BJMB
Brazilian Journal of Motor Behavior

XV[th] International Symposium on Biomechanics and Medicine in Swimming —Porto Alegre, Brazil

# The León Lewillie Lecture

# The Science of Swimming in the New AI-Driven World

RAÚL ARELLANO[1]

[1] Aquatics Lab, Department of Physical Education and Sport, Faculty of Sport Sciences, University of Granada, Granada, Spain; r.arellano@ugr.es

Correspondence to: Raúl Arellano, Universidad de Granada. **r.arellano@ugr.es**

ABSTRACT

For more than half a century, swimming resisted scientific analysis. When Leon Lewillie opened the first International Symposium on Biomechanics and Medicine in Swimming in Brussels in 1970, he identified the obstacle precisely: a barrier of water had isolated our sport from rigorous measurement. This lecture explores how this barrier became permeable. Starting with Lewillie's balloon-borne electromyograms, it follows the development through Jan Pieter Clarys's five decades of towing-tank and muscle work, which demonstrated that technique rather than body form governs resistance, and on to my own nine editions of this symposium, where the focus shifted from developing instruments to understanding the properties of water itself. The second half turns to the force now reshaping the field: artificial intelligence. Computer vision can already segment races and measure the underwater dolphin kick, while wearable sensors can classify strokes and explainable models can rank the variables behind performance and flag injury risk before it occurs. However, these tools can produce inaccurate results, hide their reasoning and have not read much of our literature. The argument is simple yet demanding: machine learning provides scale and speed, while fluid mechanics provides cause. Only their collaboration — with the scientist firmly in the loop — yields understanding. Fifty-six years after Lewillie, the water still resists; the machine is powerful, yet blind; and the researcher must still ask the right questions and stand by the answers.

**KEYWORDS:** swimming biomechanics | vortex dynamics | active drag | surface electromyography | machine learning

## INTRODUCTION

One of the most interesting aspects of being a guest speaker at the 'International Symposium on Biomechanics and Medicine in Swimming', delivering the opening lecture, is having to review numerous papers from the conference archives. Although I have all the books published by the organisers in my possession, everything is now digitised and organised thanks to Prof. Dr. Em. Jan Pieter Clarys, PhD, Professor of Experimental Anatomy at the Vrije Universiteit Brussel and a long-standing Chair of the Steering Group of the Biomechanics and Medicine in Swimming Congress. Sadly, he passed away a few years ago, but he funded this enormous task of digitising the proceedings.

I found it fascinating to revisit their contents. I believe their greatest value lies in providing young researchers with documentation that enables them to understand how science has been applied to swimming since 1970. Some might consider this information to be outdated or irrelevant today, but this could not be further from the truth. The elegant work published in these proceedings by numerous researchers since the 1970s has laid the foundations for our current investigations, avoiding the all-too-common issue of research being published whose results have been known for over 40 years. This alone makes the proceedings invaluable.

None of this would have happened without Leon Lewillie, a dedicated researcher into applying EMG to swimming. He was instrumental in the organisation of the inaugural symposium in Brussels, Belgium, in 1970, and subsequently presided over the proceedings of the second symposium in 1974. The initial volume comprises an introduction authored by L. Lewillie[1] and a foreword by James E. Counsilman[2], integrating the most advanced scientific knowledge with the practical application of swimming. Notably, within a few years, one of Counsilman's swimmers at Indiana University proceeded to secure a remarkable seven gold medals at the Munich Olympics in 1972 – a feat achieved by none other than Mark Spitz. It is important to note that many of our recent graduates and PhD holders are not acquainted with both the swimmer and the coach, indicating a significant knowledge gap regarding the intricacies of our sport. As Lewillie began his introduction, referring to the sport of swimming: 'More than any other branch of biomechanics, swimming has been isolated from scientific methods of investigation by a barrier to water'.

That barrier defined his own working life. Lewillie enjoyed recalling that nineteenth-century engineers built machines for teaching swimming without water[3]; his own career ran the other way — get the instruments wet, then keep them honest. He telemetered the electromyogram out of the pool through antennas slung from balloons and, more durably, took a maximal isometric contraction poolside before every trial, because electrode placement and skin conductivity leave absolute values meaningless even within one swimmer[4]. Aquatic EMG is still normalised that way. His triceps

traces climbed from 65% of that reference in slow crawl to 180% at sprint, relaxation almost gone, the recovering arm no longer passive — though he warned that beyond 100% the extrapolation cannot be verified[4]. With Clarys, two lamps taped to the deltoid and the distal ulna and a shutter held open showed the wrist path barely shifting as throwing force rises while the shoulder path changes a great deal[5]. The same cheap trick caught the discipline's own instruments out: tethering shrinks the front crawl hand path to about half its free-swimming diameter[3]. His 1983 review gathers the two figures that still argue for technique over volume — Clarys's active drag at twice passive and the 7% arm efficiency from flume work[3]. Presiding over the Brussels symposia and editing their proceedings[3, 6], he kept insisting that instrumentation earns its place only by leading to 'a better understanding of man'[6]. Table 1 lists all the Symposia, noting that the first two were organised by L. Lewillie.

