Artificial Intelligence World Models – for mission-critical and commercial applications –*

Prof. Dr. V. David Sánchez A., Ph.D.
Brilliant Brains, Palo Alto, California
November 2025

Abstract

An AI world model seen as a learned, internal, predictive model of the environment, that is, a model which understands rules, physics, and dynamics of its surroundings, allows to predict future states from current and past states based on observations and potential actions. A few key representative world models are first introduced for explanatory purposes without being exhaustive of the rapid evolution of this research and technology development area. Basic model examples include the Joint Embedding Predictive Architecture (JEPA) [8], see Figure 1, and its domain-specific variants {Image-, Video-, Audio-, LLM-, ...} as well as a multimodal world model called Marble [9], see Figure 2. Physical AI enables intelligent, autonomous systems including cameras, robots, and self-driving cars perceive, understand, reason, perform, and orchestrate complex actions in the physical world. For the development of physical AI for autonomous vehicles (AVs), robots, and video analytics AI agents, NVIDIA offers the Cosmos platform [10]. This platform, purpose-built for physical AI, provides generative world foundation models (WFMs), guardrails, and an accelerated data processing and curation pipeline. Figure 3(a) shows the world foundation models included: autoregressive and diffusion-based WFMs, advanced video tokenizers, and a CUDA (Compute Unified Device Architecture), AI-accelerated data pipeline for video processing and curation. Based on text, image and video prompts, Figure 3(b) shows how virtual world states are generated as videos. Omniverse is used to build physics-based geospatial-accurate scenarios, its output can then be rendered into Cosmos. which then generates photo-real physically-based synthetic data, cf. Figure 3(c).

World models are commonly used for autonomous driving, recent surveys include [11]. While requiring active driver supervision, minimal intervention, and not making a vehicle autonomous, Tesla's Full Self-Driving (Supervised) [12] can be used for quick errands, daily commutes and road trips, intelligently and accurately completing driving maneuvers including route navigation, steering, lane changes, parking and more. It is conceived and trained with real-world driving data to take care of the most stressful parts of daily driving, making the roads safer for you and others. Figure 3(a), (b), and (c) show three maneuver examples: avoiding T-bones, stopping for a pedestrian,

*This abstract has been granted permission for public release. The author has among many others a 15-year tenure as founding Chief Scientist/EiC of an Elsevier Science's AI/ML journal [1] and is the youngest "Nobel"-Prize in Engineering (IEEE Fellow) in history worldwide with the following mention "for leadership in neural and parallel computation, and pioneering contributions to autonomous space robots" [2]. As a decade civil servant at the German Aerospace Center DLR, he launched with the Office of the German Federal Minister of Research and Technology the First Federal Program for AI/ML Research and Technology Development in Germany [3], which funded multiple consortia all over the country, one of them being his consortium including DLR and Siemens Corporate R&D in Munich, Germany. He has been conceiving, designing, building, and operating advanced scalable mission-critical parallel and distributed as well as secure (NSA, commercial) computer systems for federal and state governments and the commercial industry [4], e.g., AI pipelines deployed to and executing during NASA-ESA-DLR spaceflight missions, ML pipelines for DoD classified programs, advanced AI/ML pipelines for advanced data centers and vertical on-prem, hybrid, and in the cloud applications of the State of California. For decades he has been very active in leading the development of classified DoD, e.g., mission-critical edge AI/ML computation [5], USSF, e.g., most critical Golden Dome space segment systems [6], and NASA/ESA space AI/ML programs , e.g., the development of next gen Mars helicopters [7].

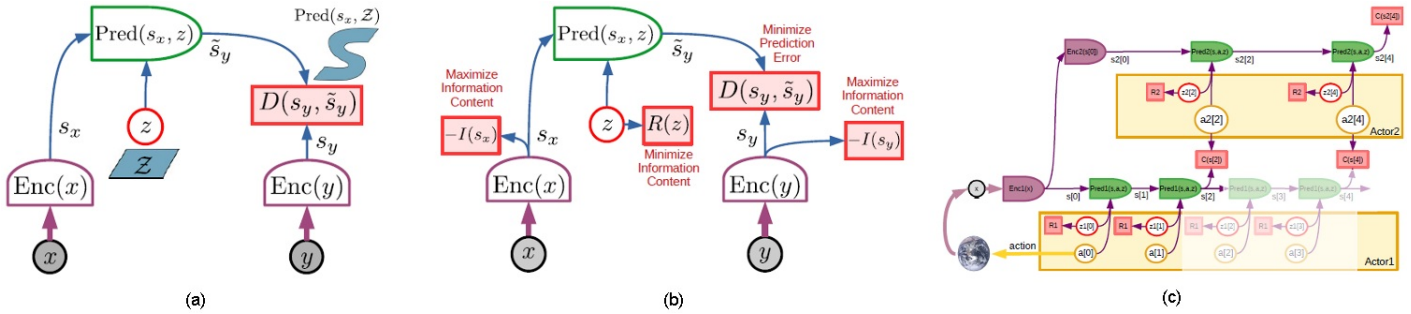

Figure 1: The Joint Embedding Predictive Architecture (JEPA) [8] (a) Generic architecture (b) Non-contrastive training (c) Hierarchical planning in an uncertain environment.

