



THE IEEE GLOBAL INITIATIVE ON ETHICS OF EXTENDED REALITY (XR) REPORT

XR ETHICS IN WAR, CONFLICT, AND DEFENSE

Authored by
Michael Middleton

Contributing Author
Mathana

TRADEMARKS AND DISCLAIMERS

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

The ideas and proposals in this specification are the respective author's views and do not represent the views of the affiliated organization.

ACKNOWLEDGMENTS

Special thanks are given to the following reviewers of this paper:

Eleni Mangina, *editorial coordinator*

Monique J. Morrow, *contributing reviewer*

The Institute of Electrical and Electronics Engineers, Inc. 3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2024 by The Institute of Electrical and Electronics Engineers, Inc.

All rights reserved. 25 October 2024. Printed in the United States of America.

PDF: STDVA27411 979-8-8557-1356-5

IEEE is a registered trademark in the U. S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated. All other trademarks are the property of the respective trademark owners.

IEEE prohibits discrimination, harassment, and bullying. For more information, visit <http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>.

No part of this publication may be reproduced in any form, in an electronic retrieval system, or otherwise, without the prior written permission of the publisher.

Find IEEE standards and standards-related product listings at: <http://standards.ieee.org>.

NOTICE AND DISCLAIMER OF LIABILITY CONCERNING THE USE OF IEEE SA INDUSTRY CONNECTIONS DOCUMENTS

This IEEE Standards Association (“IEEE SA”) Industry Connections publication (“Work”) is not a consensus standard document. Specifically, this document is NOT AN IEEE STANDARD. Information contained in this Work has been created by, or obtained from, sources believed to be reliable, and reviewed by members of the IEEE SA Industry Connections activity that produced this Work. IEEE and the IEEE SA Industry Connections activity members expressly disclaim all warranties (express, implied, and statutory) related to this Work, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; quality, accuracy, effectiveness, currency, or completeness of the Work or content within the Work. In addition, IEEE and the IEEE SA Industry Connections activity members disclaim any and all conditions relating to: results; and workmanlike effort. This IEEE SA Industry Connections document is supplied “AS IS” and “WITH ALL FAULTS.”

Although the IEEE SA Industry Connections activity members who have created this Work believe that the information and guidance given in this Work serve as an enhancement to users, all persons must rely upon their own skill and judgment when making use of it. IN NO EVENT SHALL IEEE OR IEEE SA INDUSTRY CONNECTIONS ACTIVITY MEMBERS BE LIABLE FOR ANY ERRORS OR OMISSIONS OR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS WORK, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Further, information contained in this Work may be protected by intellectual property rights held by third parties or organizations, and the use of this information may require the user to negotiate with any such rights holders in order to legally acquire the rights to do so, and such rights holders may refuse to grant such rights. Attention is also called to the possibility that implementation of any or all of this Work may require use of subject matter covered by patent rights. By publication of this Work, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. The IEEE is not responsible for identifying patent rights for which a license may be required, or for conducting inquiries into the legal validity or scope of patents claims. Users are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. No commitment to grant licenses under patent rights on a reasonable or non-discriminatory basis has been sought or received from any rights holder. The policies and procedures under which this document was created can be viewed at <https://standards.ieee.org/about/bog/iccom/>.

This Work is published with the understanding that IEEE and the IEEE SA Industry Connections activity members are supplying information through this Work, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. IEEE is not responsible for the statements and opinions advanced in this Work.

