

# **The AIWS *7-Layer Model* to Build Next Generation Democracy**

**Michael Dukakis**

**Nazli Choucri**

**Allan Cytryn**

**Alex Jones**

**Tuan Anh Nguyen**

**Thomas Patterson**

**Derek Reveron**

**David Silbersweig**

April 25, 2018

## ***Abstract***

The Artificial Intelligence World Society (AIWS) is a set of values, ideas, concepts and protocols for standards and norms whose goal is to advance the peaceful development of AI to improve the quality of life for all humanity. It was conceived by the Michael Dukakis Institute for Leadership and Innovation (MDI) and established on November 22, 2017. The World Leadership Alliance – Club de Madrid (WLA-CdM) and the Boston Global Forum (BGF) are partnered with the MDI to collaborate and develop the AIWS initiative. The President of WLA-CdM, Vaira Vike-Freiberga, serves as co-chair of AIWS activities and conferences along with Governor Michael Dukakis.

The Next Generation Democracy (NGD) is an initiative founded by WLA-CdM with the goal of “enabling democracy to meet the expectations and needs of all citizens and preserve their freedom and dignity while securing a sustainable future.” NGD is a collaboration and forum, coordinated by WLA-CdM. AIWS has partnered with WLA-CdM to promote the development of AI to support the Next Generation Democracy initiative.

To align the development of AI with the NGD initiative, the AIWS has developed the AIWS 7-Layer Model. This model establishes a set of responsible norms and best practices for the

development, management, and uses of AI so that this technology is safe, humanistic and beneficial to society.

In developing the 7-Layer Model, the AIWS recognizes that we live in a chaotic world with differing, and sometimes conflicting, goals, values and concepts of norms. Hence, the Model is aspirational and even idealistic. Nonetheless, it provides a baseline for guiding AI development to ensure positive outcomes and to reduce the risks of pervasive and realistic risks and the related harms that AI could pose to humanity.

The Model is based on the assumption that humans ultimately accountable for the development and use of AI, and must therefore preserve that accountability. Hence, it stresses transparency of AI reasoning, applications, and decision making, which will lead to auditability and validation of the uses of AI systems.

**Citation:** Dukakis, M., Choucri, N., Cytryn, A., Jones, A., Nguyen, T. A., Patterson, T., Reveron, D., & Silbersweig, D. (2018). *The AIWS 7-layer model to build next generation democracy* (BGF-G7 Summit 2018 Report). The Boston Global Forum, & Michael Dukakis Institute for Leadership and Innovation.

**Unique Resource Identifier:** <https://bostonglobalforum.org/wp-content/uploads/The-BGF-G7-Summit-Report.pdf>

**Publisher/Copyright Owner:** © Boston Global Forum, Michael Dukakis Institute for Leadership and Innovation.

**Version:** Final published version.

## The BGF-G7 Summit Report

### The AIWS 7-Layer Model to Build Next Generation Democracy



6/2018



# The Boston Global Forum - G7 Summit 2018 Report

***Michael Dukakis***

***Nazli Choucri***

***Allan Cytryn***

***Alex Jones***

***Tuan Anh Nguyen***

***Thomas Patterson***

***Derek Reveron***

***David Silbersweig***

## ***The AIWS 7-Layer Model to Build Next Generation Democracy***

**The Artificial Intelligence World Society (AIWS)** is a set of values, ideas, concepts and protocols for standards and norms whose goal is to advance the peaceful development of AI to improve the quality of life for all humanity. It was conceived by the Michael Dukakis Institute for Leadership and Innovation (MDI) and established on November 22, 2017. The World Leadership Alliance – Club de Madrid (WLA-CdM) and the Boston Global Forum (BGF) are partnered with the MDI to collaborate and develop the AIWS initiative. The President of WLA-CdM, Vaira Vike-Freiberga, serves as co-chair of AIWS activities and conferences along with Governor Michael Dukakis.