Figure 2: Marble, a Multimodal World Model [9] (a) Creating a full 3D world (b) Text-to-world generation (c) Multi-image-to-world generation.

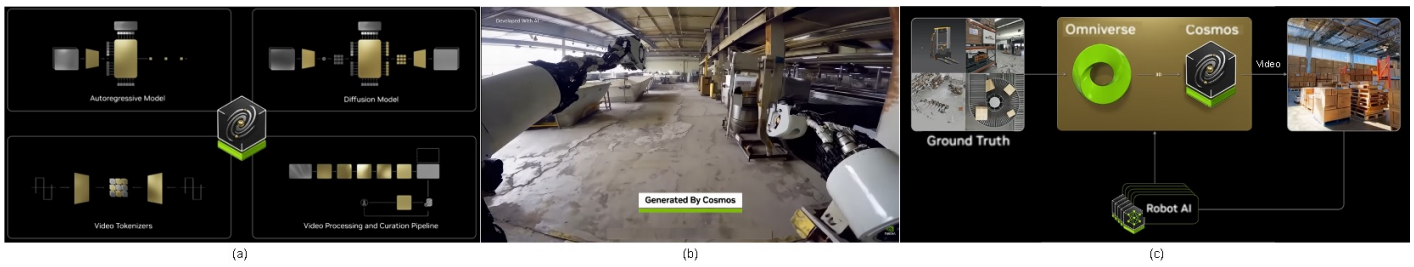

Figure 3: Cosmos, a World Foundation Model Platform for Physical AI [10] (a) Foundation models included (b) Generated virtual world states video (c) Omniverse scenarios to Cosmos to generate synthetic data.

Figure 4: Full Self-Driving (FSD) Supervised [12] (a) Avoiding T-bones (b) Stopping for a pedestrian (c) Reacting to a car door opening.

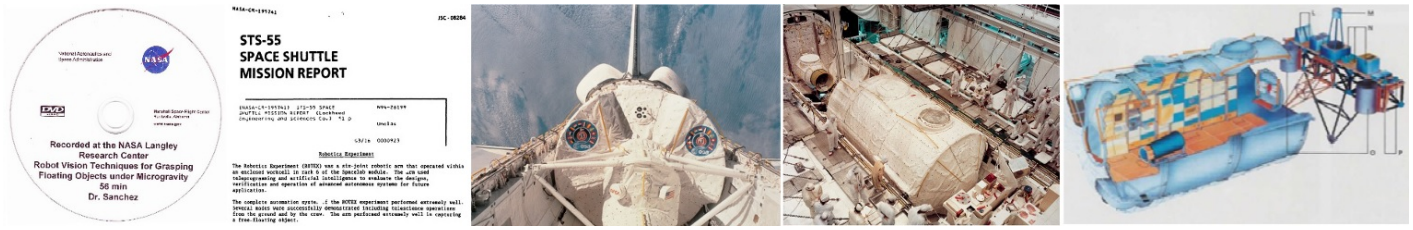

Figure 5: NASA Space Shuttle STS-55 Mission Report [14], Space Shuttle Cargo Bay, ESA Spacelab Integration [15], and ESA Spacelab Modeling.

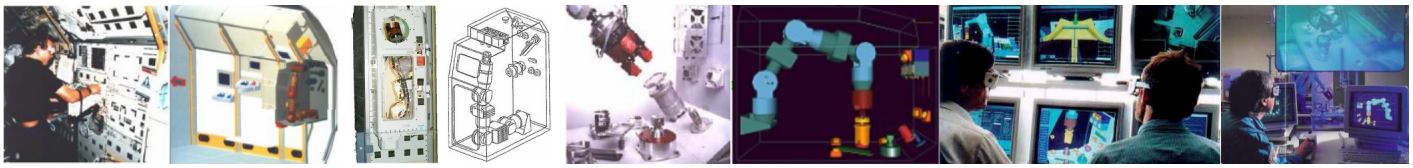

Figure 6: DLR Rotex Spacelab Rack (#6), Lab Mockup, Computer Graphics Modeling, Rendering of the entire Spacelab Workcell at the Ground Control Segment [16], Mission Operation with Stereo Display.

and reacting to a car door opening, respectively. FSD (supervised) and the attentive human driver are both looking for and avoiding hazards. If both can perform well, then together, human driver with FSD (supervised), can be expected to perform safer than a human driver on his own. The Tesla's vehicle safety report [13] shows that globally, vehicles driving with FSD (supervised) on have almost one and a half better safety record on city streets than similar vehicles driven manually.

