use, transporting weapons in close proximity to combat operations, conducting reconnaissance, and observing on behalf of armed forces.[41]

These examples illustrate the complexity and nuances involved in determining what constitutes direct participation in hostilities. Such determination often depends on the specific context and factual circumstances of each case. AWS, if involved in making such determinations, must be capable of assessing these complex situations, considering both actions and intent, and evaluating them within the broader framework of IHL. The interpretation of the rules established by the ICTY also requires a higher-level judgment. For example, judgment will be needed to determine whether the military information was transmitted for "immediate use" (qualifying as DPH) or merely for use in the more distant future (which does not qualify as DPH). Similarly, judgment will be required to decide whether weapons were transported "in the immediate vicinity of hostilities" or further away from the theater of operations, which also affects the classification of actions: actions will be classified as direct participation in hostilities only in the first case.

In order to comply with the principle of distinction using AWS, it is necessary for these systems to possess cognitive abilities inherent to humans, which allow them to assess and analyze contextual factors, as mentioned earlier. AI capable of such abilities currently exists only in theory and is referred to as General AI. At the moment, only narrow AI is available, which is also used in AWS. Narrow AI is capable of automatically performing very specific, point-specific functions or tasks at a human level or even exceeding it in some areas.

It is essential not to equate narrow AI with general AI since the process of task-solving is fundamentally different. General AI is capable of not only self-learning but also making its conclusions based on received information and environmental conditions, considering context, and even having a form of self-awareness and self-improvement. General AI can solve complex

[39] INTERPRETIVE GUIDANCE ON THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES UNDER INTERNATIONAL HUMANITARIAN LAW, ICRC, (Geneva, 2009).
[40] Prosecutor v Pavle Strugar (Appeal Judgment), International Criminal Tribunal for the Former Yugoslavia, Case No ICTY-01-42, 17 July 2008.
[41] ibid, para 77.

composite tasks, including recognizing human emotions. This can be useful in field conditions, for instance, to determine whether a person has been incapacitated or whether they are surrendering. To make a definitive judgment, general AI incorporated into autonomous weapons systems would need to process data on facial expressions, posture, spoken words, and so on.

Despite the fact that General AI is yet to become a reality, it should be noted that this field is advancing rapidly. However, there are significant discrepancies in the scientific community regarding predictions and timelines for the realization of general AI. According to experts in artificial intelligence, it's unlikely that General AI will emerge in the near future[42]. This raises the question of whether AWS can provide the necessary level of judgment for the principle of distinction at the present moment, especially in the context of the concept of direct participation in hostilities. The absence of AGI does not necessarily preclude this capability. While the development of full AGI may take a while, the artificial intelligence currently in use (Narrow AI) in AWS may still be capable of making distinctions. However, this capability is limited, and such systems should ideally be employed in specific settings, as mentioned earlier, such as exclusively for defensive operations and in areas where the presence of civilians is highly unlikely.

3. The Principle of Proportionality

Despite the fact that the principles of IHL are primarily oriented towards the protection of civilian populations, it is essential to maintain a realistic perspective on armed conflicts. In modern conditions, there are no effective means to completely eliminate the possibility of civilian casualties and injuries. This principle is established both in Article 51 of AP I and in customary IHL.

The Principle of Proportionality entails that in cases where harm is inflicted on civilian individuals, such harm should be commensurate with the military advantage gained. In essence, an attack becomes unlawful when the incidental harm to "civilian" persons is excessive as a result of the attack. IHL dictates that the use of force and the means employed must always be proportionate to the military advantage sought.[43] In the context of military advantage, it should be specific and direct, rather than abstract. Specific and direct nature of the advantage indicates that it should be significant and relatively immediate, and military advantages that may arise in the long term should be ignored.[44]

In the context of proportionality, it requires a contextual balancing of two factors: the potential harm to civilians and civilian objects on one hand, and the expected military advantage on the other. The ICTY established criminal liability for disproportionate attacks, depending on whether a person was well-informed, used available information reasonably, and expected excessive harm to civilians and objects as a result of the attack.[45] It is essential to note that the concept that there is no formula for balancing these factors is not endorsed in this study. Proportionality is not a vague notion but rather a clear directive, setting a "fixed standard" on the constraints of what commanders and soldiers can do, eliminating undesirable freedom of

---

[42] V. MULLER, N. BOSTROM, 'FUTURE PROGRESS IN ARTIFICIAL INTELLIGENCE: A SURVEY OF EXPERT OPINION' IN VINCENT (C. Muller ed), FUNDAMENTAL ISSUES OF ARTIFICIAL INTELLIGENCE (Springer, 2016).
[43] Article 51(5)(b) Additional Protocol I to the Geneva Conventions. ICRC Study on Customary International Humanitarian Law, Rule 14.
[44] J. HENCKAERTS, ICRC STUDY ON CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, 49, (1996).
[45] Prosecutor v Galic (Judgment), (International Criminal Tribunal for the Former Yugoslavia, Case No ICTY-98-29, 5 December 2003) Para 808.

action. Since assessing proportionality involves weighing competing interests, AWS must be able to anticipate the consequences of all potential decisions and the potential number of civilian casualties. They must also be responsive to changing circumstances and, subsequently, be able to calculate the military advantage and determine whether the incidental harm is acceptable.[46]

Here we will address the issue of the "frame": in order to calculate the potential collateral damage resulting from an attack by an AWS, it is necessary either to calculate the consequences of every possible action, which would take an infinite amount of time, or to make assumptions, which could potentially lead to a disproportionate attack and, consequently, a violation of IHL. Determining collateral damage is associated with assumptions because, in the conditions of armed conflict, exact certainty is rarely achievable. In cases where AWS are used in open civilian areas, due care must be taken to ensure that the information on which the assumptions necessary to determine collateral damage are based is sufficient and reliable.

Despite the prevailing scientific position[47] that the implementation of the principle of proportionality requires more than just balancing quantitative data and that "a robot cannot be programmed to replicate the psychological processes in human judgment necessary for assessing proportionality," this study does not share that view. The principle of proportionality demands a quantitative calculation, albeit a rather complex one. In this study, proportionality is examined through the lens of utilitarianism. Utilitarianism allows for the translation of ethical aspects into a specific observable dimension by replacing normative categories with observable outcomes and recognizing the moral significance of actions' consequences based on the criteria of increasing so-called "happiness." Just as the principle of utilitarianism is satisfied when an action leads to more "happiness" than "unhappiness," proportionality is satisfied when an attack results in greater military advantage than collateral damage. Utilitarianism allows for the translation of ethical aspects into a specific observable context, achieved by substituting the categories of the morally required with observable objectives. It also recognizes the moral significance of the consequences of actions, based on criteria aimed at increasing what is commonly referred to as "happiness." Therefore, just as the principle of utilitarianism is adhered to when an action leads to greater "happiness" than "unhappiness," proportionality is satisfied when an attack results in greater military advantage than accompanying harm.[48]

While maintaining the principle of proportionality, AWS must also have the capability to weigh military advantage against collateral damage. A proposed starting point for this evaluation is the methodology for assessing collateral damage estimates (CDEM[49]). A five-stage analytical system is used for assessing collateral damage, based on factors such as the area of effect of different weapon types, demographic data in the anticipated strike area, the timing of the attack and its potential impact on the likely level of civilian casualties.[50]

---

[46] M. WAGNER, 'TAKING HUMANS OUT OF THE LOOP', 159-162, (2011).
[47] J. PETMAN, 'AWS AND IHL: "OUT OF THE LOOP?"', 39, (2017).
[48] E. Winter, *'Autonomous Weapons in Humanitarian Law: Understanding the Technology, Its Compliance with the Principle of Proportionality and the Role of Utilitarianism', 183,194,* (2018) 6(1) GRONINGEN JOURNAL OF INTERNATIONAL LAW.
[49] CDEM - Collateral Damage Estimation Methodology, is a methodology developed by the United States. It is used to identify any "collateral concerns" (i.e., civilians, civilian objects, or other protected entities) within the radius of action of a weapon under consideration for use in an attack.
[50] E. Winter, *'The Compatibility of the Use of Autonomous Weapons with the Principle of Precaution in the Law of Armed Conflict'*, 240, 262, (2020) 58(2) MILITARY LAW AND THE LAW OF WAR REVIEW.

Through this methodology, utilitarianism, as originally defined by Bentham, can be applied to real values that can be used in proportionality calculations. In this approach, proportionality is assessed through a utilitarian framework, which translates ethical aspects into observable terms, focusing on the consequences of actions, based on criteria that increase "happiness." Just as utilitarianism is adhered to when an action leads to greater "happiness" than "unhappiness," proportionality is met when an attack leads to a greater military advantage than collateral damage.[51]

The key challenge here is to weigh factors that may seem incomparable. The study suggests quantifying collateral damage in terms of lives lost or injuries sustained (possibly utilizing CDEM, as mentioned earlier) and then calculating military advantage in terms of lives saved or injuries prevented (possibly using a CDEM equivalent). By comparing these two values, a reliable proportionality assessment can be made. This approach simplifies the task by making the values for autonomous systems comparable.[52]

However, this is a complex process that requires a high level of AI, which might not be readily available at the time of this study. Nevertheless, it is suggested that with time, AWS will become capable of adhering to the principle of proportionality.[53]

Furthermore, the compliance of AWS with the principle of proportionality depends on the degree of firepower they control and how much firepower is used simultaneously. For example, if an AWS uses nuclear weapons, the collateral damage could be extensive. However, if it employs a high-precision micro-projectile or a laser beam, the potential collateral damage in case of a miss would be relatively low. Therefore, it is recommended to keep the level of firepower controlled by AWS relatively low to avoid causing disproportionate damage in case of malfunctions or software errors. Additionally, since machines lack the right to self-defense, they should not have the same authority to use force and firepower as a human in the field.

4. The Precautionary Principle

This principle was first enshrined in Article 2(3) of the Hague Convention (IX) concerning Bombardment by Naval Forces in Time of War of 1907, and state practice establishes this norm as customary international law.[54] The principle of taking precautionary measures requires the adoption of all possible measures to avoid or minimize incidental harm to the civilian population. Firstly, there is an obligation to provide timely warning to the civilian population of an attack on military objects if such an attack may cause them harm, except when this is impossible due to tactical circumstances. Secondly, the principle of precaution also extends to the choice of means and methods of warfare. For example, if military advantage necessitates an attack on a military object located near a soccer stadium, the commander, in adhering to precautionary measures, must time the strike so that there is no civilian population on the soccer stadium at that moment.

It is worth noting that the requirement for taking precautionary measures applies throughout the entire process of armed deployment planning and concerns all individuals

---

[51] US Department of Defense, 'No-Strike and the Collateral Damage Estimation Methodology', (2012) CJCSI 3160.01A D-A-7.
[52] M. Newton and L. May, Proportionality in International Law, 285, (OUP 2014).
[53] M. Guetlein, 'Lethal Autonomous Weapons: Ethical and Doctrinal Implications', 126, (2005) Research Report Maxwell, AFB: Air War College, https://apps.dtic.mil/sti/pdfs/ADA464896.pdf (Last visited: 23 October 2023).
[54] ICRC Study on Customary International Humanitarian Law, 44, (1996).

involved in preparation. This includes not only commanders but potentially also weapon system manufacturers and programmers of AWS.[55] Consequently, considering the possibility of various unforeseen situations and challenges arising during the deployment of AWS and the execution of combat tasks, there is a basis for the obligation to always have a human "in the loop." In other words, an operator who will be responsible for monitoring and responding to various new situations as they develop.[56]

Adhering to caution implies measures that are highly context-dependent and subject to rapid and unpredictable changes. They involve continuous target assessment, weapon selection, timing, and method of attack. An attack must be halted if it becomes evident that it will have disproportionate consequences or if the target is no longer (or no longer remains) lawful. However, it is essential to always remember the need for ongoing reassessment, which raises the question: can AWS perform the required assessment without human intervention? The necessity for constant reassessment of circumstances, in conjunction with the existing capabilities of AI in AWS, suggests that compliance with this principle by autonomous systems may not be feasible in the near future, at least for several years.

Therefore, it appears prudent to deploy autonomous functions primarily against military targets, such as military aircraft, ships, or in situations where the risk to civilian populations is virtually non-existent. While the deployment of defensive weaponry with autonomous capabilities against enemy projectiles or missiles does not pose significant compliance issues with respect to IHL, an intriguing question arises regarding whether the deployment of fully AWS will be restricted to situations where the encounter with the civilian population is ruled out from the outset. This is because, given the fact that conflicts are increasingly becoming non-international, with no clearly defined geographical frontlines, military objectives primarily located in civilian areas, and combatants who intentionally do not clearly distinguish themselves from non-combatants, the ability of AWS to exercise caution is currently not feasible.

5. The Principle of Unnecessary Suffering

This principle is enshrined in Article 35(2) of API and essentially reiterates the customary principle of IHL, which prohibits the use of weapons, projectiles, materials, and methods of warfare that are capable of causing excessive damage or unnecessary suffering. The primary purpose of this principle is to ensure that the injuries and suffering inflicted on combatants do not exceed the necessary level required to render them ineffective. Therefore, it is forbidden to use any means or methods that cause excessive damage or inflict unnecessary suffering unless such effects are required to render the enemy combatants ineffective in any case.[57]

---

[55] W. Boothby, *'Conflict Law: The Influence of New Weapons Technology'*, 115, (2014), T.M.C. Asser Press: Den Haag.

[56] M. Trascasas, N. Weizmann, AWS under International Law, Academy Briefing No. 8, N. Weizmann, Geneva Academy of International Humanitarian Law and Human Rights, 4, (November 2014).

[57] C. Pilloud, J. Pictet, 'Protocol I – Article 57 – Precautions in Attack' in Yves Sandoz', C. Swinarski and B. Zimmermann, Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949, 686, (1987).

F. Case study: Current Challenges in the Deployment of AWS in Armed Conflicts (Autonomy in Existing Weapon Systems)

1. United Nations Expert Group on Libya: Existing Lethal Autonomous Weapon Systems - A Line Crossed?

To raise public awareness of the issue, in 2017, the Future of Life Institute released a viral video called "Slaughter Bots." [58] At that time, many experts considered it overly fantastical, believing that the emergence of such systems should not be expected before the mid-21st century. However, a few years later, the potential use of such systems is already being discussed within the UN. In March 2021, a UN expert group on Libya, where a civil war has been ongoing for several years, reported the possible use of lethal AWS, such as STM Kargu-2. The official report[59], consisting of over 500 pages, provides a detailed account of events from October 2019 to January 2021, with information regarding lethal drones presented within the context of the March 2020 battles. The report indicates that on March 27, 2020, the Prime Minister of Libya, Fayez al-Sarraj, issued an order for the "Peace Storm" operation, during which drones were used against forces associated with Haftar. According to the report, Libyan authorities employed Turkish-made STM Kargu-2 drones to strike a column of the Libyan National Army forces retreating from Tripoli. Subsequently, logistical columns and retreating armed forces were tracked and remotely engaged using combat drones or lethal AWS, such as STM Kargu-2 and other loitering munitions. The augmented capacity for operational reconnaissance included Turkish electronic warfare assets and reconnaissance, surveillance, and reconnaissance capabilities provided by Bayraktar TB-2 drones and, likely, TAI Anka-S. This, in turn, enabled the deployment of an attrition warfare strategy.[60] But this was unlike previous drone strikes. According to the incident description, the STM Kargu-2 drone could track and engage targets remotely. According to the report, “the lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true “fire, forget and find” capability.”[61]

The report does not provide information on whether there were casualties or fatalities associated with the attack. However, it does note that the drones "very effectively" assisted in causing "significant losses" to enemy anti-aircraft missile systems, which generated a flurry of sensational headlines in the media about the first-ever fully autonomous drone attack on humans in history.[62] Furthermore, the report lacks technical details, and even with the presence of such details, it is presumable that the situation would not necessarily be clarified. Despite the report containing numerous pieces of information about violations of IHL and international human rights law (IHRL), it does not mention the unlawful use of Kargu-2. Therefore, the following legal aspects that could be relevant to the deployment of STM Kargu-2 will be considered. However, for an assessment of the legality of using STM Kargu-2 in the specific context of the armed conflict in Libya, insufficient information is available.

The analysis of the report's content has revealed that UN experts do not differentiate between lethal AWS and kamikaze drones (loitering munitions). The latter, while waiting for

---

[58] "Slaughter Bots" (DUST), https://www.youtube.com/watch?v=O-2tpwW0kmU (last visited: 9 September 2023).
[59] LETTER DATED 8 MARCH 2021 FROM THE PANEL OF EXPERTS ON LIBYA ESTABLISHED PURSUANT TO RESOLUTION 1973 (2011) ADDRESSED TO THE PRESIDENT OF THE UN SECURITY COUNCIL (S/2021/229).
[60] ibid, para 60.
[61] ibid, para 63.
[62] *'Turkish Drones STM Kargu-2 Autonomously Attacked Humans*', ROBOCRAFT, (31 May 2021), https://robocraft.ru/news/4177 (last visited: 31 September 2023).

a specific target and signal to strike, patrol a particular area. For example, the Israeli IAI Harpy, developed as far back as 1980, has predecessors that are still produced, including in Azerbaijan under the name "Zarba," and they were actively used during the Nagorno-Karabakh armed conflict.[63] Yes, the deployment of loitering munitions in Libya, in this case, does not introduce any elements of novelty. Moreover, the manufacturers of these loitering munitions themselves classify them as "all-weather autonomous weapons."[64] However, the wordings featured in the UN report still indicate that there is something new at play.

Kargu-2 is a UAV created by the Turkish company STM. It utilizes machine learning algorithms integrated into the platform, enabling it to operate autonomously and be manually controlled. Unlike the Bayraktar TB2 or the Israeli loitering munition Harpy (as mentioned above), Kargu-2 is a weapon (tactical quadcopter) capable of selecting human targets and engaging them based on object classification through machine learning.[65] Although various ammunition options are available, the Kargu drone detonates the explosive charge near the target, minimizing the likelihood of collateral damage. It's worth noting that despite not having particularly impressive flight characteristics, Kargu is capable of orienting itself using visual data and reference points in the landscape, which makes it less susceptible to interference with GPS signal jamming systems.[66] The manufacturer, STM, advertises the anti-personnel capabilities of the Kargu drone in a grim video, showcasing a model of the Kargu diving towards a target amidst a group of mannequins representing people. According to the manufacturer's claims, the systems employ various facial recognition tools to identify and track potential targets, and in case they cannot fulfill their mission, the systems return to base. Of particular note in this context is the existence of separate loitering munitions (drones) in STM's product line, such as Alpagu, which presumably indicates Kargu-2's capability for fully autonomous tracking and targeting. STM's CEO, Murat Ikinci, emphasized that a squadron of 30 Kargu drones is powerful enough to destroy a military unit and a warship. Each Kargu has a specific mission. If one of the drones in the team is attacked or disabled during an operation, others replace it and carry out the assigned task, confirming the presence of AI and facial recognition systems in the drones. However, as of now, there is no confirmed case of AWS attacking humans, including by STM and Turkish authorities.