**The Next Generation Democracy (NGD)** is an initiative founded by WLA-CdM with the goal of “enabling democracy to meet the expectations and needs of all citizens and preserve their freedom and dignity while securing a sustainable future.” NGD is a collaboration and forum, coordinated by WLA-CdM. AIWS has partnered with WLA-CdM to promote the development of AI to support the Next Generation Democracy initiative.

To align the development of AI with the NGD initiative, the AIWS has developed the *AIWS 7-Layer Model*. This model establishes a set of responsible norms and best practices for the development, management, and uses of AI so that this technology is safe, humanistic and beneficial to society.

In developing the *7-Layer Model*, the AIWS recognizes that we live in a chaotic world with differing, and sometimes conflicting, goals, values and concepts of norms. Hence, the Model is aspirational and even idealistic. Nonetheless, it provides a baseline for guiding AI development to ensure positive outcomes and to reduce the risks of pervasive and realistic risks and the related harms that AI could pose to humanity.

The Model is based on the assumption that humans ultimately accountable for the development and use of AI, and must therefore preserve that accountability. Hence, it stresses transparency of AI reasoning, applications, and decision making, which will lead to auditability and validation of the uses of AI systems.



## **Layer 1: Charter and Principles: *To create a society of AI for a better world and to ensure peace, security, and prosperity***

AI “society” is the society consisting of all objects that have the characteristics of Artificial Intelligence. Any object in this society is an AI Citizen. There must be rules that govern the behaviors of these AI Citizens, as there are rules that govern human members of society. The standards and requirements for an AI citizen must also include the need to manage and supervise them. AI citizens are to be transparent in structure and process, and all are to meet AIWS Standards of AI citizenship.

- AI Citizens cannot threaten or put at risk the health, safety, dignity and freedom of any human
- AI Citizens cannot take actions which violate the law and social norms of the societies in which they are deployed
- The design and performance of an AI Citizen must be sufficiently transparent so as to expose its behavior and ensure that its behavior will not imperil in any way other AI Citizens or humans, nor violate the law and social norms of the societies in they are deployed
- The performance of AI Citizens must meet basic standards of auditability and subject to regular audits to facilitate compliance with the above.

-----

Layer 1 establishes a responsible code of conduct for AI Citizens to ensure that AI is safely integrated into human society.

## **Layer 2: Ethical Frameworks: *Guidelines for the Role of AI in Building the Next Generation Democracy***

The behavior of AI Citizens must be ethical by normal human and social standards. It must conform to the ethics codes of UNESCO and of the United Nations as whole. To be considered ethical, such behaviors must be:

- Honest, open and transparent
- People-centric: for the people, by the people, serving the people
- Respectful of the dignity of humans, their privacy, and the natural environment
- Deployed in the service of individuals, groups, and governments that are themselves ethical.
- Promote and foster tolerance.

-----

Layer 2 is based on the ethics of codes of the UN and UNESCO. Therefore, AI citizens must, first and foremost, respect human dignity, virtue and ethics. The ethics layer will also draw on best practices and ethics codes of top businesses and organizations involved in AI research such as IBM, IEEE, the Berkman Center, and MIT Media Lab.



**Layer 3: Standards: *Standards for the Management of AI Resources and Development***

Establish the AIWS Standards and convene the Practice Committee to develop, manage, and promote standards and all critical requisites of an AI citizen.

The Committee will engage with governments, corporations, universities, and other relevant organizations to facilitate understanding of AI threats, challenges etc. These entities are ultimately responsible for achieving ethical AI.

-----

Layer 3 is focuses primarily on AI development and resources, including data governance, accountability, development standards, and the responsibility for all practitioners involved directly or indirectly in creating AI.

**Layer 4: Laws and Legislation: *Laws for the Role of AI in Building Next Generation Democracy.***

Advise political leaders in crafting the best possible rules, regulations, and legislation regarding AI technologies.

This layer will follow and apply Layers 1, 2, and 3 to transform these into legal and legislative concepts.

-----

Layer 4 focuses on policies, laws and legislation, nationally and internationally, that govern the creation and use of AI and which are necessary to ensure that AI is never used for malicious purposes.