In advanced intelligent, autonomous robotics, in particular for space mission-critical applications, but not only, AI world models have been used extensively. For example, way ahead of its time (a few decades), the group under my supervision at the German Aerospace Center DLR in Oberpfaffenhofen by Munich and industrial partners as well as national space agences including NASA and ESA, successfully introduced these technologies as first to the world and launched a NASA/ESA mission demonstrating them. A substantial part of those developments have already found their application path to several industrial and commercial applications. Figure 5 shows the corrresponding NASA Space Shuttle STS-55 mission report [14], the space laboratory fitting into the Space Shuttle cargo bay ESA Spacelab being integrated [15] as well as regarding the focus of this contribution: world models, the modeling of the ESA Spacelab. Figure 6 shows an ESA astronaut working inside the ESA Spacelab, the modeling of adajacent Spacelab racks, a photograph of rack #6 of Spacelab where the DLR space robot flew to space, a modeling of that rack including the space robot and all necessary tools for the experiment, a mockup of that workcell at the DLR robotics lab for development, testing, and integration, the computer graphics modeling of that workcell, the rendering of the entire Spacelab workcell at the Ground Control Segment [16], as well as testing mission operation with an stereo display at the DLR robotics lab.

Figure 7 shows an overall block diagram of the mission including the NASA Johnson Space Center (JSC) in Houston, TX, the German Aerospace Center DLR in Oberpfaffenhofen (OP), Germany as well as the NASA's Space Shutlle, the ESA's Spacelab, the DLR's Rotex space robot and the communication channel structure, one image of the Spacelab workcell (top) and the space robot gripper stereo cameras (bottom), a block diagram of DLR's Tele Sensor Programming (TSP) used during the NASA-ESA-DLR Rotex experiment flown and used for satellite servicing applications, and finally, key DLR OP infrastructure for satellite servicing applications. A detailed description of the work environment, i.e., a highly accurate world model, is indispensable for TSP to work, i.e., for the teleoperation and the fully autonomous operation of space robots as well as for all inbetween shared control modes of operation under long communication time delays since we in principle control the future shown in the operator display. The operator could be a teleoperator on Earth, an astronaut in space, or a fully autonomous space robot brain. We demonstrated all those modes of operation during the mission flown. That means that what is being shown in the display are multisensory as

Figure 7: DLR Tele Sensor Programming (TSP) [17]: in NASA-ESA-DLR Rotex Experiment flown, One Image of the Spacelab Workcell and Space Robot Gripper Stereo Cameras, TSP Block Diagram, in DLR Satellite Servicing Application Infrastructure.

Figure 8: Autonomous Operation of Space Robot inside ESA's Spacelab aboard NASA's Spaceshuttle. Real-Time Processing of Multisensory Data including Stereo Cameras and Laser Range Finders Based on Object's Geometric and Dynamic Modeling as well as Delay Compensating Modeling for Predictive 3D Graphics Simulation.

well as control data corresponding to the future when they will be happening in space, i.e., during the experiment, the present is typically on Earth, the future is in space. The time difference between the present and the future are precisely those long communication time delays. The world model, without which TSP could not work, is either generated off-line using exact knowledge of the geometry and the location of all parts of the work cell or a sensor-based world modeling approach. This world model needs to be updated to reconcile the virtual world description with the real environment, found by the available sensors [18].

Figure 8 shows the fully autonomous approach and final catching of the floating object under microgravity in the DLR Rotex Spacelab workcell during the mission, then also that not only gripper tiny stereo cameras were operating, but also laser range finders built into the gripper fingers, a white ray is noticeable in the third pic from the left, and finally some of the computer vision processing performed overlaying tracked pixels onto the image of the object edges. To be able to process multi-sensory data including stereo camera images in real-time I conceived, designed and built the most powerful mission-critical real-time supercomputer of its time [19], it was 100x, 1000x, ... faster than the Cray because it was scalable, could do anything that general-purpose supercomputer could do and much more and much faster. Figure 9 shows a block diagram of the machine, then a photograph with a highly preliminary version of that processing engine, during the mission a larger, more powerful version was used, I included even faster processors of different types (scalable parallel distributed architecture) as well as highly optimized custom algorithm libraries, and finally testing the

Figure 9: Scalable, Mission-Critical, Real-Time Parallel Distributed AI/ML Supercomputer [19] and Real-Time Computer Vision Processing to Track and Grasp a Floating Object under Microgravity.