Returning to the events of 2020 in Libya, it is highly likely that drones possessed some capability to identify moving objects in videos, potentially including the ability to distinguish between people and other objects such as cars and buildings. However, they lacked certain other functions typically associated with full autonomy, such as the ability to prioritize targets, dynamically execute complex tactics, or make decisions in accordance with the laws of armed conflict.[67] In such scenarios, image processing may prove insufficient for detecting and identifying individuals who are unconscious or suffering from internal injuries or illnesses. On the other hand, leaving one's weapon may be programmed into an autonomous weapons system as a sign of surrender. However, there are situations where individuals may be unable to

---

[63] J. Antal, *'7 Seconds to Die: A Military Analysis of the Second Nagorno-Karabakh War and the Future of Warfighting'* (2022), https://www.airuniversity.af.edu/Aether/Book-Reviews/Article-Display/Article/3218364/7-seconds-to-die-a-mi litary-analysis-of-the-second-nagorno-karabakh-war-and-the/ (Last visited: 17 September 2023).

[64] Autonomous weapon for all Weather, https://www.iai.co.il/p/harpy (Last visited: 31 September 2023).

[65] Kargu-2 Specifications, https://www.stm.com.tr/en/kargu-autonomous-tactical-multi-rotor-attack-uav (Last visited: 31 September 2023).

[66] R. Fishman, *'Everything You Need to Know About Killer Drones from Turkey*,' TECHINSODER, https://www.techinsider.ru/weapon/712883-vse-chto-nado-znat-o-dronah-ubiycah-iz-turcii/ (Last visited: 1 September 2023).

[67] Footnote 60, para 63.

relinquish their weapons, potentially leading to misinterpretation by the machine. The use of autonomous weapons systems like Kargu-2 without the capacity to detect and provide assistance to incapacitated individuals raises questions regarding these systems' compliance with IHL. Indeed, employing such a system while aware of its inability to meet these legal requirements may be akin to an order not to spare anyone, i.e., to show no mercy or leniency and take a life in exchange for surrender, which is prohibited under customary IHL.[68]

2. GR-A1 Autonomous Security Robot: Border Safeguard or Severe IHL Violation on the Korean Peninsula

As previously noted, an operator can trust the machine and not monitor its operation properly, practicing "blind trust in the machine". It is also important to consider that in the case of the "human on the loop" concept, an AWS is capable of functioning without operator intervention and can be controlled not constantly, but only as needed, for example, as in the case of the all-weather autonomous security robot SGR-A1, which can support troops in the demilitarized zone separating North and South Korea, guarding key military facilities. The developers of SGR-A1 report that the system was created to replace people who may suffer from bad weather or fatigue. SGR-A1 has three low-light cameras,[69] thermal detectors, motion detectors, as well as image recognition software, [70]allowing it to detect targets up to two miles during the day and one mile at night.[71] Upon detecting an intruder, SGR-A1 issues verbal warnings, followed by rubber bullets, and subsequently, live metal bullets from the onboard machine gun of the AWS. SGR-A1 can operate as a "man on the loop" system. This means that the system is capable of autonomously selecting and engaging targets, but if necessary, a human operator can intervene to deactivate the system. This is fundamentally different from the concept of "human in the loop" (see above), where the weapon system awaits commands from operators to engage a target.

Due to the unstable situation in the region and the criticism directed at the developers of SGR-A1 and the authorities of South Korea for using "killer robots" to guard their borders, it is impossible to assert with certainty the presence of fully autonomous functions in this weapon system, just as it is impossible to determine the specific number of "sentries" and unsuccessful incidents that have occurred or continue to occur in the demilitarized zone of South Korea, as this data is classified from the public. Based solely on the available information at the moment, an analysis of the application of SGR-A1 reveals that the ability to address immediate threats is indeed valuable in the demilitarized zone since South Korean defense is stretched over 250 kilometers. This means that forces must respond to any threats as quickly as possible to hold their positions until reinforcements arrive and prevent an invasion.

Therefore, the "man on the loop" system is suitable for addressing this problem. Firstly, it is difficult for operators to manage multiple systems simultaneously, even in non-combat situations. In the rapidly changing battlefield scenario, the capabilities of system operators,

---

[68] ICRC Study on Customary International Humanitarian Law, Volume II, Chapter 15, Section A. Rule 46, D. Hambling, '*Israel used world's first AI-guided combat drone swarm in Gaza attacks*' New Scientist (2021), https://www.newscientist.com/article/2282656-israel-used-worlds-first-ai-guided-combat-drone-swarm-in-gaza-attacks/ (Last visited: 4 September 2023).

[69] J. Kumagai, 'A Robotic Sentry For Korea's Demilitarized Zone' (2007) IEEE Spectrum https://spectrum.ieee.org/a-robotic-sentry-for-koreas-demilitarized-zone (Last visited 4 September 2023).

[70] S. Weinberger, *'Next Generation Military Robots Have Minds of Their Own'*, BBC, (2014) https://www.bbc.com/future/article/20120928-battle-bots-think-for-themselves (Last visited 4 September 2023).

[71] Samsung Techwin SGR-A1 Sentry Guard Robot https://www.globalsecurity.org/military/world/rok/sgr-a1.htm (Last visited: 9 September 2023).

"human in the loop," to make timely decisions will be limited. Attacks or synchronized incursions in different parts of the battlefield can overwhelm operators, depriving them of the ability to track multiple threats and allocate defensive resources appropriately. This, in turn, increases the likelihood that the defensive line may be breached. Furthermore, the "human in the loop" system relies on human input before engaging a target, creating an opportunity for the target to escape. This can be costly and cause irreparable harm to military interests. For instance, in a "counter-sniper" situation, the fraction of a second needed to obtain operator approval to open fire can hinder the neutralization of the enemy.

An interesting aspect from a compliance standpoint with the laws of armed conflict is the claimed ability of the system to recognize signs of surrender.[72] For instance, upon detecting an intruder, SGR-A1, in addition to issuing verbal warnings, can recognize behavior and movements indicative of surrender.[73] However, details about how SGR-A1 recognizes surrender signs are lacking. From the manufacturer's claims in a promotional video, which suggests that this weapon system can be used on military bases and deployed on tanks, it follows that SGR-A1 could be deployed in the context of an armed conflict. This raises numerous questions, as mere capability for recognizing surrender is insufficient to comply with the laws of armed conflict during an armed conflict. For example, how does the system determine what constitutes surrender? Presumably, programming "raising hands" or "dropping weapons" as surrender signs would not be enough to enable the system to qualify certain behaviors as surrender and, therefore, to cease the use of force and adhere to the laws of armed conflict. Can the system determine if a soldier is incapacitated (which also requires the cessation of the use of force in accordance with the laws of armed conflict) and, for example, needs assistance?

Furthermore, the issue of determining a legitimate target remains relevant in this context, considering the concept of direct participation in hostilities (as mentioned above) required to comply with the principle of distinction. As repeatedly noted, modern machine learning-based systems are unlikely to make decisions considering contextual factors. For example, an autonomous system cannot distinguish a farmer in camouflage from a soldier. Adequate classification of a vehicle is also challenging since a variety of contextual and other factors may hinder this. Research[74] has shown that systems cannot recognize partially obscured objects: due to poor visibility of the wheels and front windows of a bus, the system identified the bus as a bicycle. In a similar vein, the system can easily distinguish a tank in open terrain with good lighting from a bus, but it would struggle with the same task if key distinguishing features of the tank were obstructed, for example, by trees or buildings. Weather conditions also play a significant role: studies confirm that in foggy weather, the accuracy of the AI system used to detect obstacles on roads drops to 58% compared to 92% in clear weather, which is also typical for humans.[75] Moreover, there is a high probability of deceiving AWS on the

[72] SAMSUNG SGR-1 - SECURITY GUARD, https://www.youtube.com/watch?v=xtE9hpwrDg4 (last visited: 4 September 2023).
[73] SAMSUNG TECHWIN SGR-A1 SENTRY GUARD ROBOT, https://www.globalsecurity.org/military/world/rok/sgr-a1.htm (Last visited: 4 September 2023).
[74] A. KORTYLEWSKI, Q. LIU, H. WANG, Z. ZHANG, A. YUILLE, 'COMBINING COMPOSITIONAL MODELS AND DEEP NETWORKS FOR ROBUST OBJECT CLASSIFICATION UNDER OCCLUSION', (2020), https://arxiv.org/abs/1905.11826 (last visited: 10 October 2023).
[75] Z. LIU, Y. HE, C. WANG, AND R. SONG, 'ANALYSIS OF THE INFLUENCE OF FOGGY WEATHER ENVIRONMENT ON THE DETECTION EFFECT OF MACHINE VISION OBSTACLES' (2020) https://www.mdpi.com/1424-8220/20/2/349 (Last visited: 4 September 2023).

battlefield, for instance, by placing a simulator image of a school bus on a tank, as confirmed by a similar example, which has been verified in practice.

The stated problems necessitate the need for operator intervention in the machine's operations. Since SGR-A1 functions as a "human on the loop" system, the possibility of eliminating errors remains. This means that if SGR-A1 mistakenly targets non-hostile entities, the operator can deactivate it using the "soft" or "hard kill" option. The "soft kill" option is based on a wired or wireless communication link between the remote position and the robot; if something goes wrong, the operator sends a kill signal that stops the robot's activities. The "hard kill" option is a hardware access point on the machine itself that the operator can use to manually shut it down. In the event that a system like SGR-A1 mistakenly targets civilians, the operator can disable it. However, a problem arises here: this may not save the lives of the initial civilians attacked by the machine since the operator is unlikely to foresee the wrongful targeting, as is the case with the "human in the loop" system. Nevertheless, this would prevent mass wrongful casualties as the operator could disable the machine after the initial wrongful engagement. To avoid the tragedy of losing even a few innocent lives due to incorrect targeting by the "man on the loop" system, it is necessary to deploy the weapon system exclusively in an environment where the presence of the civilian population is minimal.

Today, the demilitarized zone is so heavily fortified that there are no civilians in it, and most North Korean defectors have to travel through China, Laos, and Thailand to bypass it.[76] This makes the demilitarized zone a controlled environment, where anyone who can physically enter SGR-A1's targeting area is reasonably considered a combatant. This is why Samsung engineers programmed SGR-A1 to identify anyone in the demilitarized zone as an enemy. Undoubtedly, such a software solution cannot be used on all borders since not all borders are equally controlled environments. For example, deploying SGR-A1 would do more harm than good on the US-Mexico border, where the vast majority of border violators do not pose a military threat. However, along borders where civilians do not move or can be restricted from moving, where a controlled environment can be reasonably established, "man on the loop" systems can be used to deter invasions.

Furthermore, to prevent errors and accidents, these systems should only be deployed in defensive operations, which also minimizes the threat to civilians. However, it is essential to prevent the use of these systems for offensive purposes, which requires the commitment of states using SGR-A1 and other AWS to establish "defensive intent." Such intent can be expressed, for example, by setting a specific firing range to prevent the possibility of striking deep into the territory of neighboring countries, ensuring the deployment of systems in areas with minimal civilian presence. In this regard, an AWS deployed in densely populated cities in the context of new wars (as mentioned above) will face a high likelihood of failing to comply with the laws of armed conflict, especially the principle of distinction.

From the examples mentioned above, it becomes clear that despite the lack of technical details, the information regarding violations of IHL by these AWS, a key question still remains. To what extent can AWS adjust and regulate their own behavior after activation to the point

[76] Debra Kamin, *'How to escape from North Korea*' (2014), THE TIMES OF ISRAEL, https://www.timesofisrael.com/what-does-it-take-to-escape-north-korea/ (Last visited: 04 October 2023).

where their behavior becomes unpredictable in terms of their ability to comply with IHL? This was also confirmed during the interview with Alex Leveringhaus.[77]

III. RESULTS

Based on the adopted approach, the term "AWS" encompasses a broad category of systems capable of independently selecting, tracking, and engaging targets without constant human intervention. This approach facilitated the categorization of these systems into three distinct types: semi-autonomous, supervised autonomous, and fully autonomous systems.

In the course of research into the legal classification of AWS within IHL, a number of challenges were encountered. While there are viewpoints characterizing these systems as artificial combatants, it was discerned that their functional aspects do not influence their legal categorization. Instead, they should be considered as instruments for executing military operations, following instructions pre-programmed by humans.

When examining compliance with IHL principles, it became evident that adherence to these principles goes beyond target selection. It necessitates considering various contextual factors, particularly for the principle of distinction, which requires evaluating contextual elements, including socio-cultural factors. This level of analysis surpasses the capabilities of AI in the foreseeable future. In addressing compliance with the principle of proportionality, we employed utilitarian principles, translating ethical considerations into measurable terms, such as collateral damage and military necessity. It was concluded that AWS may excel at making these calculations more efficiently than humans. However, the need for constant contextual reassessment poses a challenge to the principle of precautionary measures.

CONCLUSION

In conclusion, the impact of AWS on armed conflicts and their compliance with IHL remains uncertain due to limited empirical data. Nevertheless, there is an imperative for the international community to proactively address these concerns. Fully AWS, in the current state of AI development, may potentially violate IHL norms if deployed without careful consideration. The risks of these systems falling into non-state actor hands are substantial, emphasizing the necessity for comprehensive regulation. Interested states, international organizations, the UN, and global civil society should advocate for consensus on specific "rules" governing the development and deployment of AWS. These rules should encompass their goals, geographical coverage, deployment scenarios, operator interaction requirements, and weaponry specifications while these systems are still in their early stages. Failing to do so could exacerbate the gap between modern technologies and the law, potentially leading to the development of uncontrolled weapons, as was the case with nuclear weapons.

[77] Dr. Alex Leveringhaus, a Ph.D. holder in the field of public administration, a research fellow at the Institute of Ethics, Law, and Armed Conflict at the University of Oxford, the coordinator of the Special Group on Ethics and Artificial Intelligence, personal interview 02.07.23.

THE INFLUENCE AND CONTRIBUTIONS OF CONFUCIANISM TO THE DEVELOPMENT OF INTERNATIONAL LAW

Weitao Gong*

Abstract: Confucianism requires individuals to take into account the fairness, justice, harmony and peaceful development of others and the whole society while developing themselves. Contemporary international law lacks this requirement of Confucianism. International law, dominated by western civilization, has neither proper regulation nor necessary condemnation toward some countries' behaviors of neglecting morality and sacrificing the common interests of mankind when pursuing their own interests. In this regard, Confucianism can play a positive theoretical role in the proper development of contemporary international law for it to return to the direction of justice and morality. Besides, Confucianism itself has a theoretical and historical connection with international law and its natural law thoughts. It also has already made positive contributions to the sustainable development of international law, world peace and the development of international human rights. Thus, it is feasible for Confucianism to contribute to the improvement of contemporary international law. The idea of building a human community with a shared future is the modern interpretation of Confucianism, which can be more systematically used as the theoretical basis for improving contemporary international law, so as to ensure the unity of powers and responsibilities of all countries and to take into account the interests of other countries in their own development.

Keywords: International Law; Confucianism; Unity of Nature and Man; Human Community with a Shared Future

* Fudan University Law School, China.

Table of Contents

Introduction.....49

I. The Development Dilemma of International Law: Causes and New Solutions.....50

A. The Development Dilemma of International Law and its Causes.....50

B. Confucianism as the New Approach to Resolve the Development Dilemma of International Law.....52

II. Core Connotation of Confucianism.....53

A. Unity of Nature and Humanity.....53

B. The Confucian Tradition of Rite and Law and Its Moral Concepts of "Self-Restraint", "Tolerance" and "Moderation".....53

C. To "Seek Harmony in Diversity" in the "Great Harmony Society".....54

III. The Feasibility of Using Confucianism as a Solution to the Development Dilemma of International Law.....55

A. Similarities Between Confucianism and the Thoughts of Natural Law....55

B. The Influence of Confucianism on International Law.....56

1. The Influence of Confucianism on the Theoretical Background at the Beginning of International Law.....56

2. The Influence of Confucianism on the Construction of International Organizations.....57

C. Contributions Confucianism Has Made to the Development of International Law.....58

1. The Contribution of "Unity of Nature and Humanity" to International Sustainable Development.....58

2. The Contribution of Confucian Rites and Laws to World Peace: the Five Principles of Peaceful Coexistence.....59

3. The Contribution of "Benevolence" to The Development of International Human Rights Law.....59

IV. Contributions of Confucianism to the Future Development of International Law.....60

A. The Idea of Building a Human Community With a Shared Future: a

Modern Interpretation of Confucianism 60

B. The Necessity to Apply the Concept of Building a Human Community With a Shared Future to the Development of Contemporary International Law 61

C. Application of the Concept of Building a Human Community with a Shared Future to the Development of Contemporary International Law 61

1. Feasibility of the Application 61

2. The Application of the Idea of Building a Human Community with a Shared Future to the Development of International Law 62

Conclusion 62

INTRODUCTION

The 21st century is in a major upheaval that global economic and security issues are becoming increasingly serious due to deglobalization. The United States, in pursuit of its own national interests, shows reluctancy to cooperate with the current multilateral international system and instead the United States intends to adopt protectionism, replacing the broad and long-term global common interests with its own demands and interests.[1] To be more specific, the U.S. government takes climate change and environmental protection as political leverage and to ensure their superiority in global digital and economic governance, national security, long-arm jurisdiction and sanctions have been abused to suppress emerging technology and economy entities especially in developing countries such as China.[2] In addition, for decades, the United States has maintained double standards on issues such as human rights, justice and peace and by means of the "Indo-Pacific Economic Framework", the United States seeks to exclude countries such as China from global and regional market chain.[3] These actions of the United States are incompatible with the purposes of the United Nations Charter (hereinafter referred to as the "Charter"). However, the current international legal system faces difficulties in regulating these actions.