TABLE OF CONTENTS

XR ETHICS IN WAR, CONFLICT, AND DEFENSE.....	5
ABSTRACT	5
1. INTRODUCTION.....	6
2. ARTIFICIAL INTELLIGENCE (AI) & ITS IMPACT	8
3. RECOMMENDATION 1	10
3.1. SUMMARY	10
3.2. CONTEXT.....	10
4. RECOMMENDATION 2	12
4.1. SUMMARY	12
4.2. CONTEXT.....	12
5. RECOMMENDATION 3	14
5.1. SUMMARY	14
5.2. CONTEXT.....	14
6. RECOMMENDATION 4	16
6.1. SUMMARY	16
6.2. CONTEXT.....	16
7. RECOMMENDATION 5	18
7.1. SUMMARY	18
7.2. CONTEXT.....	18
8. RECOMMENDATION 6	20
8.1. SUMMARY	20
8.2. CONTEXT.....	20
9. RECOMMENDATION 7	22
9.1. SUMMARY	22
9.2. CONTEXT.....	22
10. RECOMMENDATION 8	24
10.1. SUMMARY	24
10.2. CONTEXT	24
10.2.1. GENEVA CONVENTIONS	24
11. REFERENCES.....	26

XR ETHICS IN WAR, CONFLICT, AND DEFENSE

ABSTRACT

This report is the result of work within the IEEE Global Initiative on Ethics of Extended Reality (XR), a multidiscipline group of industry practitioners, ethicists, academics, researchers, educators, and technology enthusiasts. It has been written to focus on a wide range of ethical issues related to XR within the war, conflict, and defense domain. This report builds on work outlined in the “Extended Reality” chapter of the IEEE’s seminal ethics-focused publication *Ethically Aligned Design*. XR is a term used to broadly refer to a suite of immersive technologies, including virtual reality, augmented reality, and spatial computing. The scope of this report is the exploration of ethics-related issues to support the development, design, and deployment of XR applications in the war, conflict, and defense domain and the aim is to initiate expert-driven, multidiscipline analysis of the evolving XR Ethics requirements, with a vision to propose solutions, technologies, and standards in future updates. The set of recommendations within this report will hopefully contribute to the industry conceptualization of socio-technological issues, highlight concreted recommendations, and lay the groundwork for future technical-standardization activities.

MATHANA & ELENi MANGINA

CHAIR & VICE CHAIR

IEEE GLOBAL INITIATIVE FOR ETHICAL EXTENDED REALTY

1. INTRODUCTION

In 2007, Dr. Frank G. Hoffman released the seminal paper called *Conflict in the 21st Century: The Rise of Hybrid Wars*. Important in its own right, it was also a follow-up to an earlier paper written with General Jim Mattis, to whom Dr. Hoffman attributes the first public use (on September 8, 2005) of the term “hybrid warfare” [1].¹

To say that the technology landscape was different from now in 2005 is an understatement. YouTube had only been founded a few months prior, in February [2]—ironically, the same month in which *The New York Times* had published an article titled *New Model Army Soldier Rolls Closer to Battle*, in which it noted that “the Pentagon predicts that robots will be a major fighting force in the American military in less than a decade, hunting and killing enemies in combat.”[3] That article also quoted Dr. Robert Finkelstein, the President of Robotic Technology, Inc., as “telling the Pentagon it could take until 2035 to develop a robot that looks, thinks and fights like a soldier. The Pentagon’s “goal is there,” he said, “but the path is not totally clear.”

But it could be argued that the Pentagon’s path was, in fact, quite clear; after all, it was already four years since Air Force pilot Scott Swanson had become the first person to successfully kill a target via a Hellfire missile fired from a Predator drone [4]. The public use of the term “hybrid warfare” was merely evidence of a shift that had already begun, and for good reason. The events of September 11, 2001, had resulted in extraordinary changes in security and defense, with neither the public nor the Pentagon wanting to pay the price for another infamous “failure of imagination” [5]. For example, just a few weeks prior to General Mattis’ speech, China and Russia had held the first joint military exercise in their shared history, dubbed “Peace Mission 2005.” And the internet was promising to bring new forms of disruption and danger to the concept of defense. (Within a year, WikiLeaks would be founded [6], and the Chelsea Manning leak would shock the system for an additional four years [7].)