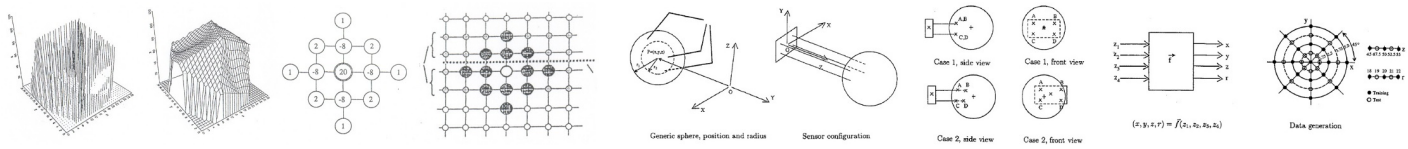

Figure 10: AI/ML World Modeling – Vision Processing [23] and AGI Training Simulation [24].

approach and catching of the flying object under microgravity with the space robot Rotex multisensory gripper. The machine provided the capability of processing all the artificial intelligence (AI) / computer vision (CV) algorithms including modeling and world model updates (geometric/static and dynamic) in real-time and delivered the continuously updating data over the VME Real-Time LAN interface to the SGI graphics rendering engines for convenience and for the aforementioned DLR's Tele Sensor Programming (TSP) approach running on additional separate general-purpose computers to operate successfully (teleoperation, fully autonomous control and all other modes inbetween). The graphics rendering could have been processed in the machine too since parallel graphics processing algorithms I had made available as well. Multiple displays were attached to the machine for displaying intermediate results, and among others, so too an interface to custom image and multisensory databases. Today lots of the processing power and respective modules could be packed into AI/ML/vision chips [20], there are though some highly specific mission-critical interfaces that would need to be added as realized in my machine to extremely successfully operate this type of missions (space, aerial, terrestrial, submarine) or embedded in miniaturized autonomous robotic agents [21], both virtual and real.

Now let us turn our attention to dynamic modeling. For visual dynamic scene understanding, for example, recursive estimation methods have been used in the so called 4D approach to dynamic vision [22] making use of spatio-temporal models of objects and processes observed. In the context of recursive estimation, the model equation of a continuous, nonlinear dynamic system is defined by $\dot{\vec{x}} = \vec{f}(\vec{x}(t), \vec{u}(t), \vec{z}(t), \vec{p}_s)$ with $\vec{x}(t)$ the n-vector of state variables, $\vec{u}(t)$ an r-vector of control variables, $\vec{z}(t)$ an n-vector of disturbances, and $\vec{p}_s$ a vector of system parameters that usually cannot be measured directly but only through an m-vector $\vec{y}(t_k)$ of output variables at discrete points t_k in time spoiled by some superimposed measurement noise $\vec{w}(t)$, whose statistics are assumed to be known sufficiently well. The standard form after the transformation of that model equation into a discrete linear state transition model for sampled data with cycle time T is $\vec{x}_k = A \cdot \vec{x}_{k-1} + B \cdot \vec{u}_{k-1} + \vec{v}_k$ where A is the $n \times n$ state transition matrix from $(k-1)T$ to kT, B is the $n \times r$ control effect matrix, and $\vec{v}(z(t), T)$ is the discrete noise term. The observable output variables $\vec{y}(k)$ depend nonlinearly on the state variables $\vec{x}(t)$ via the measurement process defined by $\vec{y}_k = \vec{h}[\vec{x}(t), \vec{p}_m] + \vec{w}(t)$, with $\vec{p}_m$ being the measurement parameters. The measurement process $\vec{h}$ in vision according to this particular approach is the perspective projection. Observability is assumed, that means that P, Q, R, the state error, process noise, and measurement noise covariance matrices, respectively, are known. When the nonlinear measurement equation is linearized around the predicted nominal state $\vec{x}_N$ of the process and the nominal parameter set $\vec{p}_N$ (q values), without the noise term, we obtain $\vec{y}^*(k) = \vec{y}_N(k) + \vec{\delta y}(k)$ with $\vec{y}_N(k) = \vec{h}[\vec{x}_N(t), \vec{p}_N, k]$, the Jacobian matrices of perspective mapping w.r.t. the state components and the parameters involved being $C_{\vec{x}} = \partial\vec{h}/\partial\vec{x}|_N$, $C_{\vec{p}} = \partial\vec{h}/\partial\vec{p}|_N$, and the measured prediction errors $\vec{\delta y}(k) = C_{\vec{x}}(k) \cdot \vec{\delta x}(k) + C_{\vec{p}} \cdot \vec{\delta p}$, expression used to detemine the deviations $\vec{\delta x}, \vec{\delta p}$ from the nominal values $\vec{x}_N, \vec{p}_N$.