Natural law, defined by Hugo Grotius, the founder of international law, as the rule of right reason which teaches us that an act is just in so far as it conforms to natural reason and moral just[4], serves as the basis of international law, though it gradually declines after the rise of positivism.[5] As a result, the development of international law lacks due justice and moral judgment to forbidden or condemn unjust behaviors. Furthermore, the revival of the new natural law school after the two world wars still fails to turn the direction of the development of international law back to the line of just and morality. The reason is that the concerning ideology of international law is still Western-centered.

Broadly defined, western centrism is an aspect of ethnocentrism, as it relates to the theory and practice of privileging western norms, categories, and narratives in the construction and development of international laws and institutions.[6] As a result, the idea of systemic western centrism commingles with the language of international law, translating values, claims, and critiques which are inherently western such as the concept of sovereignty and human rights into international law, leading those concepts and norms as dominant during the modernization and development of international law.[7] Besides, the structure of the international law is western-centered as international standards are often set by western institutions to scale and rank the non-western counterparts, and members of the international bench tend to pursue their studies in western universities prior to their appointment as judges which shows the domination

---

[1] See Zhipeng He, *Constitutional Moment: The Dilemma and Opportunity in the Development of International Law*, 34 CONTEMPORARY LAW REVIEW 14 (2020).
[2] See Joe Gould, *White House Aims to Release Overdue Security Strategies within Weeks*, DEFENSE NEWS (2022), https://www.defensenews.com/pentagon/2022/08/01/white-house-aims-to-release-overdue-security-strategies-within-weeks/ (last visited Sep 21, 2023).
[3] See Wang Shuqin, *Indo-Pacific Economic Framework Cynical Manipulative Tool of US: China Daily Editorial*, CHINADAILY (2022), https://www.chinadaily.com.cn/a/202205/25/WS628e20dba310fd2b29e5ef6e.html (last visited Sep 21, 2023).
[4] See GROTIUS HUGO, DE JURE BELLI AC PACIS LIBRI TRES (1925).
[5] See Guoqiang Luo, *The Evolution of Western Natural Law Thought*, SOCIAL SCIENCES INTERNATIONAL 33 (2008).
[6] See George Gheverghese Joseph, Vasu Reddy & Mary Searle-Chatterjee, *Eurocentrism in the Social Sciences*, 31 RACE & CLASS 1 (1990).
[7] See Sally Merry, *Human Rights and Transnational Culture: Regulating Gender Violence through Global Law*, 44 OSGOODE HALL LAW JOURNAL 53 (2006).

of western background in the actors in international law.[8] The distribution of power in the area of international law is thus biased and western centrism is exclusive to countries which do not share the dominant western values, leading to war and the ignorance and sacrifice of those countries' interests while western countries pursuing their own interests. Therefore, against the western-centered background of international law, a new methodology instead of western centrism needs to be introduced as a different approach to analyze and improve the international law system to be in line with justice, morality and equality.

Although international law is a product of Western civilization, this does not mean that the ideology of traditional Chinese culture, especially Neo-Confucianism (hereafter referred to as "Confucianism"), the Confucianism inherited and developed in Song and Ming Dynasty, has no connection with international law, nor does it mean that Confucianism has no impact or can make no contribution to the development of international law. As the essence of Confucianism, "unity of nature and humanity" leads the traditional Confucian ethics of "self-restrain", "tolerance" and "moderation" with "benevolence" as the core, as well as the thought of "seeking harmony in diversity" to reach the "Great Harmony Society". Confucianism pursues "balance" and requires that "one should not do to others what one doesn't want to do to himself" for the purpose of social fairness, justice, harmony and peace. At the same time, according to Confucianism, everyone should restrain their own behavior in a reasonable and legal manner so that while self-developing, people can balance their own interests and the interests of other people and the whole society. The ideology of Confucianism can provide positive contribution for the contemporary international law to move toward a more just and moral direction. First, the "balance" advocated by the Confucianism is precisely what contemporary international law lacks. Secondly, the Confucianism itself has theoretical and historical links with the development of international law and its ideological basis of natural law. Thirdly, Confucianism has had positive influences on the development of international law from the aspects of global sustainable development, world peace and international human rights. Having made those beneficial contributions, Confucianism can also continue to contribute to the future development of contemporary international law.

The report of *the 20th National Congress of the Communist Party of China* points out that although in modern age, traditional Chinese culture should be inherited, creatively transformed and developed.[9] As a result, in the area of international law, the idea of building a human community with a shared future has been raised as the modern interpretation of Confucianism, which will be more conducive for Confucianism to make its positive theoretical contributions to the development of international law in a systematic way so that every country can take into account the reasonable concerns of other countries when pursuing their own interests, and to promote the common development of all countries.

I. THE DEVELOPMENT DILEMMA OF INTERNATIONAL LAW: CAUSES AND NEW SOLUTIONS

A. The Development Dilemma of International Law and its Causes

International law as the conception in terms of order of the conduct of independent states including bodies of treaties, customs and other generally accepted principles governing

[8] See Salvatore Caserta, *Western Centrism, Contemporary International Law, and International Courts*, 34 LEIDEN JOURNAL OF INTERNATIONAL LAW 321 (2021).
[9] See Xi Jinping, *The Report to the 20th National Congress of the Communist Party of China*, (2022).

relations among states[10], was the law of war and peace at the outset. The thought of natural law, as the theoretical basis of international law, was also used to demonstrate and support that the wars launched by western countries were just.[11] It was until the Charter that the use of force was banned permanently. The natural law school was once almost replaced by the positive law school, but the result of the replacement was the two world wars, which failed people's high expectations for international morality and pursuit of global justice, peace and other common interests of all mankind.[12] Natural law school was thus revived, and the new school of natural law has reduced the abstraction of the natural law thought and secularized it into morality, believing that besides following the requirements of international law, countries should also take into account the universal responsibility of international morality.[13] However, the "Western-centered" value in the international system is so influential that it still dominates the contemporary international law, making it difficult for the idea of the new natural law school to regulate and condemn immoral even unlawful behaviors.

After World War II, the United Nations and other international political, economic and security institutions and the international legal system have been established, transforming the ideas of freedom, democracy, human rights, rule of law, and peaceful settlement of disputes into international laws. Indeed, those rules have played a positive role in regulating global economic and trade exchanges and have protected the interests of countries all over the world. However, some rules have been used by some Western countries as a tool for them to interfere the elections and internal affairs of other countries so as to seek geopolitical and economic interests for themselves, resulting in political instability and armed conflicts in the intervened countries.[14] For example, in global human rights governance, "responsibility to protect" requires that under the authorization of the Security Council, international community can only intervene to help when sovereign countries are unwilling or unable to protect their citizens from avoidable disasters such as slaughter, rape, and hunger.[15] However, in its first practice in Libya in 2011 and in Syria since 2015, "responsibility to protect" still serves as the tool for neo-interventionism of Western countries and cannot truly resolve humanitarian catastrophe.[16]

In addition, since the Trump administration took office, trade "war" between China and the United States has been constantly escalated by the United States, disrupting the global industrial chain.[17] In 2021, the United States even held the Summit for Democracy, trying to divide civilized countries again in the name of "democracy".[18] Pelosi, then the Speaker of the United States Congress, visited Taiwan China, on August 2nd, 2022, disregarding of China's strong opposition and solemn negotiations, seriously violating the one-China principle and the

---

[10] See Roland R. Foulke, *Definition and Nature of International Law*, 19 COLUM. L. REV. 429 (1919).
[11] See ARTHUR NUSSAUM, A CONCISE HISTORY OF THE LAW OF NATIONS (1947).
[12] See María Elósegui, *International Law and the Natural Law Tradition: The Influence of Verdross and Kelsen on Legaz Lacambra*, *in* THE THREADS OF NATURAL LAW: UNRAVELLING A PHILOSOPHICAL TRADITION 153 (2013).
[13] See Roland H. Bainton, *Review of Man and the State.*, 67 POLITICAL SCIENCE QUARTERLY 127 (1952).
[14] See Ru Sun, *China-US Competition and Their Different Ideas about International Order*, 373 CONTEMPORARY INTERNATIONAL RELATIONS 9 (2020).
[15] See BanKi-mon, *Implementing the Responsibility to Protect*, (2009).
[16] See Han Aiyong, *The Development and Evolution of Western Humanitarian Intervention after the Cold War*, STUDY TIMES, Jan. 25, 2019.
[17] See Office of the State Council, the People's Republic of China, *THE FACTS AND CHINA'S POSITION ON CHINA-US TRADE FRICTION*, CHINADAILY (2018), http://www.chinadaily.com.cn/kindle/2018-09/26/content_36981813.htm (last visited Sep 21, 2023).
[18] See Chi Ma, *Summit of Democracy Looks Back in Anxiety: China Daily Editorial*, CHINADAILY (2021), https://www.chinadaily.com.cn/a/202111/25/WS619f782fa310cdd39bc7793d.html (last visited Sep 21, 2023).

provisions of the three China-U.S. Joint Communiqués.[19] The pursuit of national interests and development of the United States is one-sided. However, the existing international legal system, though after the revival of the new natural law school, has not provided necessary legal regulations and moral condemnation against the actions of the United States and cannot prevent recidivism or imitation from other countries. By allowing one country to place its will and demands over other countries, the existing international legal system faced development dilemma, which departs from justice and morality and the legitimate routine and abandons the international legal concepts of fairness, righteousness, harmony, and peace. The development dilemma of international law also unveils the implicit value orientation of international law that is the foundation of international law still serves the interests of Western countries.

The Western-centered problem runs counter to the new international relations characterized by mutual respect, fairness, justice, and win-win cooperation, with the right of people outside of the Western countries neglected. By placing the self-interest of Western countries at priority, ignoring or even damaging the interests of other countries, interests and resources flow one-way to the West.[20] This kind of imbalance in development seriously endangers the stability of the world order.

B. Confucianism as the New Approach to Resolve the Development Dilemma of International Law

The United States has the most profound impact on the international legal system after World War II, but since the establishment of the federal government of the United States in 1789, till now, there was only 16 years in the history of the United States that it was not involved in or launched a war.[21] The failure of the new natural law school in the resolution of the development dilemma faced by the current international law shows that the international legal system needs new ideas. Unlike the United States, the gene of Chinese culture does not contain elements such as aggression, expansion, or foreign war. This is because, under the influence of Confucianism, China advocates "balance", emphasizes the cultivation of "benevolence" and always adheres to the concepts of "seeking harmony in diversity", so as to pursue coexistence with other ethnic groups and nature in peace.[22]

Fairness, justice, harmony, and peace is at the core of international law, aiming to protect the common interests of humanity. The core pursuit of international law is in consistency with the core idea of Confucianism including the "unity of nature and humanity", "self-restraint", "tolerance" and "moderation", and the thought of "seeking harmony in diversity". Those pursuit and thought all require members of the society to assume necessary obligations for the legitimate rights and interests of others, while maintaining their own interests under the premise of legitimacy and rationality, for the purpose of a harmonious world order. In this sense, Confucianism may provide us with a new approach for addressing the

---

[19] See Xinhua, *Statement by China's Ministry of Foreign Affairs on Pelosi's Visit to Taiwan*, CHINA.ORG.CN (2022), http://www.china.org.cn/world/2022-08/02/content_78353809.htm (last visited Sep 21, 2023).
[20] See Yongkou Zheng & Lanxin Fang, *Human Destiny Community Concept and Chinese Culture*, 26 HENAN SOCIAL SCIENCES 55 (2018).
[21] See Ben Norton, *US Launched 251 Military Interventions since 1991, and 469 since 1798*, GEOPOLITICAL ECONOMY REPORT (2022), https://geopoliticaleconomy.com/2022/09/13/us-251-military-interventions-1991/ (last visited Sep 21, 2023).
[22] See Michael McShane, *China's Era Marked by Peace without Hegemony*, MARTIN JACQUES (2010), http://www.martinjacques.com/when-china-rules-the-world/chinas-era-marked-by-peace-without-hegemony/ (last visited Sep 21, 2023).

development dilemma of Western-centered international law.

II. CORE CONNOTATION OF CONFUCIANISM

A. Unity of Nature and Humanity

The two traditions of Confucianism are the theory of nature and the theory of humanity under the thought of "unity of nature and humanity". The theory of nature is led by Mencius, with natural human nature as its origin and based on Confucius' theory that human nature is similar, Mencius proposed that human nature is inherently good while practice made it apart; the theory of humanity is represented by Xunzi, who proposed that human nature is inherently evil, believing that only humanistic education is the foundation for a person to become a man, and thus, emphasizing education and norms.[23] There was once a conflict between the theory of nature and humanity as for the human nature. The metaphysicians in the Wei and Jin dynasties who inherited Mencius' theory of nature thought that if human nature was not inherently good, it would be useless to learn classics according to Xunzi's view, meaning that education could not make the wicked good under the theory of humanity.

It was until Neo-Confucianism during the Song and Ming dynasties, the conflict between the two traditions was resolved through the methodology of rational thinking. Neo-Confucianism's cognition of the idealized and metaphysical world was based on the theory of nature, and their cognition of the real world was based on the theory of humanity. The "nature" in the thought of "unity of nature and humanity" was regarded by the Neo-Confucianism as the operation law of the universe and the highest principle of everything in the world, while the "humanity" as a solution to the current social situation and practical problems from a humanistic perspective by applying the principle of "nature".[24]

To understand the thought of "unity of nature and humanity" from the perspective of syllogism, "nature" is the major premise, as the precondition for being legal and reasonable; "rite" and "law" transformed from the principle of "nature", are the minor prepositions applicable to "human" problems; and the "human" conclusion is the application of the "rite" and "law" to practical social problems.[25] It can be seen that the thought of "unity of nature and humanity" in Confucianism is not irrational or illogical as some western scholars have conceived[26], but contains logic derived from practical rationality. The thought of "unity of nature and humanity" is thus not empty talk nor transcendental meditation, but it is based on social reality, with the priority to solve practical political and ethical issues.[27]

B. The Confucian Tradition of Rite and Law and Its Moral Concepts of "Self-Restraint", "Tolerance" and "Moderation"

The intermingled relationship between the rite and law has always been a prominent feature of Confucianism because the norms in traditional Chinese society are not just laws, but

---

[23] See Shunfu Shen, *On the Two Traditions of Confucian Philosophy*, 147 STUDIES OF ZHOUYI 81 (2018).
[24] See Youyun Lv & Minghui Xiong, *The Argumentation Model and Rational Foundation of Confucian Philosophy*, 277 FUJIAN TRIBUNE 85 (2015).
[25] See Chenjun You, *The Influence of Confucian Moral Idea on Traditional Chinese Litigation Culture*, 436 LAW SCIENCE 135 (2018).
[26] See Carl B. Becker, *Reasons for the Lack of Argumentation and Debate in the Far East*, 10 INTERNATIONAL JOURNAL OF INTERCULTURAL RELATIONS 75 (1986).
[27] See Sumin Xu, *A Comparative Study of Chinese Philosophy and Western Philosophy under Divine Light: On Matteo Ricci's The Truth of God, Nicolas Longobardl's On the Soul and the Tao and Malebranche's A Dialogue*, 206 JOURNAL OF SUN YAT-SEN UNIVERSITY (SOCIAL SCIENCE EDITION) 52 (2007).

also rites.[28] Rites, as the social values and basic norms that law needs to maintain, is not only about morality and customs, but also about national systems, cultural orders, and social systems. Its purpose is to build an ideal harmonious and orderly society.[29] Under the guidance of the thought of "Unity of Nature and Humanity", the Confucian tradition of rite and law integrates principle of nature, law of nation and relation of human altogether. As from the perspective of syllogism, rite is the result of the projection of the principle of nature to social facts, and the law of nation is derived from the ethical practice of the rite to regulate the relation of human. Human relations then follow the principle of nature, providing legitimacy for various circumstances in life.

Taking rite as the standard, people living in society can only achieve the core idea of "benevolence" by "self-restraint", "tolerance" and "moderation". The so-called "self-restraint", as a performance of "benevolence", refers to the act of conforming one's thoughts and actions to the provisions of "rite"; the moral concept of "tolerance" also revolves around "benevolence", and requires that when dealing with interpersonal relationships, one should be considerate, and should not do to others what one does not want to do to himself; the purpose of "self-restraint" and "tolerance" is to seek "moderation" since moderation means impartiality and neutrality.[30] The moral values of "self-restraint", "tolerance", and "moderation" constitutes the ethical principles of rite, requiring everyone to assume responsibility for themselves and for others and the society as well. In this way, a balanced state of all parties can be maintained by means of "benevolence" so as to achieve social harmony.

C. To "Seek Harmony in Diversity" in the "Great Harmony Society"

As for the "human" in the "unity of nature and humanity", Zhu Xi, a Neo-Confucianism scholar in the Song and Ming dynasties, once interpreted Mencius' claim that "people are the most important, the state is secondary while the monarch is the least", meaning "the state is based on the people, and the state is also built for the people" since people are the foundation of a nation.[31] This people-oriented philosophy is more evident in the "Great Harmony Society" ideology. The ideology advocates that "people should not only be close to their own relatives, nor to their only children, and by making the strong useful, having their children to grow into man, support can be provided for the widow, widower, the childless, those who are disabled and who lost their fathers at an early age and every single member of the society who are in old age."[32] This reflects emphasis on the right of every people in the society. The ideology of making the strong useful emphasizes the selection of talents and abilities in politics. It is for the sake of a guarantee of a stable and harmonious society, peaceful coexistence and common progress of humanity has been highlighted in the "Great Harmony Society" ideology since the ideology believes that harmony can be sought through diversity; people should do their best to achieve common interests and common development of all people regardless of the differences.[33] In fact, this is also a reflection of the core of "unity of nature and humanity" that

---

[28] See Victor H. Mair, *ELEVEN. Language and Ideology in the Written Popularizations of the Sacred Edict*, *in* ELEVEN. LANGUAGE AND IDEOLOGY IN THE WRITTEN POPULARIZATIONS OF THE SACRED EDICT 325 (2020).
[29] See Zhiping Liang, *Ethics and Law: Cultural Conflict in the Era of Legal Transplantation*, GUANGXI NORMAL UNIVERSITY PRESS 60 (2015).
[30] See Lai Chen, *Confucian Culture and National Revival*, ZHONGHUA BOOK COMPANY 84 (2020).
[31] See Desheng Cao, *President Emphasizes Cultural Confidence*, CHINADAILY (2021), https://www.chinadaily.com.cn/a/202103/24/WS605a7519a31024ad0bab11d3.html (last visited Sep 21, 2023).
[32] See Juyou Sun, *The Confucian Thought of Great Harmony and the Construction of a Community with a Shared Future for Humanity*, 37 DONGYUE TRIBUNE 63 (2016).
[33] See Siwei Song & Jiancheng Sun, *The Confucian Thought of Great Harmony and the Contemporary Concept of Harmonious Society*, SHANDONG SOCIAL SCIENCES 58 (2009).

is to pursue harmony and peace between man and nature, man and man, and man and society.