Of course, it is not necessary to reach so far back for evidence of technology disrupting security. The events of the past two years have been so bold and fast-moving that this paper has had to be updated monthly before finally giving in and submitting it for publication, knowing it might quickly become outdated. Simply put, the release of ChatGPT on November 30, 2022, has brought around a great awakening in everyday people about just how quickly and fundamentally artificial intelligence (AI) might change their lives.

Currently, the world is in a period of extraordinary competition between companies, countries, and random interests for who, how, when, and why this power will be developed, wielded, and shared. The focus of this paper is on extended reality (XR), but AI will be a fundamental part of the discussion, especially given that it has an unprecedented thirst for data that XR is uniquely suited to quench—perhaps in exchange for giving XR the path

¹ Numbers in brackets correspond to the references listed in Section 11.

to ubiquity and acceptance it has long been missing.

Not to be misinterpreted—technology has taken humanity to extraordinary places and allowed for widespread wealth and peace that was literally once unimaginable. But it also can, and has, been used to kill, neutralize, and restrain human beings on a level that grows increasingly concerning. As one draft of this paper was being finalized, *+972 Magazine* published an article describing a previously-secret Israeli AI program called “Lavender” and claimed that the nation’s army “almost completely relied” on the AI program during the first weeks of the Israel-Hamas War, giving “sweeping approval for officers to adopt Lavender’s kill lists, with no requirement to thoroughly check why the machine made those choices or to examine the raw intelligence data on which they were based.” *The Guardian* newspaper published a response in turn from the Israel Defense Forces (IDF), which directly stated that “the IDF does not use an artificial intelligence system that identifies terrorist operatives or tries to predict whether a person is a terrorist,” and that “analysts must conduct independent examinations, in which they verify that the identified targets meet the relevant definitions in accordance with international law and additional restrictions stipulated in the IDF directives.” [8]

It is becoming increasingly important that all people—as human beings, and not as political animals or agents of corporations—need to understand that once these phenomena are normalized, they will forever change relationships with one another, as well as technology itself. In his 2014 book, *Superintelligence: Paths, Dangers, Strategies*, Nick Bostrom says, “This is quite possibly the most important and most daunting challenge humanity has ever faced. And—whether we succeed or fail—it is probably the last challenge we will ever face.” He is speaking of superintelligence, but XR professionals should be keenly aware of how closely aligned to this world they are and will be.

The speed with which new XR applications and ethical problems arise is astonishing, but it seems that many folks in XR are all too keenly aware of the challenges being faced; guidance is already available from government and other organizations. For example, there are key crossovers to the US Department of Defense’s Ethical Principles for Artificial Intelligence [9], as well as to the express or implied points of view of governments around the world.

General Matts’ speech came almost sixty years after Morton Heilig unveiled his *Sensorama* machine (by many accounts, one of the first virtual reality (VR) systems), and fifty years after Ivan E. Sutherland’s famous *The Ultimate Display*, which ends with the following quote: “The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked [10].”

2. ARTIFICIAL INTELLIGENCE (AI) & ITS IMPACT

This paper is first and foremost about XR; however, it must be acknowledged that a special relationship has already formed between XR and artificial intelligence (AI)—a relationship that is so important that the final draft of this paper was delayed (among other reasons) in order to see the final framework of the European Union’s AI Act, which passed unanimously on March 13, 2024 [11].

Of special interest is Title II, Art. 5, which deals with prohibited AI systems and includes numerous references to XR concerns expressed in this paper, including a system that “deploys subliminal techniques beyond a person’s consciousness to materially distort a person’s behavior in a manner that causes or is likely to cause that person or another person physical or psychological harm.” [12]. It was not feasible to fully review all AI impacts and concerns, but two simple concepts warranted inclusion, as discussed in the following paragraphs.

As mentioned previously, AI has an unprecedented thirst for data that XR is uniquely suited to quenching. Many existing historical data sources have been heavily mined, and—for the moment—synthetic data generated by AI is not providing the same measure of quality. Even when it improves, it will still fail to provide a full and faithful account of ongoing human life; where better to turn than to XR and its ability to capture enormous, deeply intimate volumes of data about people and systems on a level never before possible?