We now outline the entire recursive estimation process. The process starts with an initial hypothesis $\vec{x}^*(0)$. Then find an estimate for a Gaussian probability distribution of this initial state $p(\hat{\vec{x}}_0) = \det(2\pi P_0)^{-1/2} \exp[-\frac{1}{2}(\vec{x} - \hat{\vec{x}}_0)^T P_0^{-1} (\vec{x} - \hat{\vec{x}}_0)]$, setting the mean value to $\hat{\vec{x}}_0 = \vec{x}^*(0)$ and P_0 is the $n \times n$ error covariance matrix. If the state components are independent of each other: $P_0[i,j] = 0$ for $i \neq j$ and $P_0[i,j] = \sigma_i$, the state variance components, for $i = j$. Then, incrementing the time index, the state variables $\vec{x}^*(k)$ can be predicted using the dynamic plant model $\vec{x}_k = A \cdot \vec{x}_{k-1} + B \cdot \vec{u}_{k-1} + \vec{v}_k$, and the predicted nominal measurement values $\vec{y}^*$ (no noise) can be computed using the the measurement model $\vec{y}^*(k) = \vec{h}[\vec{x}^*(k), \vec{p}_m]$, the noise model assumed being white noise: $E[\vec{w}] = 0, E[\vec{w}^T \cdot \vec{w}] = R$. Now

we can compute the prediction errors using actual measurement values $\vec{y}(k)$ and $\vec{\delta y}(k) = \vec{y}(k) - \vec{y^*}(k)$ and improve the state estimate using the innovation step defined by $\hat{\vec{x}}(k) = \vec{x^*}(k) + K(k) \cdot \vec{\delta y}^*(k)$. We will then use this new estimate for the state variables in the next recursion loop. $K(k)$ is a gain matrix, which in the stochastic scheme (Kalman filter), as opposed to the deterministic scheme (Luenberger observer) with no noise modeled, is called the Kalman-gain matrix. In this case, the dynamic plant model entails noise described by $E[\vec{v}] = 0, E[\vec{v}^T \cdot \vec{v}] = Q$. During the process called filter tunning, all assumed uncorrelated covariance values need to be appropriately determined. With this in mind, the process loop entails: computing the expected values using $\vec{x}_k^* = A(k-1) \cdot \hat{\vec{x}}_{k-1} + B(k-1) \cdot \vec{u}_{k-1}$. Then predict the expected error covariance matrix using $P^*(k) = A(k-1) \cdot P(k-1) \cdot A^T(k-1) + Q(k-1)$. Now compute the expected measurement values $\vec{y}^*(k) = \vec{h}[\vec{x}^*(k), \vec{p}_m]$ and the Jacobian matrix $C = \partial \vec{y}^* / \partial \vec{x}|_N$. Then compute the gain matrix for the prediction error feedback using $K = P^* \cdot C^T \cdot \{C \cdot P^* \cdot C^T + R\}^{-1}$ (minimizes the covariance matrix of the estimation error $P = E[\delta \hat{\vec{x}} \cdot \delta \hat{\vec{x}}^T]$ given the underlying constraints), update the state variables per innovation using $\hat{\vec{x}}(k) = \vec{x^*}(k) + K(k) \cdot [\vec{y}(k) - \vec{y}^*(k)]$, and update the error covariance matrix using $P(k) = P^*(k) - K(k) \cdot C(k) \cdot P^*(k)$. This concludes the basic version of the filter for explanatory purposes. There are several improvements when applied in practice.

Now we briefly present a key model of three-dimensional objects and how to generate their representation from noisy sensor data. Figure 10 shows a modeling of modeling approach for three dimensional visible surface reconstruction, also multi-resolution-capable, to model natural and man-made objects, based on a sparse map of depth data computed for example by the stereo vision processing of a pair of images. Shown are the sparse depth data generated when processing real image data from a stereo camera system and the reconstructed parameterizable surface function. Next the structure of one type of a so called computational molecule of the model is shown and then its placement on the border between two adjacent processors while performing parallel processing of the model to achieve real-time capability. For this modeling of modeling approach, the energy functional of the associated variational problem is of the type $\mathcal{E}(v) = \mathcal{S}(v) + \mathcal{P}(v)$, $\mathcal{S}(v)$ and $\mathcal{P}(v)$ are stabilizing and penalty functionals respectively, with $\mathcal{S}_{\rho\tau}(v) = \frac{1}{2} \int \int_\Omega \rho(x,y) \left\{ \tau(x,y)(v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2) + [1 - \tau(x,y)](v_x^2 + v_y^2) \right\} dxdy$ and $\mathcal{P}(v) = \frac{1}{2} \left\{ \sum_{i \in D} \alpha_{d_i} [v(x_i, y_i) - d(x_i, y_i)]^2 + \sum_{i \in P} \alpha_{p_i} [v_x(x_i, y_i) - p(x_i, y_i)]^2 + \sum_{i \in Q} \alpha_{q_i} [v_y(x_i, y_i) - q(x_i, y_i)]^2 \right\}$. ρ and τ are continuity control functions over the visual field Ω. In the underlying physical model, ρ and $(1 - \tau)$ represent the spatially varying surface cohesion and tension, respectively. d, p, q are the depth values and the components of the surface normal $\vec{n}(x_i, y_i)$, respectively, for further details including its FEM discretization, numerical computation, and parallel algorithms you can see [23]. It is a modeling of modeling approach because you model the model itself first in its basic capabiliies which btw. potentially makes it more adequate than current generic neural network models using deep learning and others, at least for certain modeling requirements and with higher modeling accuracy and computational throughput. Then the resulting model can model (any) generic three-dimensional object , i.e., its three-dimensional surface, efficiently.