III. The Feasibility of Using Confucianism as a Solution to the Development Dilemma of International Law

Confucianism, guided by the thought of "unity of nature and humanity", gave birth to the ethical principles of rite and the rules of law in ancient China. The ideology of Confucianism, especially people-oriented thought of "Great Harmony Society", is not, as some scholars believe, irrelevant to international law or even contrary to the idea of international law. As the foundation of the birth and development of international law, the thought of natural law can serve as a bridge to bring Confucianism and the international law together since there has been theoretical connections and historical influence between Confucianism and the international law.[34]

A. Similarities Between Confucianism and the Thoughts of Natural Law

Since the beginning of international law, natural law has been regarded as the basis of international law. Vitoria believes that the law of nations is the law established among all countries based on natural reason; Grotius also explicitly places natural law as the basis of international law and believes that natural law is the order of legitimacy and reason and with natural law as the standard, any act consistent with reason and nature is morally just, otherwise, it is morally evil; Pufendorf also believes that the law governing the relations between countries is natural law; Wolff deduces from the natural international law that all countries have the same basic rights and obligations, and are naturally equal with no privileges among each other; Vattel believes that there is an internal natural law that constrains national moral consciousness, supplemented by external treaties and customs, both of which constitute international law.[35] The thought of natural law, as the foundation of international law, can be summarized as followed that the legitimacy of natural law is generally considered to originate from an absolute and unified existence, serving as the basis and the value criterion for secular law and political system. Natural law is thus the standard to judge whether positive law conforms to justice and morality. The purpose of international law is the same as natural law that is to seek fairness, justice, harmony and peace, since the idea of natural law is indispensable to international law. Lauterpacht regards the general legal principles in the Statute of the Permanent Court of International Justice as the essence of natural law in a modern sense and believes that the general legal principles are the main source of law of international law.[36] Therefore, natural law is still playing an important role in the contemporary international law.

The logic of "unity of nature and humanity" is similar to that of natural law. As mentioned above, from the perspective of syllogism, "unity of nature and humanity" can be understood in this way that in order to deal with the practical social problems, the principle of "nature" as the major premise is summarized as "rite", the ethical principles by people through rational observation, and "rite" is integrated into the rules of "law", which is then applicable to "man", human issues, as the conclusion. In this sense, the absolute and unified existence as the source of the legitimacy of natural law can be analogized to the principle of "nature" in "unity of nature and humanity", and the positive law that conforms to justice and morality under the doctrine of natural law, can be analogized to "rite" and "law". If one violates rite, they will be

[34] See Yang Zewei, *Western International Law and China's Confucianism in the 19th Century - Collision and Integration*, 13 J. HIST. INT'L L. 285 (2011).
[35] See BARDO FASSBENDER & ANNE PETERS, THE HISTORY OF INTERNATIONAL LAW (2012).
[36] See H. Lauterpacht, *The Grotian Tradition in International Law*, 23 BRIT. Y.B. INT'L L. 1 (1946).

despised as immoral by society. If they also violate "law", they should bear corresponding legal responsibilities. It is the same in the international legal system, which includes both international moral concepts at the level of international communication and legal norms as the embodiment of international moral concepts such as customs and treaties formed in practice. Violations of international morality will be internationally condemned, while violations of international legal norms will result in corresponding legal consequences. Therefore, the natural law ideology of the international law is similar to that of the "unity of nature and humanity" of Confucianism.

B. The Influence of Confucianism on International Law

1. The Influence of Confucianism on the Theoretical Background at the Beginning of International Law

International law emerges during the Renaissance and the Age of Enlightenment. At that time, Europe was undergoing the rise of capitalism and the decline of feudalism. Philosophy at that time also was faced by a dilemma whether the rational trend of thought should replace God's authority and the concept of nature should take the place of apocalyptic theology. It happened that during the Renaissance and the Age of Enlightenment, Neo-Confucianism was spread to Europe by Jesuits traveling between China and the West.[37]

The Jesuits went to China to preach Christian culture, hoping to convert non-Christian in Asia to believers. They developed especially high regard for the Chinese civilization and its Confucian philosophy in their reports as having established belief in God by the natural light of reason rendering Chinese easier to convert; for the purpose to support their conclusion in the report, the Jesuits translated classical texts of Confucianism into Latin and published in the seventeenth-century under the title *Confucius Sinarum Philosophus*; these reports and translations were widely read in Europe in the latter part of the seventeenth century, leading to a considerable body of literature and enthusiasm on the great civilization of Asia had built, even foreshadowing the East-West dialogue and finally integrated Confucianism deeply into the ideological debates of the Enlightenment period as the philosophical tools to overthrow the medieval religious theocracy.[38] This is because the ideology of Confucianism is people-oriented and rational. To put it in a simple way, the process of "nature" acting on "human" is not theological or apocalyptic, but rather derived from rational observation by human.

After the introduction of Confucianism in Europe, there were two kinds of repercussions. One kind of European philosophers attacked Confucianism for its atheism, such as the opposition of Jesuits to Neo-Confucianism; the other kind supported Confucianism for it was materialistic, atheist and naturalistic, such as Pierre Bayle, Voltaire, Leibniz and Quesnay.[39] Pierre Bayle believes that although the Confucianism was atheism, it still forms a school of thoughts that is self-consistent in logic and results in fine social governance, meaning that Christian theology is not the only prerequisite for a good social order; Voltaire also uses Confucianism as a weapon against the Christian Church, and contends that the reason why Confucianism achieves satisfied results in social governance is not because its thought is based on faith, but because its thought is based on reason, which is the core idea revived by the Age of Enlightenment; Leibniz holds the same view, believing that China has formed a natural belief

[37] See D. E. MUNGELLO, THE GREAT ENCOUNTER OF CHINA AND THE WEST, 1500–1800 (4th ed. 2012).
[38] See J. J. CLARKE, ORIENTAL ENLIGHTENMENT: THE ENCOUNTER BETWEEN ASIAN AND WESTERN THOUGHT (2003).
[39] See QIANZHI ZHU, THE INFLUENCE OF CHINESE PHILOSOPHY ON EUROPE (1999).

based on rationality rather than revelation; Quesnay believes that only by dispelling all unnatural and artificial restrictions, such as theological restrictions on freethought, economic prosperity, social happiness and harmony can be achieved.[40] From this, we can tell that the first connection between Confucianism and international law, is the theoretical background at the birth of international law was influenced by Confucianism and its rational thinking. Natural law, as the ideological basis of international law, may thus also be driven by this rational trend to break away from divine law and become a more modern standard to judge whether positive law conforms to justice and morality during the Renaissance and the Age of Enlightenment.

2. The Influence of Confucianism on the Construction of International Organizations

There is also a connection between the "Great Harmony Society" ideology of Confucianism and the establishment of international organizations. Christian Wolff, who once put forward the idea of "world federation", has a strong interest in Chinese philosophy, and his interest stems from his teacher Leibniz since Leibniz admires Chinese culture and it is believed that Leibniz's theory is deeply influenced by the thought of "Great Harmony Society"; besides, Wolff also gave a speech on the "Practical Philosophy of the Chinese" in 1721, expressing his worship of Confucianism.[41] The "world federation" constructed by Wolff was based on natural law. He believes that the "world federation" is the basic principle of natural law, which can bring people from all over the world together, from family society to feudal society, collective society to civil society, and finally to the Great Society belonging to all mankind; in this society, individuals have reached an actual contract, so they must abide by the natural law underlined by the society; at the same time, the moral system of the "world federation" should be orderly and meticulous, so all individuals and collectives who pursue perfection and happiness can fulfill their responsibilities in the whole social system to play their due role.[42] Based on this, Wolff constructed a universal society with the goal of achieving human harmony in his theory of natural law.

It is not difficult to conclude that the construction of the "world federation" is similar to the "Great Harmony Society" of Confucianism and from Wolff's preference of Chinese philosophy, it can be inferred that the "Great Harmony Society" thought of Confucianism exerts influences on Wolff's idea of "world federation". The concept of "world federation" was put into practice to some extent in the form of the League of Nations and the United Nations in the 20th century. Due to the similarity between the "Great Harmony Society" of Confucianism and the concept of the "world federation", a deep connection between Confucianism and the construction principles and ultimate goals of international organizations has been established that is because of the influence of Confucianism to Wolff and his "world federation", Confucianism is deeply connected to the construction of international organizations since the "world federation" is the original construction principles of international organizations, and from this perspective, Confucianism can also be considered as the underlying principles of international organizations.

However, as Samuel Huntington has said, what makes the culture, ideology, and other soft power of one country more attractive is the victory of the country's economic, military, and other material strength.[43] During the Renaissance and the Age of Enlightenment, China's

[40] See J. J. CLARKE, *supra* note 39.
[41] See J. J. CLARKE, *supra* note 39.
[42] See FASSBENDER AND PETERS, *supra* note 36.
[43] See SAMUEL P. HUNTINGTON, THE CLASH OF CIVILIZATIONS AND THE REMAKING OF WORLD ORDER (2011).

prosperity was sought after by Western countries, and it was believed that the traditional Chinese culture and ideological system represented by Confucianism were worth learning from. However, after the primitive accumulation of capital and the plunder of the Age of Great Navigation, the strength of Western countries greatly improved. Therefore, after the 18th century, Western enthusiasm for Confucianism and Chinese ideology gradually cooled down.[44] However, this does not mean that Confucianism is unrelated to the development of international law. On the contrary, to this day, traditional Chinese ideology represented by Confucianism still can contribute for the improvement of international morality and the development of contemporary international law to be in line with justice and morality.

C. Contributions Confucianism Has Made to the Development of International Law

Confucianism is related to international law in both theoretical and historical dimensions. Furthermore, Confucianism has also made positive contributions to the practice and the development of international law, rendering it feasible to resolve the development dilemma of international law.

1. The Contribution of "Unity of Nature and Humanity" to International Sustainable Development

The thought of "Unity of Nature and Humanity" believes that man and nature are interlinked in essence that everything and every people in the world all follow the principle of nature to achieve the harmonious coexistence of man and nature.[45] The thought also advocates and attaches importance to tolerate the development of all things in the world which is in accordance with the principle of nature since it is necessary to ride the trend of technological revolution and industrial transformation, seize the enormous opportunity in green transition, and let the power of innovation drive us to upgrade our economic, energy and industrial structures, and make sure that a sound environment is there to buttress sustainable economic and social development worldwide.[46] To balance economic development and environmental protection is also the requirement of 2030 Agenda for Sustainable Development of United Nations.

Ecological issues are important issues that affect the future and destiny of mankind. Though the idea of "unity of nature and humanity" originated from ancient Chinese civilization, it still has adherence to the modern value of harmonious coexistence of man and nature. Under the guidance of the thought of "unity of nature and humanity", not only does China contribute to the path of China's sustainable development, but it also has a positive impact on international biodiversity, prevention of species loss, and reduction of carbon emissions issues.[47] The thought of "unity of nature and humanity" can help mitigate the sharp confrontation between man and nature and to maintain eco-balance. It should not remain only as the value basis for China but should also be the theory reference for global sustainable development strategy to help coordinate with the international environmental and economic governance and the green and low-carbon transformation of the international community.

---

[44] See J. J. CLARKE, *supra* note 39.
[45] See Shuqiao Li & Zhaoxia Kong, *A Community with a Shared Future for Mankind: The Extension of Wisdom in Chinese 'Harmony' Culture*, 29 PACIFIC JOURNAL 49 (2021).
[46] See Jinping Xi, *For Man and Nature: Building a Community of Life Together*, CHINADAILY (2021), https://language.chinadaily.com.cn/a/202104/23/WS6082197fa31024ad0bab9c28.html (last visited Sep 22, 2023).
[47] See Di Zhou, *China's New Concept of Development from the Perspective of the Sustainable Development Goals*, 6 BRAZ. J. PUB. POL'Y 236 (2016).

2. The Contribution of Confucian Rites and Laws to World Peace: the Five Principles of Peaceful Coexistence

The Five Principles of Peaceful Coexistence put forward by China share the same connotation with the Confucian concept of "self-restraint", "tolerance" and "moderation". The Five Principles of Peaceful Coexistence require that all sovereign states do not infringe upon each other and do not interfere with each other's internal affairs, which is the embodiment of "self-restraint"; the requirement of mutual respect for sovereignty and territorial integrity, equality and mutual benefit is the reflection of the concept of "tolerance", and the ultimate goal of peaceful coexistence can be achieved by "moderation"; the Five Principles of Peaceful Coexistence emphasize the unity of rights and obligations of states, and represent the aspirations and propositions of emerging countries to establish a new international order; therefore, the Five Principles of Peaceful Coexistence are consistent with the provisions of the Charter, and also complement and develop the principles of the Charter, serving as significant contributions derived from Confucianism to the development of international law and to the establishment of a new type of fair and reasonable international relations.[48] Today, the Five Principles of Peaceful Coexistence still serve as the guideline for China in international relations and have been generally accepted by the international community.[49] In the year 1954, the Five Principles of Peaceful Coexistence were formally proposed as norms of international relations.[50] Ever since then, the notion has been put into practice constantly. In 1960, China and Myanmar properly resolved the border issue under the guidance of the Five Principles of Peaceful Coexistence, signing the first border treaty of the People's Republic of China and the first peace and friendship treaty between Asian countries.[51] Till today, the Five Principles of Peaceful Coexistence still plays an important role in the establishment of the new international political and economic order, including the promotion of the Paris Agreement on addressing climate change and the global governance in international relations through comprehensive rule of law.[52]

3. The Contribution of "Benevolence" to The Development of International Human Rights Law

In the development of human rights, especially the contemporary construction of human rights, Chinese culture has its special significance, especially for the Universal Declaration of Human Rights (hereinafter referred to as the "Declaration"), which claims to be a unique product of multicultural integration including Chinese culture. After the establishment of the United Nations, the Commission on Human Rights was designated to draft the Declaration. P. C. Chang, the Chinese representative who was then the vice chairman of the Human Rights Commission, introduced Confucianism to the drafting process of the

---

[48] See Liu Zhenmin, *Following the Five Principles of Peaceful Coexistence and Jointly Building a Community of Common Destiny Editorial Comments*, 13 CHINESE J. INT'L L. 477 (2014).
[49] See Shishi Li, *Continuously Enriching the Contemporary Connotation of the Five Principles of Peaceful Coexistence*, PEOPLE'S DAILY, Jun. 28, 2014.
[50] See Lin Han & Tao Meng, *70 Years of Diplomacy: The Five Principles of Peaceful Co-Existence*, CGTN (2019), https://news.cgtn.com/news/2019-09-23/70-years-of-diplomacy-The-Five-Principles-of-Peaceful-Co-Existence-KdktDEM2Q0/index.html (last visited Sep 28, 2023).
[51] See Guide Jia, *The Practice of International Law and Contributions of Chian in the Seventy Years Since the Founding of the People's Republic of China*, JOURNAL OF INTERNATIONAL ECONOMIC LAW 1 (2020).
[52] See Huikang Huang, *From the Five Principles of Peaceful Coexistence to Building a Human Community with a Shared Future, Contributing Chinese Wisdom to the Reform of Global Governance and International Rule of Law - Commemorating the 50th Anniversary of China's Restoration of Its Legitimate Seat in the United Nations*, JOURNAL OF INTERNATIONAL LAW 1 (2021).

Declaration.[53] He contended that the thought of "benevolence" and "seeking harmony in diversity" of Confucianism means fairness between two people, and shows compassion and understanding of another person's position and viewpoint without giving up one's own position and viewpoint.[54] Based on this, P. C. Chang transformed the above ideas into "conscience" and implemented it in the Preamble, Article 1 and Article 18 of the Declaration. The integration of the expression "conscience" into the Declaration is recognized as an example of the most important thought in Confucianism being introduced into the Declaration. The so-called "conscience" is related to the consideration of justice and morality which is the standard for natural law to judge positive law. It is because the connotation of Confucianism and natural law thoughts are interlinked that Confucianism such as "benevolence" and "seeking harmony in diversity" can finally be transformed into international practices such as the Declaration.

IV. CONTRIBUTIONS OF CONFUCIANISM TO THE FUTURE DEVELOPMENT OF INTERNATIONAL LAW

The theoretical connection between Confucianism and international law, the historical influence of Confucianism on international law and the positive contributions Confucianism has made toward international law have been shown above. However, these connections, influences and contributions are not systematic. As a modern interpretation of Confucianism, the idea of building a human community with a shared future can help integrate Confucianism more systematically into international law and ultimately, to resolve the development dilemma of international law, and to turn the direction of the development of international law back to the track of justice and morality.

A. The Idea of Building a Human Community with a Shared Future: a Modern Interpretation of Confucianism

In March 2018, the idea of building a human community with a shared future was enshrined into the Constitution of China. The idea of building a human community with a shared future requires that one country while self-developing should take into account the legitimate concerns of other countries, and calls on people of all countries to work together to maintain durable peace, universal security, common prosperity, and to build an open, inclusive, clean and beautiful world, so as to safeguard the interests of all countries and their people.[55] These connotations of the idea of building a human community with a shared future are actually modern inheritance, innovation and interpretation of the essence of Confucianism. The idea of building a human community with a shared future emphasizes that each sovereign states is a member of a community with a shared future and thus, should fulfill their obligations according to international law. This is the requirement of "self-restraint". Besides, in accordance with the idea of building a human community with a shared future, each sovereign states should take into account the reasonable concerns of other countries when pursuing their own interests and should not damage the interests of other countries based on their own interests and preference. This reflects the concept of "tolerance" and the pursuit of "moderation" underlined by the "Great Harmonious Society" as well as the "unity of nature and humanity". Applying the idea of building a human community with a shared future, as the modern systematic interpretation of Confucianism, to the context of international law means that all participants in international

---

[53] See FACING HISTORY AND OURSELVES, FUNDAMENTAL FREEDOMS: ELEANOR ROOSEVELT AND THE UNIVERSAL DECLARATION OF HUMAN RIGHTS (2010).
[54] See MARY ANN GLENDON, A WORLD MADE NEW: ELEANOR ROOSEVELT AND THE UNIVERSAL DECLARATION OF HUMAN RIGHTS (First Edition ed. 2002).
[55] See HUIKANG HUANG, CHINA'S DIPLOMACY AND INTERNATIONAL LAW (2019).

law should abide by the corresponding obligations of international law and restrain their own behavior while enjoying their rights. At the same time, "double standards" will be prevented because "moderation" requires impartiality.