For XR, this will likely be a transformative boon, providing the level of prominence, growth, and ubiquity its supporters have long felt it deserves. And the more data it feeds AI, the more articulately AI will be able to help it gather, analyze, and leverage new data in ways that are not currently understood. What also remains unknown for AI is at what speed and at what level of power each sea-change will come—and what the resulting effects will be for all human beings.

In this, there is no shortage of disparity in optimism or pessimism. Marc Andreessen’s *The Techno-Optimist Manifesto* bluntly states, “We believe that we are, have been, and will always be the masters of technology, not mastered by technology,” [13] while the *Global AI Moratorium* site calls for a “global moratorium on large AI training runs until the technical problem of AI alignment is solved,” before noting further:

“Today, cutting-edge AI systems are artificial neural networks: millions to trillions of numbers we automatically adjust until they start achieving a high score on some metric. We don’t know what these numbers represent. We do not know how the resulting AI systems work or what their goals are.” [14]

There is much to be optimistic about, but that is no excuse to avoid risk mitigation, nor is it a reason to allow for

the second point: the danger of normalization. Every day, decisions are being made and normalized in theory and on the battlefield in ways that will be extremely difficult to unwind.

Consider Lori Robinson and Michael O'Hanlon's 2020 commentary "The evolution of artificial intelligence and future of national security" in which they shared the following (condensed for relevance):

"The vice chairman of the Joint Chiefs of Staff, General Paul Selva, has recently argued that the United States could be about a decade away from having the capacity to build an autonomous robot that could decide when to shoot and whom to kill—though he also asserted that the United States had no plans actually to build such a creature.

Even if General Selva's terminator is not built, robotics will in some cases likely be given the decision-making authority to decide when to use force. This highly fraught subject requires serious careful ethical and legal oversight, to be sure, and the associated risks are serious. Yet the speed at which military operations must occur will create incentives not to have a person in the decision-making loop in many cases." [15]

Nine countries are known to possess nuclear weapons; a few additional countries host them on behalf of those nine [15]. XR professionals and consumers should be deeply aware of the costs and associated threats to global security that have come from these weapons in the past eight decades, and cognizant of their own current power, ability, and responsibility to have a say in what an XR-equipped future looks like in the critical fields of war, conflict, and defense—especially given that this technology (while not an apples-to-apples match against nuclear weapons) will certainly not be restrained to a few countries or interest groups.

3. RECOMMENDATION 1

3.1. SUMMARY

XR use in war and conflict must respect individual personhood.

It should never:

- Alter the apparent number of living persons physically present in a place.
- Change the appearance of an actual person to belie their nature, such as to make them appear in a uniform that does not represent their allegiance.
- Change their state, such as to make the injured appear uninjured or vice versa.
- Alter their assets, such as to add a weapon that is not there or remove one that is.
- Alter the state of their assets or situation, such as to add or remove heat signatures that could indicate the active/inactive state of their weapons, or to add or remove the appearance or signature of chemicals or other malicious environmental hazards that would change the impression of their status in the eyes of a VR operator.

3.2. CONTEXT

In 2019, three researchers from the University of New Haven published one of the first and most important papers on immersive virtual reality attacks [17]. They classified such attacks into four main areas, one of which was as follows:

“Overlay Attack: Any attack that *overlays* unwanted images/video/content on a player’s VR view. The player will have no option to remove the content. This attack includes persistent images as well as content that remains fixed in virtual space.”

Author’s note: Not to be confused with a similar concept with the same name in mobile devices where users are unknowingly interacting with an active and malicious window, over a trusted program.

While they were speaking to the concept of a game player who is perhaps being bombarded with advertisements, malicious, or harmful content, a line can be drawn to serious ethical perils in the zone of national defense or even the administration of justice/law enforcement within a community.