We now present an example of how to generate training and test data in a virtual world via simulation to build artificial brains. The case exemplifies furthermore a way at an appropriate level of abstraction of how biological brains might attempt to learn capabilities as opposed to approaching a given task by simulating mathematical models of the world and solving the given task in a different logical, analytical way. Figure 10 also shows, this time on its right half, a robot three-finger gripper with four distance sensors – distance measurements are $z_i, i = 1 \cdots 4$ – and a generic ball (sphere) of radius r to be grasped in an idealized environment. The innovative generation of training and test data in a supervised learning setting is then outlined. Different situations are shown in which the distance measurements from the surface of the ball can occur; front and side views are shown and the corresponding data is generated in this virtual world via simulation. Once a self-designing neural network is trained and tested with the simulated data, the multivariate function $(x, y, z, r) = \vec{f}(z1, z2, z3, z4)$ has been approximated, and such neural network embedded into the learning robotic system can provide the robot gripper control with the necessary information to approach and grasp a ball of any size given sensor measurements. The description of the fully automatic creation of the neural network architecture from the training and test data [25] is beyond the scope of this paper.

The approach can be applied to other objects to be grasped. The generation of training and test data entails a certain roaming of the environment. So seen, this can be viewed as a generalizing of generalizing approach [26]. The inner generalization of the approach is to appropriately approximate the multivariate functions. The outer generalization of the approach extends the approach to the grasping of other objects.

In this report, the author analyzes and reports on the further development of previous and current AI world models, including some new, innovative of his own. In the wake of an accelerated international competition to dominate AI technologies, the U.S. Congress is seeking to establish a unified federal strategy for AI and modernize the government's research and regulatory approach to the technology [27]. Government regulations have been controversial, some might be necessary and some potentially harmful [28]. The matter is so delicate that for example, the current U.S. Administration (Executive) is apparently drafting an executive order that would direct the Justice Department to sue states that pass laws regulating AI [29]. At the U.S. DoD too, now the Department of War (DoW), there has been a recent realignment of the Chief Digital and Artificial Intelligence Officer (CDAO) to under the Office of the Under Secretary of Defense for Research and Engineering (USD R&E) to accelerate the Department-wide AI transformation [30]. In the private sector, the AI surge is the biggest market boom the world has seen [31]. The market capitalization of alone the 6 world's most highly valued companies {Nvidia ($5.1T), Apple ($4.0T), Microsoft ($4.0T), Alphabet ($3.3T), Amazon ($2.4T), Meta ($1.9T)}, all heavily involved in AI, add up to $20.7 trillion US dollars. A core aspect of the AI boom has been so far to invest in AI factories, that is, massive data centers hosting AI supercomputers made up of AI super chips processing current models, predominantly LLMs.

Much more powerful AI than currently available, including Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI), is claimed to become reality within the next few years [32]. There is on the other end, severe doubts about those expectations, at least in that short period of time [33]. That GenAI and LLMs are highy useful technologies, that human cognition is currently still beyond reach, and a more solid, subtantiated description of the AGI/ASI status quo is provided in [34]. Among others, multiple Natural Language Processing (NLP) tasks can be addressed using LLMs. Transformers have become the dominant architecture in those domains. If we go beyond language to behavior, i.e., develop artificial learning systems including intelligent, autonomous, learning robotic systems, then we need Large Behavior Models (LBMs) or the equivalent which learn from expert demonstrations diverse tasks including action patterns and contextual interactions. Rather than word prediction, emphasis is on actions, choices, and preferences and their derivations after training when exposed to unseen observations. LBMs enable us to train learning robots with much less data. Diffusion policies and human demonstrations (imitation) have been the key for those developments [35]. Some efforts are currently noticeable within the AI R&D, to move from generative models for language and images [36] to the development of world models that can simulate spatial relations and reason about the environment before interacting with it [37]. In particular, advances in scene rendering have been recently achieved, volume rendering is the generation of images from discrete samples of volume data. Techniques include, e.g., splatting introduced in [38], a feed-forward algorithm, i.e., a method that attempts to map each volumen element into the image plane and not the other way around (feed-backward method). It renders directly rectilinear volume meshes. 3D Gaussians are used for scene representation and high-quality real-time novel-view synthesis captured with multiple photos or videos [39]. They preserve desirable properties of continuous volumetric radiance fields for scene optimization. 3D Gaussian splatting (3D-GS) is for example used in the Marble world modeling approach. As was shown in the aforementioned NASA/ESA/DLR space mission, the degree of sophistication of the world models I designed and deployed has been extremely high even compared to current efforts. The further development of those applied world models, including as previously shown x-of-x world modeling and generalizing approaches at higher levels of abstraction closer to human cognition capabilities, is herewith being pursued to factually accomplish the next big step towards more powerful AI that can be even more utilitarian to all aspects of human life. Some examples of these developments enabling for instance robots to become more