B. The Necessity to Apply the Concept of Building a Human Community with a Shared Future to the Development of Contemporary International Law

Different from the zero-sum game of the Western-centered ideology, which excludes others and promotes hegemony, the idea of building a human community with a shared future takes the conflict of interest between countries with different cultures into consideration. The idea calls for common interests, promoting common and harmonious development to achieve common prosperity regardless of differences, which embodies the idea of inclusiveness and integration and the pursuit of world harmony and equality, which is in line with the demand for the unity of powers and responsibilities of sovereign states in contemporary international law.[56] From this perspective, the idea of building human a community with a shared future can provide positive theoretical values for maintaining the just and moral development of international law so as to pursue a fair, just, harmonious and peaceful world order.

C. Application of the Concept of Building a Human Community with a Shared Future to the Development of Contemporary International Law

1. Feasibility of the Application

The concept of building a human community with a shared future is not a fresh start. Since ancient Greece, western philosophers, thinkers and jurists have been exploring different forms of human communities. The "human community with a shared future" is not a "world republic" or a "unified" world government that transcends the state, nor is it an abstract surreal utopia. The idea of building a human community with a shared future emphasizes the rights and responsibilities each and every sovereign states hold for each other. It aims for the establishment of a more fair and harmonious international order that corresponds to justice and morality. Plato, Aristotle, Voltaire and Montesquieu all believe that people cannot achieve self-sufficiency by themselves since the natural gregariousness of humanity leads people to social life, which is the essence for people to survive together; Grotius, the father of international law, also stressed that while pursuing self-interests, we must also recognize the common interests of the international community and the human community in order to live together peacefully.[57] This is consistent with the core idea of Confucianism.

The new natural law school's moral simplification of natural law does not mean that law, especially international law, is just morality without coercion. It shows that moral legitimacy is an integral part of international law in that the assumption of international legal obligations is not only legal, but also moral, which correspond to the "rite" and "law" of Confucianism. At the same time, the humanization of international law emphasizes the protection of the interests of every individual in the international community, which corresponds to the people-oriented approach of Confucianism.

In conclusion, the idea of building a human community with a shared future is not only consistent with sustainable development, peaceful development, people-oriented approach and

[56] See GLENDON, *supra* note 56.
[57] See Zhongfa Ma, *Evolution of Idea of Building a Community of Shared Future for Mankind and Its Implied Thoughts of International Law*, 43 JOURNAL OF LIAONING NORMAL UNIVERSITY (SOCIAL SCIENCE EDITION) 1 (2020).

other international law theories, but also the condensation, generalization and advancement of these international law theories. Therefore, the idea of building a human community with a shared future has the feasibility to turn the development of international law back to the legitimate development of justice and morality and will be conducive to solve the development dilemma of contemporary international law, which is now constrained by the interests of Western countries and deviates from justice and morality.

2. The Application of the Idea of Building a Human Community with a Shared Future to the Development of International Law

When a country joins international organizations or concludes bilateral or multilateral treaties with other countries, it means that the country assumes corresponding obligations in accordance with the articles of association of the organization or the provisions of the concluded treaty, and the country should limit its own rights and freedom in certain aspects. Firstly, in terms of security, all countries should "self-restrain" and attach importance to both of their own security and common security in the development of other countries. To address various security challenges faced by humanity, especially non-traditional security crises, members of the international community should avoid being induced by traditional alliance and Cold War thinking, which goes against the goal of mutual benefit and common security among all parties.

Secondly, in terms of economy and ecology, every country should be tolerate and adhere to the principles of free trade and openness of the international economic system, and prevent deglobalization, which will restrict global trade to countries with similar values and systems, and lead to the "Matthew Effect" that the strong gets stronger and the weak gets weaker, as well as resource monopoly, leading to the impediment of the vitality of global economy and the equal rights to development.

Finally, in terms of ideology, an international morality based on fairness, justice, harmony and peace should be established by means of Confucianism. Communication and cooperation are necessary to prevent "democracy" and "freedom" from becoming tools that only serve the interests of western countries.[58] As members of the human community with a shared future, all sovereign countries in the world have their own inherent interests such as sovereignty, equality and independence, as well as basic obligations such as mutual respect, cooperation, and good faith performance. To reach the "Great Harmonious Society", all countries should avoid sacrificing the legitimate concerns of other countries when pursuing their own development. The fairness, justice and peace of the world order can thus be durable.

CONCLUSION

Globalization itself is a complex combination, and to this day, there are still situations where a few countries especially western countries will use their leading position to seek privileges for themselves or ensure the maintenance of their existing privileges. The influence of Western-centered ideology results in colonialism, culture expansionism, military expansionism and competitiveness, the adverse effects of which drag international law into development dilemma. In order to reduce the impact of "privilege" on the legitimate development of contemporary international law, a universal value system based on common human challenges and different cultures, and a new type of international ideology centered on

[58] See Peigong Liu & Hongze Shan, *The Universal World' to Ideal of 'All Things in One'——On Human Destiny Community in Confucianism and Its Contemporary Value*, 27 HENAN SOCIAL SCIENCES 31 (2019).

mutual benefit and reciprocity to address international issues need to be proposed. Unlike the western ideology of privilege derived from religion resulting in the consequence that international law serves for western privileges, the Confucianism has always focused on the moral obligation of each person living in society. The idea of building a human community with a shared future, as a modern interpretation of Confucianism, can more systematically integrate Confucianism into international law and to resolve the current development dilemma.

"The patriotism is not just about the state with no individual living in it and is not just about a country not involved in the world."[59] Every country in the international law system has the same legal and moral obligation to itself and to other countries and the international community. In this complex and interdependent era, no country can be independent. Confucianism can enrich the idea of international law, form an innovative international moral concept that protects the interests and concerns of all participants in the international community, and promote the international legal system to be more open, inclusive, fair, just and harmonious in the aspect of security, economy and ecology and ideology, so as to promote the peaceful and legitimate development of contemporary international law.

[59] See QICHAO LIANG, A RECORD OF MY TRAVEL IMPRESSIONS IN EUROPE (2014).

THREE MORAL CHALLENGES OF SURVEILLANCE CAPITALISM IN THE METAVERSE

You Zhang*

Abstract: 2021 has been called the first year of the metaverse, which is an independent virtual digital world that is both imitative and transcendent to the real physical world. Many tech giants claim that this technological innovation will bring huge opportunities and dividends to society, but some critics believe that it will also pose challenges to the current social ethics. The important moral challenges may stem from the data issue posed by the metaverse, which is believed to build an unequal relationship between users and service providers due to data-intensive technologies such as VR. This paper argues that the unequal relationships in data have caused the typical consequences of what Zuboff calls surveillance capitalism and posed three major moral challenges to our society, including alienation, exploitation, and domination. According to the Marxist account, alienation and exploitation arise from the existence of digital labor and the monopoly of means of production in the metaverse, while by referring to the Republican account of liberty, the emergence of domination can be attributed to the inequality of data possession.

Keywords: Metaverse; Surveillance Capitalism; Alienation; Domination; Exploitation

* School of Law, Huazhong University of Science and Technology, China.

Table of Contents

Introduction 66

I. Metaverse and Surveillance 66

A. Definition of Metaverse 67

B. Surveillance in the Metaverse 68

C. Metaverse and Surveillance Capitalism 70

II. Alienation in the Metaverse 72

A. Marx's Account of Alienation 72

B. Digital Labor in the Metaverse 73

C. Alienation in the Metaverse 74

III. Exploitation in the Metaverse 77

A. The Marxist Account of Exploitation 77

B. Instruments of Production in the Metaverse 78

C. Exploitation in the Metaverse 79

IV. Domination in the Metaverse 81

A. Freedom of Non-Domination 82

B. Domination in the Metaverse 83

C. Common Knowledge and Domination 86

Conclusion 86

INTRODUCTION

Many old-school people today lament that face-to-face communication has gradually moved away from us, especially during the COVID-19 pandemic of the past three years, when social distancing control policies made more people engage in online interactions. While ambitious businessmen are seizing the opportunity to promote a new concept that had long existed in the works of science fiction --the Metaverse. This concept was originally proposed by novelist Neil Stephenson in his novel *Snow Crash*. And it was subsequently brought to the film screen, the virtual world OASIS in the sci-fi movie "Ready Player One" must have impressed every movie fan, which presents an attractive blueprint for the Metaverse.

2021 has been called the first year of the metaverse. In March of that year, game developer Roblox went public, proposing in its prospectus the concept of the metaverse as a sustainable, shared 3D virtual space, while social media giant Facebook changed its name to Meta in October of the same year, demonstrating Zuckerberg's ambition to participate in the development of metaverse. On one hand, the transformative proposition of the metaverse is thought to bring great convenience to our lives in the fields of education, and healthcare, and will become a platform for leisure, office, and social shared by all of humanity.[1] On the other hand, many scholars have launched critical studies on the risks of the technical aspects implied.[2]

However, the influence of the metaverse would not be limited within the field of science and technology. According to Kranzberg's laws, "Technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves, and technology can have quite different results when introduced into different contexts or under different circumstances".[3] It implies that amid technological development a "digital revolution" is creeping in, which will bring us about new power and political relations. Thus, this paper hopes to provide a preliminary analysis of the moral risks that the metaverse will pose, based on a few political philosophy theories. Since the metaverse is a complex concept, in the limited space of this dissertation I will focus my attention on three potential moral risks posed by the problem of data in the metaverse.

This paper is divided into four sections. Section I introduces the definition of the metaverse and the surveillance capitalism that exists in it. Sections II to IV will discuss how surveillance capitalism that exists in the metaverse will lead to three moral challenges: alienation, domination, and exploitation.

I. METAVERSE AND SURVEILLANCE

There is no consensus on the definition of the metaverse, as it is something that is still in its infancy and "may shift at any time as it continues to be built and used".[4] Similarly, scholars also have different opinions on what challenges the metaverse will bring to our society.

---

[1] See Sang-Min Park & Young-Gab Kim, *A Metaverse: Taxonomy, Components, Applications, and Open Challenges*, 10 Ieee Access 4209, 4225 (2022).

[2] For example, some scholar point out that "AR causes relatively high effort and mental demand compared to conditions without AR." Nanna Xi, et al, *The challenges of entering the metaverse: An experiment on the effect of extended reality on workload*, Inform Syst Front 659, 672 (2022).

[3] Melvin Kranzberg, *Technology and History: "Kranzberg's Laws"*, 27(3) Technol. Cult. 544, 545(1986) .

[4] Jooyoung kim, *Advertising in the Metaverse: Research Agenda*, 21 (3) J. Interact. Advert 141, 142 (2021).

For example, from a health perspective, some scholars argue that the metaverse will have a negative impact on individuals' physiology and psychology; [5] Other scholars are concerned about the accessibility of the metaverse, believing that it will exacerbate the inequality seen by different groups of people;[6] Some scholars have also focused their research on the issue of avatars, arguing that avatars in the virtual world will pose a influence to agent identity.[7] Besides, another controversial issue in the metaverse is the data issue, and existing research generally believes that the metaverse relies on a large number of sensors such as live microphones, cameras, and eye trackers to collect more sensitive information than traditional systems.[8] It puts metaverse users under more comprehensive surveillance than the real world, and this surveillance is believed to bring new power relationships. This paper will focus on the data issues, the ethical challenges of the metaverse due to data, and the surveillance it brings.

A. Definition of Metaverse

Some scholars view the metaverse as "a visual world that blends the physical world and digital world";[9] while best-selling author and businessman Ball considers it as "a massively scaled and interoperable network of real-time rendered 3D virtual worlds"; [10] Legal practitioners Burger-Smid views that "it is a set of interconnected, always-on virtual environments that allow a person to effectively transcend the physical world".[11] Although each of these person has their unique description of the metaverse, by way of summary we can find that they all agree that the metaverse is essentially a digital virtual world that is parallel to the real world but can interact with the real world.

The main disagreement on the definition of the metaverse lies in the characteristics of the metaverse. Mystakidis, for example, describes this virtual world in terms of a perpetual and persistent multi-user environment.[12] While Ball considers this virtual world as a persistent, real-time rendered 3D world.[13] Despite the wide variety of these adjectives, we can distinguish

---

[5] According a survey, "some of the fatigue symptoms related to VR use, such as 'motion sickness, eye strain, headaches, nausea, and dizziness due to heavy VR headset'". Besides, many psychological effects (depression and anxiety, addiction, social isolation, and abstinence from real, physical life) would also occur in users of metaverse. Ghada Refaat El Said, *Metaverse-Based Learning Opportunities and Challenges: A Phenomenological Metaverse Human–Computer Interaction Study*, 12 (6) Electronics 1379, 9 (2023).

[6] For example, some scholars argue that "most VR headsets are currently designed for middle-aged adults, without considering the needs of people in the K-12 age range (4 to 18 years) or older individuals.....The same applies for people with physical or sensory disabilities, but also for those with low levels of digital literacy skill". Matteo Zallio & P. John Clarkson, *Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments*, 75 Telematics Inf 101909, 7 (2022).

[7] *See* Do Yuon Kim et al., *Avatar-mediated experience in the metaverse: The impact of avatar realism on user-avatar relationship*, 73 J. Retail. Consum. Serv. 103382, 1-11 (2023).

[8] *See* Ghada, supra note 5, at 8; Yogesh K Dwived et al., *Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy*, 66 Int. J. Inf. Manage 102542, 8 (2022).

[9] Yuheng zhao, et al., *Metaverse: Perspectives from graphics, interactions and visualization*, 6(1) Vis. Inform 56, 56(2022).

[10] Matthew L. Ball, *Framework for the Metaverse*, MatthewBall.vc (Sep. 3, 2023, 4:23 PM), https://www.matthewball.vc/all/forwardtothemetaverseprimer.

[11] Ahmore Burger-Smidt, *The metaverse and data privacy: Will regulation keep up?*, Werksmans Attorneys (Sep. 3, 2023, 8:23 PM), https://www.werksmans.com/legal-updates-and-opinions/the-metaverse-and-data-privacy-will-regulation-keep-up/#:~:text=The%20metaverse%20is%20a%20set%20of%20interconnected%2C%20always-on,physical%20world%20is%20shared%20amongst%20all%20of%20us.

[12] Stylianos Mystakidis, *Metaverse*, Encyclopedia (Basel, Switzerland) 486, (2022).

[13] Matthew, *supra* note 10.

these features into two categories by deconstructing the term metaverse itself. The root "verse" stands for universe or world, while the prefix "meta" denotes transcendence. Features like Real-time, persistent, and multi-user let the user feel that they are engaging in a complete world rather than a physical world derivation. In other words, the difference between the metaverse and the traditional Internet is that the immersive performance of its content makes users feel that they are in a new space similar to the physical world. Another part of the features makes the metaverse different from the physical world, such as the digital feature, which makes the practice in the metaverse free from the limitations of the laws in the physical world. Therefore, it can be said that the metaverse is a transcendence of the physical world.

In short, this dissertation defines the metaverse as an independent virtual digital world that is both imitative and transcendent to the real physical world and can interact with the real world.

B. Surveillance in the Metaverse

The existence of the metaverse and the novel experiences it brings us to depend on some key technologies or infrastructures. "The Metaverse is an end-users-oriented integration of various layers of Information Technology (IT), where Human–Computer Interaction (HCI) would be the core technology."[14] The main technologies include Extended Reality (XR),[15] the Internet of Things (IoT),[16] cloud computing,[17] and 5G communication.[18] But those technologies are also considered data-intensive, which means that their application will expand the scope of data collection and increase the amount of data collected. According to Bailenson's research, using a VR device for 20 minutes would leave 2 million unique body language records.[19] Due to the embodied nature of these devices, the data they capture is bound to have strong personal attributes; While Cloud computing makes the collection, processing, and application of these data out of the control of users.[20] So many scholars believe that this will pose a huge threat to personal privacy, for example, Wang argues that there may be a risk of

[14] *See* Ghada, *supra* note 5, at 1.
[15] Extended Reality (XR) is one of the most crucial technologies in the metaverse, which can be subdivided into Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) technologies. *See* Stylianos, *supra* note 12, at 487.
[16] Haptic Internet, a new type of IoT technology, is considered to have an important place in the metaverse. It is a project that allows humans and machines to interact with their surroundings in real time through haptics while on the move and within a specific spatial communication range. *See* Gerhard P. Fettweis, *The Tactile Internet: Applications and Challenges*, 9 (1) Ieee Veh. Technol. Mag. 64, 64-70(2014).
[17] "In industrial Metaverse applications, massive computational resources are consumed to build physically accurate simulation environments. Prospective consumer applications will also challenge computing power requirements..... An explosively growing number of digital assets are crowding into the Metaverse". In this context, cloud computing or cloud network flow technology can effectively reduce the computing and storage burden of local devices. *See* Yang Cai et al., *Compute- and Data-Intensive Networks: The Key to the Metaverse*, 2022 1st International Conference on 6G Networking (6GNet) 1, 1-8(2022).
[18] Due to the large communication load generated by the above technologies, fast and stable data transmission technology is essential for the metaverse. 5G technology is considered a necessary infrastructure for the metaverse, "5G wireless technology is the 5th generation of wireless technology. It provides wireless communication with almost no limitations." Simar Preet Singh et al., *Fog computing: from architecture to edge computing and big data processing*, 75 J. Supercomput 2070, 2082 (2019).
[19] *See* Jeremy Bailenson, Experience on Demand: What Virtual Reality Is, How It Works, And What It Can Do, n.p. (1st edition, 2018).
[20] It has been criticized for building a "digital enclosure" to prevent users from accessing their data. *See* Mark Andrejevic, *Meta-Surveillance in the Digital Enclosure*, 20 Surveil Soc. 390, 390(2022).

privacy invasion during the collection, transmission, processing, and storage process.[21]

In the metaverse, user data is continuously obtained and utilized by service providers through various sensors, which is summarized by contemporary scholars as "surveillance". The so-called surveillance refers to: "scrutiny of individuals, groups, and contexts through the use of technical means to extract or create information";[22] Or "gathering of some form of data connectable to individuals".[23] One of the most famous theories of surveillance is Foucault's so-called panopticism, which was used to analyze 17th-century models of social discipline.[24] As technological, business, and policy changes accelerate the flow of information, concerns about surveillance and privacy increase. In this context, Lyon proposed the concept of data surveillance, arguing that such surveillance would not only limit people's freedom but also create long-term social differences.[25] Since then, Zuboff has proposed the theory of surveillance capitalism based on the development of big data technology, analyzing how surveillance constitutes a new relation of production.[26] In the metaverse, this surveillance as a breach of privacy will continue and become a more serious problem. There are two main reasons for this.