There is great danger in the risk of AI development, devoid of the appropriate ethical and regulatory frameworks. Public and private entities are already considering ways to regulate AI. Regardless of what they may accomplish, there is further risk that the rate at which AI advances will outpace the continuing development of these frameworks. The goal of this layer is to guide leaders in these endeavors so that their work is effective and timely.

**Layer 5: International Policies, Conventions, and Norms: *Global Consensus***

To be effective, the development of AI in the support of humankind depends on a global consensus. International conventions, regulations, and agreements for AI development in support of Next Generation Democracy are therefore essential for the success of AIWS.

This layer will promote the adoption of the AIWS ethics, standards and legislative proposals consistent with, and integrated in, international law through conventions, regulations, treaties, and agreements.

-----

Layer 5 will focus on the global application and diffusion of AIWS-established norms and concepts. The responsible development and use of AI depends on acceptance by the global community. If even one state or actor uses AI irresponsibly or maliciously, the threat they could pose would be significant, and cannot be accepted, consensus .

AIWS also calls upon the leaders of all G7 nations to sign an agreement on the Ethical Development and Deployment of AI. Such an agreement would prohibit the development of autonomous AI weapons and mandate that AI be developed only for peaceful purposes. The threat posed by a potential AI arms race is alarming and states must take action now to prevent such a possibility.



## Layer 6: Public Services and Policymaking: *Engage and Assist Political Leaders*

AI can itself be used to aid in achieving the legislative and policy goals that promote its peaceful and constructive use. It can assist political leaders in effective and practical decision-makings by providing AI-based evaluations, data, and suggestions to solve social and political issues. This will ensure that all parties are informed and make the best decisions possible.

This layer will help to shape applications for Next Generation Democracy.

-----

Layer 6 emphasizes the role AI should play in providing analysis and data to inform political leaders. While AI per se cannot perform the functions of leadership, they will prove invaluable assets by providing assistance for human leaders. Examples of current AI projects for policymaking include SAM, the world's first AI politician, created and operating in New Zealand, and GROW360 in Japan.

## Layer 7: Business Applications for All of Society: *Engage and Assist Businesses*

As AI is deployed to be used by business, industry, and private citizens, it is essential to that AI technologies remain benevolent and free from risks of misuse, error, or loss of control. Therefore, it is imperative to work with the private sector in developing best practices for the applications of AI in society.

Layer 7 emphasizes the applications of AI and the services it can (and does) provide to citizens. AI is already being sold to, or tested for, consumer use in a variety of sectors. This includes fully autonomous vehicles, smart home assistants (e.g., Alexa and Google Home), and others. It also includes more subtle uses in social media, aviation, and other large sectors. With AI becoming more integrated in the lives of the average citizen, the technology will increasingly change our society.

Through Layer 7, and the Model as a whole, AIWS hopes to ensure that inviting AI into our lives will have positive effects.

## Plan 2020 and the Future

By February 20, 2020, all seven layers of the Model will be completed. Future applications and users of the final Model: AIWS International Court, AIWS University, AIWS Healthcare, AIWS Public Transportation, AIWS Policy Makers, and AIWS Political Leaders. These applications will ensure that AIWS norms for AI are adopted broadly and responsibly.

## Next Steps: Actions to be Taken

The AIWS Standards and Practice Committee is established to:

- Update and collect information on threats and potential harm posed by AI
- Connect companies, universities, and governments to find ways to prevent threats and potential harm
- Engage in the audit of behaviors and decisions in the creation of AI
- Create both an Index and Report about AI threats - and identify the source of threat
- Create a Report on respect for, and application of, ethics codes and standards of governments, companies, universities, individuals, and all others
- Work with the UN to call for an AI Peace Treaty, to AI, similar to the Chemical Weapons Convention that prohibits the creation, stockpiling, and use of those weapons
- Work with AI experts on a consensus announcement that "AI experts not engage in any work for, or participate in projects developing AI weapons".