human and obtain superhuman capabilities are presented in [40] including my work in the context of the NASA's Mars Exploration Program (MEP) to deliver the next gen autonomous Mars helicopters for the NASA-ESA Mars Sample Return (MSR) Campaign.

References

[1] V.D. Sánchez. *Neurocomputing 50th volume anniversary*. Neurocomputing, 50, ix, 2003. https://profdrvdsaphd.lima-city.de/documents/Neurocomputing50thAnniversary.pdf

[2] V.D. Sánchez. *IEEE Fellow Award – "for leadership in neural and parallel computation, and pioneering contributions to autonomous space robots"*. 1995. https://profdrvdsaphd.lima-city.de/documents/IEEEFellow.pdf

[3] V.D. Sánchez. *Personal written communication with the Office of the German Federal Minister of Research and Technology BMFT – Start of the First German National AI/ML Research and Technology Development Program*. 1988.

[4] V.D. Sánchez. *Advanced Automation for mission-critical Information Technology past AIOps*. November 2024. https://profdrvdsaphd.lima-city.de/documents/AdvancedAutomationMisCriITPastAIOps.pdf

[5] V.D. Sánchez et al. *The Design of a Real-Time Neurocomputer Based on RBF Networks*. Neurocomputing 20 (1998), 111–114.

[6] V.D. Sánchez. *Next Generation Missile Defense Systems including the Golden Dome*. May 2025. https://profdrvdsaphd.lima-city.de/documents/NextGenerationMissileDefenseSystemsInclGoldenDome.pdf

[7] V.D. Sánchez. *Colonizing the Red Planet*. November 2022. https://profdrvdsaphd.lima-city.de/documents/MarsColonization.pdf

[8] Y. LeCun. *A Path Towards Autonomous Machine Intelligence*, New York University, Courant Institute of Mathematical Sciences, Version 0.9.2, June 27, 2022.

[9] World Labs. *Marble: A Multimodal World Model*. 2025. https://www.worldlabs.ai/blog/marble-world-model/

[10] NVIDIA. *Physical AI – NVIDIA Cosmos*. 2025. https://www.nvidia.com/en-us/ai/cosmos/

[11] T. Feng et al. *A Survey of World Models for Autonomous Driving*. 10 Sep 2025, version v4 https://doi.org/10.48550/arXiv.2501.11260

[12] Tesla. *Full Self-Driving (Supervised)*. 2025. https://www.tesla.com/fsd/

[13] Tesla. *Full Self-Driving (Supervised) – Vehicle Safety Report*. 2025. https://www.tesla.com/fsd/safety/

[14] National Aeronautics and Space Administration (NASA). *STS-55 Space Shuttle Mission Report*. NASA-CR-195741, July 1993.

[15] European Space Agency (ESA). *Spacelab*. In BR-250, ESA Achievements – more than thirty years of pioneering space activity, 108–113, June 2005.

[16] Deutsches Forschungszentrum für Luft- und Raumfahrt (DLR). *ROTEX (1988-1993)*. https://www.dlr.de/en/rm/research/robotic-systems/hands/rotex-1988-1993/

[17] G. Hirzinger. *Multisensory shared autonomy and tele-sensor programming — Key issues in space robotics*. Robotics and Autonomous Systems 11 (1993) 3–4, pp. 141–162.

[18] K. Landzettel et al. *The Telerobotics Concepts for ESS*. https://www.robotic.dlr.de/fileadmin/robotic/bernhard/ess94.pdf

[19] V.D. Sánchez. *Eine parallel-verteilte Architektur für Rechnersehen und Telerobotik*, in M. Baumann und R. Grebe (Hrsg.): Parallele Datenverarbeitung mit dem Transputer, Reihe Informatik Aktuell, Springer-Verlag, Berlin, 1993, pp. 295–303.