Firstly, the object of surveillance has changed. In the real world, the subject being monitored is a physical person, while in the metaverse it is monitored both the physical person and the user's avatar in the virtual world. For real-world surveillance objects, due to the application of VR, AR, and other data-intensive technologies, the amount of data obtained by the metaverse has greatly increased, and the scope of data collected has also expanded. Egliston argues that those devices would collect the data including the user's behavioral characteristics like head and eye movement, the user's pulse, and breathing, and may also include the data of the surrounding environment.[27] Therefore, the metaverse will lead to an increase in the intensity of surveillance of users in the real world. Unlike the real world, where human information needs to be captured by sensors, the behavior of avatars in the metaverse is composed of data and inevitably leaves digital traces. While real-world individuals will be able to escape surveillance temporarily, avatars in the metaverse will be under ubiquitous surveillance by what Zuboff calls the Big Other. In other words, due to the data nature of the avatar, its monitoring can be continuous and uninterrupted.

Secondly, the means of surveillance have changed, and surveillance in the metaverse era is more automated, which has led to a reduction in the cost of surveillance. During the Cold War, the GDR's Ministry of State Security (also known as Stasi) was considered one of the most efficient intelligence services in the world, reportedly employing 274,000 people to Maintain its domestic surveillance system. But in the metaverse, the efficiency of data collection is greatly improved with the help of big data technology, which allows Meta to maintain its global business with only 70,000 employees. Thus, technological disruption usage

---

[21] Yuntao Wang et al., *A Survey on Metaverse: Fundamentals, Security, and Privacy*, 25 Ieee Commun. Surv. Tutorials 319, 334-335 (2023).
[22] Gary T. Marx, Windows into the Soul: Surveillance and Society in an Age of High Technology, 20 (2016).
[23] Gary T. Marx, *Surveillance Studies*, International Encyclopedia of the Social & Behavioral Sciences 733, 733 (2015).
[24] *See* Michel Foucault, *Discipline and Punish: The Birth of the Prison*, 195-230 (2nd ed. 1995).
[25] *See* David Lyon, *Surveillance as Social Sorting: Privacy, Risk and Digital Discrimination*, 1-28 (2002).
[26] *See* Shoshana Zuboff, *Big other: Surveillance Capitalism and the Prospects of an Information Civilization*, 30 (1) J. Inf. Technol. 75, 75-89 (2015).
[27] *See* Ben Egliston & Marcus Carter, *Critical questions for Facebook's virtual reality: data, power and the metaverse*, 10 (4) Internet Policy Rev. 1, 9 (2021).

of new tools has just brought new "material artifacts, software, and automated processes" for observation.[28] Meanwhile, AI and big data technology have also brought about the reduction of surveillance costs, and when surveillance is included in commercial operations, this cost reduction will inevitably lead to an increase in its application. The next sub-section will discuss how commercially operated surveillance will reshape production relations.

C. Metaverse and Surveillance Capitalism

Ubiquitous surveillance in the metaverse not only means that citizens' privacy is at risk, but it can lead to a more far-reaching influence on social and economic systems. Current research generally refers to this new system as "surveillance capitalism" or "data capitalism". Zuboff describes the former as a "new form of information capitalism that aims to predict and modify human behavior as a means to produce revenue and market control".[The] latter has been described as "a system in which the commoditization of our data enables an asymmetric redistribution of power that is weighted toward the actors who have access and the capability to make sense of information".[29] Taken together, these two similar theories (hereafter use "surveillance capitalism" to refer to them) aim to describe the fact that personal information has become a resource of economic value, and on this basis form a new capitalist operating model that profits through the "appropriation and commercialization of personal data".[30]

A further description of surveillance capitalism can be developed from the premise and mode of operation of its emergence. Data resourceization is the main premise of surveillance capitalism, which means that data has become a new means of production, which can be used to produce and create wealth on the one hand, and to gain political power on the other hand. Economically, the commercial value of data has been proven by many studies, such as a study that found that DDEM (Data-Driven Marketing Economy) added $156 billion in revenue to the U.S. economy and fueled more than 675,000 jobs in 2012 alone.[31] Other scholars estimate that the value of the 2.1 trillion pieces of monetizable content-personal data at the time of Facebook's IPO was about 5 cents per data point or around $100 per user.[32] The asymmetry formed in data control will also lead to the asymmetry of knowledge and power[33] since communication and information have been historically recognized as key sources of power.[34] Like factory machines and financial capital, personal data has become the means of production that generates profit and power in surveillance capitalism.

Zuboff describes the operation of surveillance capitalism in her book. When someone uses Google as a search engine, this produces a "behavioral surplus", this collateral data created by the users when they use the engine, including the keywords, the number and pattern of search terms, how a query is phrased, spelling, punctuation, dwell times, click patterns, and

---

[28] Andrew B. Whitford, *Surveillance and privacy as coevolving disruptions: reflections on "notice and choice"*, 6 (1) Policy Des. Pract. 13, 16 (2023).
[29] Sarah Myers West, *Data Capitalism: Redefining the Logics of Surveillance and Privacy*, 58(1) Bus. Soc. 20, 20 (2019).
[30] Bernd Carsten Stahl et al., *Surveillance Capitalism*, Ethics of Artificial Intelligence 39, 39(2023).
[31] John A. Deighton & Peter A. Johnson, *The Value of Data: Consequences for Insight, Innovation & Efficiency in the U.S. Economy*, 77 (2013), https://www.ipc.be/~/media/documents/public/markets/the-value-of-data-consequences-for-insight-innovation-and-efficiency-in-the-us-economy.pdf.
[32] Jonathan Cinnamon, *Social Injustice in Surveillance Capitalism*, 15(5) Surveill Soc. 609, 614 (2017).
[33] Shoshana Zuboff, The Age of Surveillance Capitalism: The Fight For a Human Future at the New Frontier of Power, 311(2019).
[34] *See* Manuel Castells, *Communication, Power and Counter-power in the Network Society*, 1 Int. J. Commun 238, 238(2007).

location.[35] Companies like Google and Facebook "recognize the gold dust in the detritus of its interactions with its users and took the trouble to collect it up from what is often referred to as the 'data exhaust' ".[36] This data can be analyzed to provide a detailed profile of each user so that it has the ability to accurately predict users' behavior and judgment. Finally, this processed data is sold as a product to advertisers and helps them accurately advertise to consumers. [37] Advertisers are willing to buy this data because it could "nudges" —deliver a particular message to a particular person at just the moment when it might have a high probability of actually influencing his or her behavior—to make the consuming decisions advertisers expect.[38]

The metaverse has become a hotbed for surveillance capitalism because of the increased bargaining power of service providers in it. Due to increased regulation stealing user and personal information without a bottom line is no longer a good option. The process of obtaining personal information by the service provider is a transaction process, the user obtains the service for free or at a lower price, and the service provider obtains the user's personal information as a price.[39] In the real world, people need to provide less information and have more alternatives, such as the option to use cash if they do not wish to leave a record of their card purchases. But in the metaverse, these alternatives disappear, if you decline this deal, you may decline metaverse service.

Furthermore, due to "network effects", the more services the metaverse can offer and attract more customers, the greater people's dependence on it. When most people choose to accept the service, many of life's necessary activities move to the platform, this makes the rest of the people have to join to survive. With the development of the metaverse, it will involve almost all aspects of life, entertainment, office, etc., it is foreseeable that it will become an important public place for people's social participation in the future, and the lack of social participation is considered to have a negative impact on personal physical and mental health.[40] So in this case, metaverse service providers have gained monopoly-like status, making their bargaining power with users greatly enhanced, and many users have to make deals with it despite not liking its data policy.

This section argues that surveillance, and surveillance capitalism, will continue to exist and be further strengthened in the metaverse. This fact also implies that the metaverse has many moral challenges, and the following discussions will be made around three of the main moral hazards – alienation, exploitation, and domination.

---

[35] *See* Paul Michael Garrett, *'Surveillance Capitalism, COVID-19 and Social Work': A Note on Uncertain Future(s)*, 52(3) Brit J. Soc. Work 1747, 1749(2022); Also *see* Shoshana *supra* note 33, at 67.
[36] Shoshana, *supra* note 33, at 68.
[37] According to a study, "Google ads were wildly successful as a means for monetizing the company's search business: By perfecting an auction model for pricing and selling ads, its revenues grew rapidly with year on year growth rates in advertising revenues of 514% in 2002 and 246% in 2003." Sarah *supra* note 29, at 32.
[38] *See* Shoshana, *supra* note 33, at 77-78; Aron Darmody & Aron Darmody, *Manipulate to empower: Hyper-relevance and the contradictions of marketing in the age of surveillance capitalism*, 7 (1) Bd & S. 205395172090411, 3(2020); José Van Dijck, *Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology*, 12(2) Surveill Soc. 197, 200(2014).
[39] *See* Shoshana *supra* note 33, at 212.
[40] Nikhil Venkatesh, *Surveillance Capitalism: a Marx-inspired account*, 96(3) Philosophy 359, 381-382(2021).

II. ALIENATION IN THE METAVERSE

One of the risks that surveillance capitalism creates in the metaverse is alienation. In the Marxist critique of capitalism, labor alienation is an influential argument, which Marx elaborated in his *Manuscript of Economics and Philosophy in 1844* As a form of capitalism, surveillance capitalism may also have alienation problems, but its specific form should vary according to the social context. Therefore, this section will also analyze the hidden alienation risks in the metaverse in combination with the "digital labor" theory proposed by contemporary scholars.

A. Marx's Account of Alienation

The general concept of alienation is considered to be a description of a social or psychological problem, "a problematic separation between a self and other that belong together".[41] The word alienation originated from the Latin noun alienatio, which means "take away", or "remove".[42] This concept was originally used to refer to an individual's estrangement from God, the legal transfers of ownership rights, and mental derangement, and later it was first introduced to philosophical discussion by Social contract theorists.[43] Inspired by Hegel, who is considered to be the creator of the concept of alienation that we are familiar with,[44] Marx argued that "man does not experience himself as the acting agent in his grasp of the world, [the world] stand above and against him as objects, even though they may be objects of his own creation".[45] In Marx's concept of alienation, there are two main contents, the first is objectification, where part of the subject is separated from the subject and becomes the object; The second is domination, where the external object has power over the subject.

In the *1844 manuscript,* Marx discussed in detail how alienation manifests itself in labor and division of labor. He argued that the capitalist division of labor would lead to the problematic separation of man as subject from four objects: the product, the process of labor, the species-being, and fellow humans. It has been argued that there is an internal causal relationship between these four alienations and that the alienation of the labor process is the initial cause of the other three alienations.[46] Marx believed that in the process of labor, "labor is an external thing to the worker...he does not affirm himself but denies himself in labor...to torture himself physically and destroy his spirit".[47] The proletariat, who has gained formal freedom in capitalist society, has nothing but their labor, so they can only choose to sell their labor to capitalists for a living.[48] They can only engage in mechanical and tiring labor when the process of labor is no longer a free and self-fulfilling process for the proletariat, but out of their control and becomes a tormenting force of aliens.

The direct consequence of the alienation of the labor process is the alienation of the

---

[41] David Leopold, *Alienation*, The Standford Encyclopedia of Philosophy, (Sep. 7, 2023, 11:18 AM), https://plato.stanford.edu/entries/alienation/.
[42] Hamid Sarfraz, *Alienation: a theoretical overview*, 12 Pak. J. Psychol. Res. 45, 45(1997).
[43] Those theorists included Grautius, Hobbes, Locke, and Rousseau. *See* David *supra* note 41; *Id.* at 46.
[44] *See* Erich Fromm, Marx's Concept of Man 47(2004).
[45] *Id.* at 44.
[46] *See* Xiaomang Deng, *laodong yihua jiqi gengyuan [Labor alienation and its source]*, 1983(03) Zhongguo Shehui Kexue 155, 155(1983).
[47] Karl Marx, Karl Marx: Selected Writing 85 (David McLellan ed. 2nd ed.2000).
[48] *See* G. A. Cohen. On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy, 163–182 (1st ed. 2006)

products of labor. Marx argues that in the wages labor system, "wages, therefore, are not a share of the worker in the commodities produced by himself. Wages are that part of already existing commodities with which the capitalist buys a certain amount of productive labor-power".[49] This means that the product of labor condensed by the labor of the worker does not belong to the worker, but becomes a separate object.[50] Moreover, the product of labor became a dominant force hostile to the workers. Marx argued that "[t]he worker becomes all the poorer the more wealth he produces..... The devaluation of the world of men is in direct proportion to the increasing value of the world of things".[51] These products flow back into the hands of workers in the form of commodities, who have to spend money to buy the goods they need to survive, demonstrating the domination of the product over its subject.

The remaining two forms of alienation are the result of alienated labor. Non-alienated labor is a "free conscious activity",[52] and this is a characteristic of human beings, hence what Marx called the species-being of human beings. In capitalist society, workers cannot freely and consciously participate in labor due to the pressure of survival, and can only use labor as a means of maintaining physical survival. In other words, alienated labor becomes an obstacle to the development of unique human abilities. Meanwhile, once the worker sees his species-being as a means to an end, he necessarily sees others as a means to an end. Capitalists see workers as a means to profit, while workers must compete with each other for the chance and price of labor.[53] So the otherwise united humans become mutually isolated individuals, and each individual opposes the others to increase their own interests.

B. Digital Labor in the Metaverse

Marx's above statement is based on his observation of industrial capitalist society in the 19th century, but it seems difficult to find an obvious wage labor relationship between the platform and users of the metaverse, which seems to mean that the theory of labor alienation is difficult to apply in it. Users seem to act as consumers rather than a producer in the metaverse, since they exchange money for goods and services entering the metaverse, such as buying VR equipment, buying special decorations for their avatars, and so on. But as far as today's free social media situation is concerned, social media platforms are very welcoming to more users to accept their free services, which seems to mean that the platform is profitable in the process of users enjoying online services. Some scholars refer to this user online activity as "Internet prosumption",[54] or "consumption work",[55] to reveal the two-way interest relationship between users and platforms.

Fuchs uses the concept of "digital labor" in its book to refer to the process of users creating value for Internet platforms.[56] Its labor takes the form of unpaid activities of users, such as generating and sharing content, interacting with others, and contributing data. From the

---

[49] Karl Marx, *Wage Labour and Capital*, 45 (2) Ninet. Century Prose 319, 320(2018).
[50] *See* Dan Smith Swain et al., The Oxford Handbook of Karl Marx, 363(2019).
[51] Karl, *supra* note 49, at 86.
[52] Karl, *supra* note 47, at 90.
[53] *See* Jan Kandiyali, *The Importance of Others: Marx on Unalienated Production*,130(4) Ethics 555, 562(2020).
[54] *See* George Ritzer & Nathan Jurgenson, *Production, Consumption, Prosumption: The nature of capitalism in the age of the digital 'prosumer'*, 10(1) J Consum Cult. 13, 13(2010).
[55] *See* Ursula Huws, *The reproduction of difference: gender and the global division of labour*, 6(1) Work Organisation, Labour and Globalisation 1, 4(2012).
[56] *See* Christian Fuchs, Digital Labour and Karl Marx 246 (1st ed. 2013).

accumulation logic of surveillance capitalism above, we can know what the Internet platforms have in common is that "they use a business model that is based on targeted advertising and that turn users' data (content, profiles, social networks, online behavior) into a commodity."[57] These data as commodities must have their producers, and these data are records of the user's online activities, which could be recognized as the production process of their personal data. Digital labor takes ideas and human subjectivity as the object, users externalize it through action or choice, and it is fixed in the form of personal data formed as a product.

Digital labor will be widespread in the metaverse for two main reasons. First, the metaverse aims to build a durable, and immersive virtual space, which means that it is designed to entice users to spend more time in it with more concentration. According to a survey, in Web 2.0, the time people access the Internet through mobile phones or computers has reached about 6 and a half hours per day in 2022.[58] The metaverse is believed to be more attentive to users than Web 2.0, and some scholars have proposed the concept of immersive time (ImT), which is the conscious, deliberate, and dedicated time spent using a headset and other accessories to continually engage in the metaverse.[59] When a user surfing the traditional Internet, he is still able to engage in other activities and multitasking (like talking with others), while the user's metaverse experience is almost completely disconnected from the real world and they could focus solely on the virtual experience. It is precisely because of this characteristic that many people who are disappointed with the real world choose to spend more time in the metaverse to escape reality.[60]

Second, the metaverse has the idea that it should be decentralized and emphasize user-generated content (UGC). These characteristics mean that metaverse users are not only experienced in this virtual world but should be builders or even regulators. Therefore, metaverse users will be encouraged to participate in the creation of metaverse content more frequently, which increases the actual data labor workload of metaverse users. In short, both immersive experiences and UGC in the metaverse make users spend more energy in the metaverse. While users experience the metaverse, they also participate in digital labor, and in the process, there is also a risk of alienation due to the separation of personal data from their creator, which will be discussed in detail in the next section.

C. Alienation in the Metaverse

First of all, the process of users logging in and experiencing the metaverse in the metaverse replaces the traditional labor process. Therefore, the alienation of digital labor in the metaverse originates from this alienation of metaverse experience. The first thing to acknowledge is that there is a non-alienated metaverse experience, which is an experience driven by individual free will and self-realization.[61] For example, when we crave social, we

---

[57] *Id*. at 247.
[58] *See* Simon Kemp, *Digital 2023: Global overview report*, Datareport, (Sep 8. 2023, 10:44 AM), https://datareportal.com/reports/digital-2023-global-overview-report.
[59] *See* Emmanuel Mogaji et al., *Immersive time (ImT): Conceptualizing time spent in the metaverse*, 72 Int. J. Inf. Manage. 102659, 2(2023).
[60] *See* Andrew Kuo et al., *Brave new World of Warcraft: a conceptual framework for active escapism*, 33(7) J. Consum. Mark. 498, 498(2016); Dai-In Danny Han et al., *Virtual reality consumer experience escapes : preparing for the metaverse*, 26(4) Virtual Reality 1443, 1443-1458(2022).
[61] Kandiyali define fully unalienated production as "work that (1) involves self-realization, that is, the exercise, development, and manifestation of our individual powers; (2) satisfies another's need; (3) is conducted with the

can chat with people from all over the world through the metaverse platform, and when we need entertainment, we choose to immersively experience the variety of games in the metaverse.

However, marketing practitioners have recognized that the more people stay online, the more this surfing time gives them access to data about consumers, and the more they can serve consumers with marketing messages.[62] The current operators of social media and online games have adopted this strategy. For example, *Fortress Night*, which is considered to be one of the video games closest to the metaverse, is set up with daily tasks. Players can get rewards if they log in to the game every day and complete certain tasks, which incentive ensures the number of daily online players. There is a study that predicts that intensifying the feeling of embodiment in the metaverse will make the metaverse more addictive than the traditional Internet,[63] which will make the metaverse experience no longer guarantee the self-realization or autonomy of the user. Addiction not only frees the metaverse experience from the control of the subject, but also has many adverse effects on the subject's physical and mental health, like mental illnesses such as isolation and anxiety, and physical health issues such as cyber sickness, eye soreness and trouble focusing, impaired hand-eye coordination, reduced depth perception, increased reaction time, loss of balance, and prolonged nausea.[64]

The alienation of individual experiences in the metaverse will directly lead to the alienation of their products – personal data. Users would generate vast amounts of personal data in the metaverse, much of which is described as "exhaust" because it is useless to users. However, Venkatesh believes that these personal data, such as users' likes on Facebook, search records on Google, etc., have become a digital commodity with huge economic value.[65] Huge profits drive internet companies to grab a lot of personal information, just as Zuboff argues that "Google knows far more about its users than they know about themselves".[66] Consequently, this sensitive data is out of the control of its subject and becomes the asset of the data company, with Facebook reportedly owning $97.7 billion in assets in 2012 in the form of 2.1 trillion pieces of user personal data.[67] Even more troubling is that in the endless data grabbing, Internet platforms have acquired asymmetrical control over personal information, which allows platforms to provide us with more sophisticated and personalized advertising[68]. Many users are unconsciously influenced by personalized advertising to make decisions that benefit advertisers. Therefore, personal data in the metaverse is not only separated from its subject but increasingly becomes a force opposed to its subject.

The alienation of digital labor in the metaverse will also lead to the alienation of species-being. This means that many times the user's experience of the metaverse does not develop its ability as a quasi-essence, but only as a means to other ends. From the previous discussion, digital labor is a kind of labor like agricultural farming and industrial manufacturing,

---

intention of satisfying another's needs; (4) is used and appreciated by that other; and (5) is performed freely". Jan, *supra* note 53, at 571.

[62] *See* Emmanuel, *supra* note 59, at 4.

[63] Miguel Barreda-Ángeles & Tilo Hartmann, *Hooked on the metaverse? Exploring the prevalence of addiction to virtual reality applications*, 3 Front. Immunol. 1, 1-9(2022).

[64] *See* BEIS, *The safety of domestic virtual reality systems: A literature review*, Gov.Uk, (Sep 9. 2023, 4:51 PM.), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/923616/safety-domestic-vr-systems.pdf.

[65] *See* Nikhil, *supra* note 40, at 366-367.

[66] Shoshana, *supra* note 26, at 83.

[67] *See* Jonathan, *supra* note 32, at 614.

[68] *See* Mark Andrejevic, *Surveillance in the Digital Enclosure*. 10 (4) Communication Review 295, 313 (2007).

so Focus argues that "information creation is itself a work process..... digital work on social media is a specific form of informational work".[69] Just like labor, experiencing the metaverse freely and consciously is considered non-alienated participation, receiving medical training, distance education, etc. are believed to foster a relaxed and creative learning culture that is inclusive, active, and experimental,[70] as well as improves cognitive processing, mental elaboration, and imagery ability.[71] However, social demands and fears of isolation often push users to experience the metaverse as a means rather than an end. In an information society, digital media has become an important means of communication for many people. "Isolation of an individual from communication and social networking will ultimately result in either death or an animal-like existence."[72]

Finally, users' alienation from their fellow humans is also a broader social consequence in the metaverse. This happens both between metaverse service providers and users, as well as between users and users. As far as the former is concerned, Zuboff monitors the relationship between surveillance capitalists and users as a "formal indifference".[73] Because in surveillance capitalism, surveillance capitalists only need to hire a few technicians, they do not depend on the support of a large number of users, so they do not have to pay as much attention to the interests of users as traditional capitalists consider the welfare of workers. In addition, in Focus's case study of Facebook, he noted that "alienation of the instruments of labor also means in the context of Facebook that the users do not own and control the platform".[74] Therefore, in the metaverse, the reciprocity between the platform and the user will be weakened, and the user will become a means for the platform to make profits.

In the case of the latter, since everyone can shape themselves through their digital avatars and digital properties, the relationship between users in the Metaverse will be hidden behind the veil of data. This brings us to Marx's theory of commodity fetishism, since commodities and people have a direct social connection, while people and people are indirectly connected, the relationship between people is obscured by the relationship between people and things.[75] Data in the metaverse is likely to become a new fetish due to its status as a necessary medium for communication, and the alienation between humans behind it is believed to cause addiction, social isolation, and abstinence from real, physical life.[76] A Metaverse user with gorgeous fashion may be sought after by everyone. And humble virtual dwellings may be despised by neighbors. This will cause many people to compare their virtual properties, such as the fanatical worship and pursuit of game currency or game equipment by players in online games.

In summary in a metaverse that is still fostered by surveillance capitalism, users will inevitably face alienation from their data, their behavior, their species-being, and fellow humans. Such alienation would undoubtedly be a moral risk of the metaverse itself and hinder

---

[69] Christian, *supra* not 56, at 250, 254.
[70] *See* Stylianos, *supra* note 12, at 490.
[71] *See* Vanja Bogicevic et al. *Virtual reality presence as a preamble of tourism experience: The role of mental imagery*, 74 Tourism Manage. 55, 56(2019).
[72] Christian, *supra* not 56, at 254.
[73] Shoshana, *supra* note 26, at 76.
[74] "After Facebook's initial public offering, its 12 executive officers and directors controlled together 61.1% of the class B stock (Facebook Registration Statement, Form S-1)." Christian, *supra* not 56, at 256.
[75] *See* G. A. Cohen, Karl Marx's Theory of History: A Deference, 115-133(1979).
[76] *See* Mel Slater et al., *The Ethics of Realism in Virtual and Augmented Reality*, 1 Front. Virtual Real 1. 7-8 (2022).

the pace of human self-realization and prosperity.

III. EXPLOITATION IN THE METAVERSE

Surveillance capitalism in the metaverse will not only cause the challenge of alienation but also have another major drawback of capitalist society, that is, economic exploitation between opposing classes. Marx constructed his theory of exploitation through the theory of surplus value, which this section combines with the aforementioned theory of digital labor to discuss the potential risks of exploitation in the metaverse.

A. The Marxist Account of Exploitation

Exploitation is a very controversial concept, with Kymlicka arguing that the ordinary definition of exploitation is "taking unjust advantage of others".[77] However, different political philosophical theories have different views on the definition of unjust. It has also been argued that exploitation does not necessarily mean injustice, and that exploitation is unjust only when it is conditioned by other conditions -- such as violating the freedom or other rights of others.[78] Vrousalis summarizes the general structure of exploitation as "A exploits B if and only if:(1) A benefits, (2) from a social relationship with B, and (3) by taking advantage of B".[79] This deconstruction can be applied to most accounts of exploitation, including the Marxist accounts of exploitation referred to in this section.

Marxist account of exploitation refers specifically to the value (in the form of a product) extracted by the capitalist from the worker's labor, over and above the remuneration (in the form of a wage) for the worker's labor.[80] Marx argues that in the commodification of labor in a capitalist society, workers are free to sell their labor, so workers and capitalists are in a relationship of wage labor. In addition, capitalists only employ workers when they can extract "surplus value" from them, so the capitalist can obtain the benefits produced by the workers in the labor. Finally, taking advantage means "A does that by taking advantage of certain important features of B, features that are central to B's person, her life or well-being".[81] The reason this non-reciprocal surplus value transfer can exist in the free exchange of labor markets is that the "worker, whose only source of income is the sale of his labor-power, cannot leave the whole class of buyers, i.e., the capitalist class".[82] Returning to Vrousalis' structure, Marx's account of exploitation can be summarized as follows: capitalist gain surplus value from wage labor, by taking advantage of worker's lack of means of production.

Of the above three points, in addition to the fact that the employment relationship is easier to understand, the other two points are worth further description. First, the theory of surplus value argues that the worker's one day can be divided into two parts, The worker gets paid for his labor in the first part of the day in the form of a wage, when a reciprocal relationship is maintained between the worker and the capitalist. In the rest of the day, however, he keeps

---

[77] Will Kymlickam, Contemporary Political Philosophy : An Introduction, 178 (2nd ed. 2002).
[78] *See* Allen Wood, *Unjust Exploitation*, 54 (S1) South. J. Philos 92, 92-93(2016).
[79] *See* Nicholas Vrousalis, *Exploitation: A primer,* 13(2) Philos. Compass.e12486, 1-14 (2018).
[80] Will, *supra* note 77, at 177-187.
[81] Nicholas, *supra* note 79, at 2.
[82] Karl, *supra* note 49, at 9.

working but gains no pay. His labor during this time is called "surplus labor".[83] The value of the product produced in the "surplus labor" is called "surplus value". Exploitation is the process of transferring this "surplus value" from the worker to the capitalist without compensation, and this transfer is considered to be non-reciprocal -- it is A receiving something from B without giving an equivalent in return.[84] Thus, the Marxist account of exploitation holds that capitalists benefit from surplus value.

Marx believed that, in fact, workers' labor under capitalism was neither truly voluntary nor entirely for the benefit of the workers themselves. It is not truly voluntary, because workers are forced. After all, they lack ownership of the means of production.[85] The means of production are all the material conditions necessary for people to engage in the production of material materials, that is, the sum of the means of labor and the objects of labor, which can generally include land, plant, machinery and equipment, tools, raw materials, and so on. Productivity can only be generated if labor and means of production are combined. In capitalist society, workers control only their own labor, while the means of production are monopolized by capitalists, so workers cannot complete production alone, but need to sell their labor power.[86] Thus, capitalists enjoy an advantage in the labor market by virtue of their possession of the means of production.

B. Instruments of Production in the Metaverse

Although we discussed earlier that there is a new form of labor in the metaverse - digital labor, this is still not enough to prove that there is Marx's account of exploitation in the metaverse. Certain conditions need to be met for exploitation to exist, and Roemer lists three points in his paper: (1) unequal ownership of the capital stock, (2) labor markets, and (3) scarcity of capital relative to the labor available for employment.[87] From the previous discussion, similar to the traditional labor market, capitalists buy workers' labor with wages, and in the metaverse, surveillance capitalists exchange users' data with metaverse services. Both labor and personal data are considered commodities and are formally freely traded in both markets. Therefore, there is a personal data market in the metaverse, which satisfies Roemer's second condition. The first and third conditions seem less clear in surveillance capitalism, and it deserves a longer discussion.

In Roemer's model, capital combined with labor produces products -- corn in his case. In Marx's account, the productive forces are the unity of means of production and labor, which includes three elements: all labor, instruments of production; objects of production. Thus, what Roemer calls inequality and relative scarcity of capital stock can also be considered inequality and scarcity of means of production. In the metaverse, if the user's ideas and human subjectivity as the object of labor, the user's experience in the metaverse is a labor process, and the personal data with advertising value as the product, then we still omit an important element in the description of the production in the metaverse, that is, the instruments of production. The large amount of raw personal data generated by users is called digital exhaust because they cannot directly generate value, and only large amounts of aggregated processed personal data

[83] Matt Zwolinski et al., *Exploitation*, The Stanford Encyclopedia of Philosophy, (Sep 8. 2023, 10:05 PM) , https://plato.stanford.edu/archives/sum2017/entries/exploitation/.
[84] *See* Nicholas, *supra* note 79, at 4.
[85] Matt, *supra* note 83.
[86] Allen, *supra* note 78, at 102.
[87] John E. Roemer, *Socialism Revised*, 45 (3) Philos. Public Aff. 261, 268 (2017).

have commercial value. Like the machines that mine ore and the furnaces that smelt it, the sensors needed for data collection and the computing power needed for data processing are the instruments of production in surveillance capitalism.

Data collecting and processing capacity will be very unequal between metaverse service providers and users. For example, data processing capacity depends on the one hand on data centers with millions of "virtual servers" and the other hand on data scientists who master new methods related to predictive analysis, reality mining, life pattern analysis, etc.[88] These means of production thus require large upfront investments, and only a small number of Internet giants have this capability. In addition, this ability is also considered relatively scarce. Some scholars argue that in surveillance capitalism society is divided into three "data classes": those who create data, those who collect it, and those who can analyze it.[89] The analysis class is a very lean team of people who have the technical expertise to engage in personal data analysis. While computer coding and data analysis skills have always been scarce area, their size has shrunk dramatically with the rapid rise of big data and the development of new tools and algorithms to handle large unstructured data sets.[90] Therefore, in the metaverse, surveillance capitalists unequally possess relatively scarce data collection and analysis capabilities, which allows them to meet Roemer's remaining two conditions.

C. Exploitation in the Metaverse

In the first two sections, we discussed the Marxist account of exploitation and the conditions that support exploitation in the metaverse. This section will build on previous discussions and discuss specific forms of exploitation in the metaverse, which is the process of transfer of surplus value. As mentioned earlier, the source of surplus value in the capitalist system of production is the unpaid labor of workers. Accordingly, in the current information society, many scholars argue that there exists "digital free labor",[91] and "unpaid jobs".[92] This form of labor is considered an important means of exploitation in advanced capitalist societies. Because Internet users engage in many value-creating activities without being commensurately remunerated. These free digital labors include: accessing the site, writing messages, participating in conversations", free labor of Internet tinkers in the "open source" movement,[93] or searching, discovering, and evaluating city locations in Google Maps.[94] By this, users consume their own private time to generate data with economic value. Personal data can be used for advertising and UGC could attract more people to visit.

While the user's experience creates tremendous value for service providers, they could not enjoy the benefits directly. Cookie technology is a widely adopted technique in user data acquisition practices, first introduced in Netscape's Navigator 1.1 browser, which enables

---

[88] Shoshana, *supra* note 26, at 80.
[89] *See* Lev Manovich, *The Promises and the Challenges of Big Social Data*, Debates in the Digital Humanities 460, 470 (2012); Mark Andrejevic, *The Big Data Divide*, 8(1) Int. J. Commun 1673, 1673 (2014).
[90] *See* Thomas H. Davenport & D.J. Patil, *Data Scientist: The Sexiest Job Of the 21st Century*, 90 (10) Harv. Bus Rev. 70, 70-78 (2012).
[91] Tiziana Terranova, *Free Labor: Producing Culture for the Digital Economy*, 18(2) Technol. Cult. 33, 33 (2020).
[92] Christian, *supra* not 56, at 255.
[93] *See* Tiziana, *supra* note 91, at 49.
[94] *See* Alexander Tarr & Alexander Tarr, *Will Review for Points: The Unpaid Affective Labour of Placemaking for Google's "Local Guides"*, 123(1) Feminsit Rev. 89, 89 (2019).

servers to track user activity in a way that facilitates e-commerce.[95] Today, almost all of our websites receive a similar inquiry from Cooike: "To improve your experience, we (and our partners) store and/or access information on your connected terminal on your connected terminal with your consent to all of our websites and apps".[96] This means that in transactions to obtain personal data, the user only receives "better service" by providing personal data, and the advertising revenue actually generated by this data is hidden behind the veil. This is similar to the statement that the price of labor as a commodity is the wages rather than the value of the commodity it produces, which justifies exploitation under the morality and laws of capitalist society.

Although some scholars point out that users receive immaterial rewards after selfless labor, digital rewards such as "badges" or higher "user level".[97] Besides others argue that people could get recognized during digital labor, since when audiences actively engage in the production of media content, they would have the opportunity for authentic self-expression.[98] But these rewards do not eliminate exploitation, just as Marx argued that raising workers' wages could not eliminate capitalist exploitation. Unlike classical political economists such as Ricardo, who believed that surplus value arises from the unfair exchange of labor and wages between workers and capitalists, Marx argues that surplus value arises from the inequality of ownership of the means of production. Although surveillance capitalists can give users many other rewards, they cannot own the final data product because they lack the means of production in the data economy. If and only if the value of the final data product is higher than the cost of providing the service, the surveillance capitalist will make a deal with the user. Just as Fisher argues "less alienation creates more exploitation",[99] surplus value and exploitation will persist despite the user receiving some reward.

Exploitation has been proven by many scholars to exist in Web 1.0 or Web 2.0, and the next task of this subsection is to discuss whether exploitation will continue or even intensify in the metaverse. The answer to this question is yes, due to the strengthening of the extraction of absolute surplus value and relative surplus value in the metaverse. In Marx's account, the history of capitalist production can be seen as the history of capitalists striving to raise the rate of surplus value and workers fighting each other vigorously against each other. "The capitalists may increase the amount of surplus value extracted from the working class by two means: (1) by absolute surplus value -- extending the working day as long as possible, and (2) by relative surplus value -- by cutting wages."[100] The characteristics of the metaverse will facilitate the extracting of these two means.

First, the immersion and persistence of metaverse experiences will lead to enhanced capture of absolute surplus value. The most direct way to intensify exploitation is to strengthen the extraction of absolute surplus value, that is, to increase the full value produced by each worker without changing the amount of necessary labor. In traditional capitalist wage labor, this is reflected in increased work intensity and longer working hours; While in surveillance

---

[95] *See* Sarah, *supra* note 29, at 27.
[96] This text comes from the database *Wiley online.*
[97] Valeria Pulignano et al., *Why does unpaid labour vary among digital labour platforms? Exploring socio-technical platform regimes of worker autonomy*, Hum. Relat. 19 (2023).
[98] *See* Eran Fisher, *How Less Alienation Creates More Exploitation? Audience Labour on Social Network Sites*, 10 (2) Triplec 171, 182 (2012).
[99] Eran, *supra* note 98, at 171.
[100] *Surplus Value*, Encyclopedia of Marxism, (Sep 10. 2023, 5:28PM), https://www.marxists.org/glossary/terms/s/u.htm.

capitalism, it is manifested in increasing the amount of time users spend online and increasing the density of data extraction. On one hand, in the context of the metaverse, a truly immersive virtual environment created by cutting-edge technologies and a stable and enduring virtual environment allowing individuals to inhabit and transform by actively participation could attract users to spend more time in it.[101] On the other hand, Immersive relies on cutting-edge technologies such as augmented reality, real-world graphics rendering, and digital twins are data-intensive technologies,[102] meaning that the metaverse will collect more personal data per unit of time.