[20] V.D. Sánchez. *Modern Machine Learning Technology*. December 2017. https://profdrvdsaphd.lima-city.de/documents/ModernMachineLearningTechnology.pdf

[21] V.D. Sánchez. *Architecting Next-Generation Generative AI/ML Models and Agentic AI Solutions*. August 2025. https://profdrvdsaphd.lima-city.de/documents/ArchitectingNextGenAIML AgenticAISolutions.pdf

[22] E. Dickmanns. *The 4D-approach to dynamic machine vision*. In Proceedings of the 33rd IEEE Conference on Decision and Control, Lake Buena Vista, FL, 14–16 December 1994.

[23] V.D. Sánchez. *Parallelverarbeitende dreidimensionale Oberflächenrekonstruktion*, in J. Hektor und R. Grebe (Hrsg.): Parallele Datenverarbeitung mit dem Transputer, Reihe Informatik Aktuell, Springer-Verlag, Berlin, 1994, pp. 238–249.

[24] V.D. Sánchez. *Beiträge zum überwachten Lernen reellwertiger Funktionen mithilfe neuronaler Netze*. Deutsches Zentrum für Luft- und Raumfahrt (DLR) Bericht, unveröffentlicht, 1993.

[25] V.D. Sánchez. *Searching for a solution to the automatic RBF network design problem*. Neurocomputing 42 (2002), 147–170.

[26] V.D. Sánchez. *Künstliche Intelligenz und Maschinelles Lernen zur Lösung ungelöster Probleme in der Mathematik (in German)*. March 2025. https://profdrvdsaphd.lima-city.de/documents/ArtificialGeneralIntelligenceMachineLearningOnMathematics.pdf

[27] K. Smith. *House Bill Seeks to Create Unified Federal AI Strategy as Global Competition Accelerates*. ExecutiveGov November 26, 2025. https://www.executivegov.com/articles/house-ai-strategy-bill-jen-kiggans/

[28] V.D. Sánchez. *Deep Impact of Advanced AI Technology Developments – Government Regulatory Measures –*. September 2023. https://profdrvdsaphd.lima-city.de/documents/Artificial IntelligenceRegulation.pdf

[29] The Washington Post. *White House drafts order directing Justice Department to sue states that pass AI regulations*. Updated November 20, 2025. https://www.washingtonpost.com/technology/2025/11/19/trump-order-ai-sue-states/

[30] CDAO. *CDAO Re-alignment to USD(R&E) Accelerates AI Transformation at DoD*. August 20, 2025. https://www.ai.mil/Latest/News-Press/PR-View/Article/4281147/cdao-re-alignment-to-usdre-accelerates-ai-transformation-at-dod/

[31] BBC. *The contradiction at the heart of the trillion-dollar AI race*. November 19, 2025 https://www.bbc.com/news/articles/cvgvynlxqdyo/

[32] G. Grossman. *AGI is coming faster than we think – we must get ready now*. Venture Beat, November 10, 2024. https://venturebeat.com/ai/agi-is-coming-faster-than-we-think-we-must-get-ready-now/

[33] W.D. Heaven. *How AGI became the most consequential conspiracy theory of our time*. In An MIT Technology Review series, The New Conspiracy Age, October 30, 2025. https://www.technologyreview.com/2025/10/30/1127057/agi-conspiracy-theory-artifcial-general-intelligence/

[34] V.D. Sánchez. *On Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI)*. January 2025. https://profdrvdsaphd.lima-city.de/documents/OnArtificialGeneralIntelligenceAndArtificialSuperIntelligence.pdf,

[35] V.D. Sánchez. *AI/ML Methods to Develop Superior Next Gen Autonomous Learning Robot Systems for Industrial Terrestrial and Space Applications*. December 2024. https://profdrvdsaphd.lima-city.de/documents/AIMLNextGenAutonLearnRobots.pdf

[36] V.D. Sánchez. *Advanced Gen AI & LLM Foundations and Applications – Paving the way to a more powerful and diverse ML –*. December 2023. https://profdrvdsaphd.lima-city.de/documents/AdvancedGenAILLMs.pdf.

[37] G. Wang. *AI Has Mastered Words And Images. Now It's Entering The Physical World*. Forbes, Updated November 14, 2025.

[38] L.A. Westover. *Splatting: A Parallel, Feed-Forward Volume Rendering Algortihm*. Ph.D. thesis, The University of North Carolina at Chapel Hill, July 1991.

[39] B. Kerbl et al. *3D Gaussian Splatting for Real-Time Radiance Field Rendering*. ACM Transactions on Graphics (TOG) 42 (2023) 4.

[40] V.D. Sánchez. *Enabling Robots to become more human and obtain superhuman capabilities for terrestrial and space applications using advanced AI/ML*. November 2024. https://profdrvdsaphd.lima-city.de/documents/EnablingSuperhumanRobotsUsingAdvancedAIML.pdf