Second, the extraction of absolute surplus value has limits and provokes resistance, so a more moderate way is to extract relative surplus value, that is, to increase the proportion of surplus labor by shortening the necessary labor time. In other words, reduces the value of labor. So a more moderate way is to extract relative surplus value, that is, to increase the proportion of surplus labor by shortening the necessary labor time. This entails lowering the price of labor, which is determined by the time to produce a fixed set of commodities to reproduce workers' labor capacity from one day to the next.[103] In traditional capitalism, this is often done through technological innovation to increase productivity. Current research has found that data-intensive companies are trying to replace real-world data with synthetic data, or data produced artificially at a time when surveillance is increasingly encountering social and regulatory resistance.[104] This will reduce the need for data from users while meeting the needs of machine learning. In other words, the technology enhances the capture of relative surplus value by reducing the amount of data necessary to sustain metaverse services.

Finally, the concept of "Factory Planet" also reveals the expanding risk of exploitation in the metaverse.[105] The exploitation of user labor in the metaverse is indicative of a phase of capitalism in which we find an "all-ubiquitous factory that is a space of the exploitation of labor".[106] Exploitation in industrial production is limited to working hours, while exploitation in domestic work is limited to the family. While social media and mobile internet have further expanded the scope of exploitation to the entire planet, the metaverse may further strengthen this trend, not only in scope but also in intensity as mentioned earlier. In this ubiquitous factory, users who lack data processing capabilities cannot meet needs such as social participation on their own and have to trade personal data for services in personal data transactions. Surveillance capitalists who control the means of production gain surplus value by selling personal data, which constitutes exploitation in the metaverse.

IV. DOMINATION IN THE METAVERSE

The previous section argues that differences in data processing capabilities have led to a relationship of economic exploitation between users and service providers in the metaverse. However, the difference between the two goes beyond data processing capabilities to control the data itself. The lack of control over personal data is believed to create a power gap that

---

[101] Shahper Richter & Shahper Richter, *What is novel about the Metaverse?*, 73 Int. J. Inf. Manage 102684, 6-7(2023).
[102] Ben, *supra* note 27, at 6-7.
[103] Francisco Paulo Cipolla, *The Mechanism of Relative Surplus Value*, 50(1) Rev. radic. polit. econ. 116, 121 (2018).
[104] James Steinhoff, *Toward a political economy of synthetic data: A data-intensive capitalism that is not a surveillance capitalism?*, New Media Soc, 1 (2022).
[105] Nick Dyer-Witheford, Digital Labour, Species Being and the Global Worker, 485 (2010).
[106] Christian, *supra* note 56, at 279.

creates a dominant relationship between users and service providers. The massive personal data grabbing of metaverse service providers and the existence of so-called "digital enclosure"[107] make domination the third moral challenge brought by the metaverse to our society. This section will analyze this issue in conjunction with the republican account of freedom and privacy theory.

A. Freedom of Non-Domination

Before discussing the dominance that exists in the metaverse, it is necessary to review the general theory of domination advocated by republicanism. Among the various schools of political philosophy, the most in-depth discussion of domination should be the republican scholars, who argue that the existence of dominance is the cause of the erosion of individual freedom. Republican theorists argue that the existence of domination is responsible for the deprivation of individual freedom, which is a critique of liberal non-interventionist accounts. Liberals assert freedom as an exclusion from external interference. Isaiah Berlin refers to this idea as negative freedom, which is "an area within which a man can act unobstructed by others".[108] This idea became the dominant theory in political philosophy after Bentham and William Paley[109] but has recently been challenged by republicanism. For republicans, this account of freedom fails to cover all situations in which it is appropriate to say that individual freedom has diminished. They argue that individual liberty could also be derogated without any actual interference, and summarized their account as "freedom as non-domination".

Republicans believe that domination is typically expressed in the relationship between "liber and servus, citizen and slave".[110] For its definition, one of the most widely known version comes from Pettit who argues that "someone dominates or subjugates another, to the extent that (1) they can interfere (2) with impunity and at will (3) in certain choices that the other is in a position to make".[111] In his later writings, he summarized dominance in a relatively short sentence "B has a power of interfering in the choice that is not itself controlled by A".[112] Thus, the unfreedom of the republican account depends on a persistent state of power, not on whether a particular decision is interfered with. Laborde and Maynor conclude that "domination is a function of the relationship of unequal power between persons, groups of persons, or agencies of the state".[113]

So the understanding of power relationship is the key to understanding domination. Pettit argues that it has three characteristics. Firstly this power is a capacity to interfere with the choices of others. The interference here is similar to the liberal account, including "removal, replacement, or misrepresentation of options".[114] However, Pettit gives it three restrictions: (1) such interference must be aimed at worsening rather than improving the situation of others; (2) it must occur through the intention, or at least negligence, of a party; (3) the ability of interference must be a practical ability, which we might call -- a capacity that is more or less

---

[107] Mark Andrejevic, *Surveillance in the Digital Enclosure*, 10(4) Communication Review 295, 297 (2007).
[108] Isaiah Berlin. et al. Liberty : Incorporating Four Essays on Liberty, 121 (2002).
[109] *See* Neil, Hopkins, *Freedom as Non-Domination, Standards and the Negotiated Curriculum: Freedom as Non-Domination*,49 (4) J. Philos. Educ. 607, 609(2015).
[110] Philip Pettit, Republicanism: A Theory Of Freedom And Government, 32 (1997).
[111] *Id.* at 53.
[112] Philip Pettit, On the People's Terms: A Republican Theory and Model of Democracy, 50 (2012).
[113] Cecile Laborde & John Maynor, Republicanism And Political Theory, 5 (2008).
[114] Andrew Roberts, *A republican account of the value of privacy*, 14 (3) Eur. J. Political Theory 320, 325(2015).

ready to be exercised -- not a capacity that yet to be fully developed.[115] Secondly, this ability must be arbitrary, which means that it fails to track the avowable interests of the subject, and the operation of this ability is unchecked, which means that the only constraint to it is the operator's own will.[116] Finally, domination needs not to be comprehensive, and a qualified dominator only needs to have the discretionary capacity to interfere with certain choices of others.

Republicans use the cases of "benevolent slave owners" and "rule by law" to support their account of freedom as a more appropriate alternative to liberial's. The former describes a situation in which freedom remains unfree without interference, while the latter describes a situation in which freedom is not diminished despite interference. The former assumes that a kind slave owner never interferes with the choice of his slaves, under the analysis of the liberal account, these slaves are undoubtedly free which contradicts intuition. Republicans could better handle this situation, who argues that one person has dominating power over another -- does not require that the person who enjoys such power actually interferes -- could lead to unfreedom.[117] The latter assumes that there are two states, one governed by many laws, and the other governed by an arbitrary monarchy. In a liberal account, citizens of the former state have less freedom since they are interfered with by laws. While republicans argue that as long as the law is not an instrument of the arbitrary will of any individual or any group, but respects the common interests and ideas of people,[118] the rule by law does not pose the risk of domination.

B. Domination in the Metaverse

Although service providers promote the metaverse as a free community, on the one hand, due to its decentralization, personal negative freedom will be less interfered with by public power, on the other hand, its virtual nature can allow users to open a second life with infinite possibilities, thereby enhancing personal positive freedom. However, behind the appearance of freedom in the metaverse, due to the acquisition and possession of users' personal data by service providers, there is an implicit risk of domination. Domination that exists in the metaverse presupposes that there is a specific power relationship between the user and the service provider, and this power relationship needs to give the latter the ability to interfere arbitrarily with the former. This arbitrary power is considered to come from the loss of privacy caused by the service provider's possession of the user's personal data.

Privacy issues are considered to be one of the most important challenges that the metaverse poses to our society. A survey of 300 developers on the Agora platform, one of the builders of the metaverse, revealed that 33% of respondents cited data privacy and security as the biggest hurdles the metaverse must overcome.[119] Fernandez & Hui argue that the technology of supporting metaverse introduces new ethical and privacy dilemmas as continuously sensing devices are made to expose users to privacy at sensory, behavioral, and

[115] *See* Philip, *supra* note 110, at 53.
[116] *Id*. at 59.
[117] *See* Philip, *supra* note 110, at 64.
[118] *Id.* at 36-41.
[119] *See* Team Agora, *SURVEY: Developers Cite Data Privacy and Security and Disinformation and Hate Speech as Top Metaverse Challenges*, Agora Newsroom, (Sep 11. 2023, 4:59 PM), https://www.agora.io/en/news/survey-developers-cite-data-privacy-and-security-and-disinformation-and-hate-speech-as-top-metaverse-challenges/.

communication levels by collecting large amounts of metadata.[120] Other scholars believe that in the metaverse, malicious parties can collect sensitive information from the online profiles and public information of social network users, in addition, data leaks and unintentional data releases, or incorrect security or privacy protection configurations may also cause privacy problems.[121]

Research by privacy scholars shows that the loss of privacy leads to changes in power relations. The traditional liberal view is that the value of privacy lies in the guarantee of autonomy since it leaves us the physical and psychological space that we need to reflect on ourselves and our lives.[122] A new view of privacy is that the value of privacy lies in providing agents an antipower.[123] The concept of antipower is raised by Pettit, who argues that antipower is what comes into being as the power of some over others is actively reduced and eliminated.[124] Pettit believes that resources are the main factor influencing antipower and that bullies always have more resources at their disposal than those who are bullied.[125] Privacy guarantees individual's control over his or her own data and information, which as a resource in the information society will increase the antipower of its owner.

In the metaverse, when users' privacy is violated by service providers and lose control of their personal data, they will lack the antipower to resist domination. Accordingly, due to the lack of antipower restrictions, service providers have arbitrary powers. Zuboff argues that "user dependency is the fulcrum of the surveillance capitalist project..... Most people find it difficult to withdraw from these utilities, and many ponder if it is even possible".[126] In the metaverse, a large number of services need to be provided by a large amount of data aggregation, and users who do not master data often have to rely on service providers. As more and more functions will be accessed in the metaverse, users' avatars or accounts will become increasingly important to individuals. When the data that maintains the avatar and account is in the hands of the service provider, the survey provider can easily make the user submit by denying access, deleting the data, banning the account, etc., and the user lacks the counter-power to fight it.

Arbitrary power also means that it will be exercised with the ultimate goal of maximizing the interests of the service provider, which is equally evident in surveillance capitalism. Zuboff argues that there is "radical indifference" in surveillance capitalism, which means that Surveillance capitalists lack organic reciprocity with those who are sources of consumers or employees.[127] Since traditional capitalist production requires a large number of employees and consumers composed of residents, capitalists need to promote the welfare of the population to obtain stable employment and expand consumption. The number of employees needed by surveillance capitalism is greatly reduced, and its buyer becomes

---

[120] *See* Carlos Bermejo Fernandez & Pan Hui, Life, *the Metaverse and Everything: An Overview of Privacy, Ethics, and Governance in Metaverse*. arXiv:2204.01480, (2022), https://arxiv.org/abs/2204.01480.
[121] *See* Yan Huang et al., *Security and Privacy in Metaverse: A Comprehensive Survey*, 6(2) Big Data Min. Anal. 234, (2023).
[122] *See* Andrew Roberts, *Forewords Why Privacy and Domination?*, 4 (1) Eur. Data Prot. Law Rev. 5, 6 (2018).
[123] *See* Bryce Clayton Newell, Forewords · Privacy as Antipower: In Pursuit of Non-Domination, 4 (1) Eur. Data Prot. Law Rev. 12, 14 (2018).
[124] *See* Philip Pettit, *Freedom as Antipower*, 106(3) Ethics 576, 588 (1996).
[125] *Id.* at 589.
[126] Shoshana Zuboff, *Surveillance Capitalism and the Challenge of Collective Action*, 28 (1) New Labor Forum 10, 24-25(2018)
[127] *Id.* at 21.

advertisers. As a result, the welfare of metaverse users is no longer important to service providers, which will make them more arbitrary in using their power.

As we have seen, the arbitrary power that constitutes domination needs to be able to interfere with the choices of others. In the metaverse, the power that service providers acquire through user data is also considered to be able to interfere with user choices in three forms. First of all, the survey provider can make the user voluntarily give up this option by imposing a large enough sanction on the option he does not want to see. Users are considered to have a chilling effect when they choose under surveillance, which makes them voluntarily abandon some options that they believe will bring negative results.[128] This deterrence may be further heightened in the metaverse because the user's avatar is tied to more interests than current Internet accounts. It is conceivable that a digital avatar tied to all of your bank accounts, social accounts, and gaming accounts would bring far more damage to the user if blocked than if a single functional account were blocked. So some scholars argue that as we move into an increasingly automated, networked world, our freedoms will be limited as governance improves.[129]

Secondly, when the user's choice is known to the survey provider, the latter can preemptively delete some options that are not favorable to it, so that although the user has the formal freedom of choice, his or her choice has been interfered in essence. And this form of interference is more common in the Metaverse. Since the Metaverse is a virtual digital world, many physical obstacles are no longer obstacles in it. This on the one hand facilitates the extension of the user's positive freedom, but on the other hand, makes it easier for the service provider to practice interference. For example, if a service provider does not want a user to access a certain area, then they can close it off by simply entering a few commands, but in the real world, it may be costly to build a fence or hire guards. Thus, removing the user's option also becomes easier in the metaverse.

Finally, in terms of misrepresentation, Pettit argues that successful manipulation "will affect the exercise of your cognitive capacity to choose between certain options even if it leaves your objective capacity in place. By means of manipulation, I may succeed in getting you to choose as I wish."[130] Thus, if survey providers know users' strategies, fears, and weaknesses, they can devise counter-strategies, manipulate and nudge users in the direction of choices they think are more desirable, or coerce users into taking options that survey providers prefer.[131] Personalization of content by service providers to users is considered to be a typical example of manipulation, where the service provider changes the content of the item pushed by obtaining the user's search history thus inducing the consumer to spend money.[132] In the metaverse, user data is more fully grasped by service providers, which will lead to further strengthening of their ability to misrepresent.

---

[128] *See* Shoshana, *supra* note 26, at 82.
[129] *See* Jack M. Balkin,*Room for Maneuver: Julie Cohen's Theory of Freedom in the Information State*, Jrslm. Rev. Legal Stud. 79, 82 (2012) ; Julie E. Cohen, What privacy is for. 126(7) Harv. Law Rev. 1904, 1918-1926 (2013).
[130] Philip, *supra* note 112, at 55.
[131] *See* Andrew Roberts, *A republican account of the value of privacy*, 14 (3) Eur. J. Political Theory 320, 336 (2015).
[132] *See* Neil M. Richards, *The dangers of surveillance*, 126(7) Harv. Law Rev. 1934, 1939 (2013).

C. Common Knowledge and Domination

Van Dijck argues that "a large number of people -- naively or unwittingly -- trust their personal information to corporate platforms".[133] These people seem unaware or unconcerned about possible domination and are happy to trade personal data for metaverse services. Pettit believes that when the three conditions of domination are satisfied, the society will form common knowledge, that is, the dominated person realizes that he is in a dominated relationship.[134] And because of the existence of this knowledge, the dominated will take the initiative to cater to the dominators' preferences to avoid actual interference. Therefore, the dominated will voluntarily give up some options that they would have chosen, so even no actual interference constitutes a limitation on autonomy. Thus common knowledge seems to be a necessary condition for the restriction of freedom in domination without actual interference.

This view seems to cut the link between domination and freedom, but it is a misunderstanding of republican non-domination freedom. Non-dominant freedom is not concerned with the autonomy of the agent in the face of specific choices but with the power relationship between the agent and others. In other words, freedom of non-dominance is concerned about the state of being a free person, while freedom of non-intervention is concerned about freedom in concrete choices. Therefore, the identification of freedom in the republican account does not depend on whether there is a loss of autonomy in specific choices, but on whether there is domination in the power relations in which the agent is placed.[135] The simple fact that we depend on others—others have acquired the power to interfere arbitrarily—undermines our freedom. Roberts and Pettit also acknowledge that an exception to common consciousness is the manipulation of individual and group options of the dominated, where common knowledge may not exist but the dominated is still influenced by power.[136] Manipulation in the metaverse is also considered to be the riskiest means of interference. Users are often not aware of the significant value of their data, but their choices are still substantially at risk of interference by the service provider.

CONCLUSION

In conclusion, this dissertation argues that the metaverse, as a virtual digital world is both reductive and transcendent to the real physical world and can interact with the real world. It allows our activities to transcend the physical barriers of the real world to some extent but still poses many moral risks to our society. Because of its digital properties, the risk of surveillance of users in the metaverse is greatly increased, and this gives rise to a new mode of accumulation, which Zuboff calls surveillance capitalism, and which is thought to pose three ethical challenges.

The technologies that the metaverse relies on (e.g. XR and IoT) capture large amounts of data, and cloud storage technologies prevent users from controlling over their data. This process puts metaverse users at risk of alienating themselves from their data, behavior, class nature, and others. And in the metaverse due to the alienation of users and service providers into two opposing classes, this class division and opposition is seen to create two other moral

---

[133] José, *supra* note 38, at 197.
[134] *See* Philip, *supra* note 110, at 59-60.
[135] *See* Christian Dahl & Tue Andersen Nexø, To Be Unfree 37–54 (2014).
[136] *See* Andrew Roberts, *Privacy, Data Retention and Domination: Digital Rights Ireland Ltd v Minister for Communications*, 78 (3) Mod. Law Rev. 535, 544 (2015).

challenges. Economically, the greater data processing capacity of service providers enables them to exploit users' "digital labor" in data-service transactions with them. Politically the service provider dominates the user because it holds a lot of personal data about the user and acquires the arbitrary right to interfere with the user's decisions at will.

Since the metaverse itself is a nascent and rapidly evolving matter, any exploration of it at this stage is based on limited observations and appropriate conjectures. Although this dissertation argues that it will pose great challenges to our society, many developing technologies may also be called effective measures to avoid these risks, and research on this issue will also be my future focus.

This page intentionally left blank.



www.ijlet.org

La Nouvelle Jeunesse