**Table 1.** History of the International Symposium on Biomechanics and Medicine in Swimming (1970–2026). The Symposium — the oldest scientific event in aquatic sports — has been held every four years since 1970. For each of the fifteen editions the table gives the edition number (Roman and ordinal), the year, the host city and country, and the proceedings editors; for the forthcoming 2026 edition, the chair and local organizing committee are listed instead. *Country* denotes the host country of the congress.

| Ed. | Year | Host (city, country) | Editors / Organizers |
|---|---|---|---|
| **I (1st)** | 1970 | Brussels, Belgium | Lewillie, L.; Clarys, J. P. |
| **II (2nd)** | 1974 | Brussels, Belgium | Clarys, J. P.; Lewillie, L. |
| **III (3rd)** | 1978 | Edmonton, Canada | Terauds, J.; Bedingfield, E. W. |
| **IV (4th)** | 1982 | Amsterdam, Netherlands | Hollander, A. P.; Huijing, P. A.; de Groot, G. |
| **V (5th)** | 1986 | Bielefeld, Germany | Ungerechts, B. E.; Wilke, K.; Reischle, K. |
| **VI (6th)** | 1990 | Liverpool, United Kingdom | MacLaren, D.; Reilly, T.; Lees, A. |
| **VII (7th)** | 1994 | Atlanta, United States | Troup, J. P.; Hollander, A. P.; Strass, D.; Trappe, S. W.; Cappaert, J. M.; Trappe, T. A. |
| **VIII (8th)** | 1998 | Jyväskylä, Finland | Keskinen, K. L.; Komi, P. V.; Hollander, A. P. |
| **IX (9th)** | 2002 | Saint-Étienne, France | Chatard, J.-C. |
| **X (10th)** | 2006 | Porto, Portugal | Vilas-Boas, J. P.; Alves, F.; Marques, A. |
| **XI (11th)** | 2010 | Oslo, Norway | Kjendlie, P.-L.; Stallman, R. K.; Cabri, J. |
| **XII (12th)** | 2014 | Canberra, Australia | Mason, B. |
| **XIII (13th)** | 2018 | Tsukuba, Japan | Takagi, H.; Ohgi, Y.; Sengoku, Y.; Gonjo, T. |
| **XIV (14th)** | 2023 | Leipzig, Germany | Staub, I. |
| ***XV (15th)* *** | *2026* | *Porto Alegre, Brazil* | *Castro, F. A. S. (Chair); Zacca, R.; Geremia, J.* |

** Edition XV (Porto Alegre, 2026) is forthcoming; proceedings not yet published. Sources: BMS database (https://bms.sport-iat.de/) and BMS 2026 website (https://www.bms2026.com.br/).*

## Form, Force and Specificity: Jan Pieter Clarys and Five Decades of Aquatic Measurement

Despite never having had the opportunity to meet Professor Leon Lewillie in person, I was fortunate enough to have met Jan Pieter Clarys. He was a generous individual who gave me the opportunity to observe directly the nature of the research he was conducting in his laboratory — an experience that will forever remain etched in my memory. This kind of behaviour is indicative of the highest calibre of scientists.

The concepts proposed by J.P. Clarys commenced with an enquiry derived from the discipline of naval architecture and concluded with its dismantling. Between 1970 and 1979, at the 200-metre towing tank in Wageningen, Clarys undertook a study in which he measured the resistance of towed and self-propelled swimmers against a full

anthropometric description of each subject [7,8,9]. Two results emerged that were at odds with the prevailing perspective. The drag recorded at a depth of 60 cm below the surface exceeded that at the surface (13.14 kg against 11.01 kg at 1.9 $m \cdot s^{-1}$), indicating that the elimination of wave-making resistance incurs greater pressure and eddy resistance costs than it saves [8]. Furthermore, the flow around the body exhibits turbulence throughout the competitive velocity range, thereby resulting in frictional drag being proportionally trivial. No association was found between passive drag and body surface area, nor with the H2/S ratio [9]. The magnitude of active drag surpassed that of passive drag by a factor of 1.5 to 2, thereby precluding the possibility of inferring propulsive force from towed measurements. A mere four anthropometric variables demonstrated a correlation with active drag, and none of these variables exhibited a strong correlation (r = 0.40–0.46). The form of the given subject was not accounted for in terms of resistance; instead, the factors of alignment, technique and even water temperature were considered relevant factors in the experiment [9, 10].

That conclusion set the second problem. If technique governs resistance, the object of measurement must become the musculature that produces it, and the Brussels Swimming EMG Project was built to that end[11]. Its lasting contribution is methodological rather than descriptive. Integrated signals were planimetered and expressed as a percentage of maximal isometric contraction, which rendered patterns comparable between swimmers of unequal ability and yielded normalised diagrams for twenty-five superficial muscles of the front crawl[12, 13]. The instrumentation followed the same reasoning: two-channel telemetry gave way to active electrodes with preamplifiers above 1 GΩ and multichannel FM recording, since immersion itself attenuates the detectable output of a contracting muscle[11]. The substantive finding was unexpected. What separated competitive from non-competitive swimmers lay not in the arms but in the trunk and pelvic girdle — latissimus dorsi at 92.3 against 23.7 per cent of maximal isometric contraction, gluteus maximus pars inferior at 122.4 against 31.2[11, 12]. Total efficiency in the front crawl, measured over the same period, stood below 5 per cent[14].

The subsequent analysis of the normalised electromyography data provided a criterion for the determination of the specificity of the given exercise. Furthermore, the aforementioned approach was applied consistently over the course of the following two decades, irrespective of the commercial standing of the device under evaluation. The dry-land apparatus demonstrated deficiencies in all aspects of the study, including cycle durations, signal amplitudes, peak counts, and coordination, which deviated from the established swimming pattern [15]. Tethered swimming, the most common type of hand paddle, and the MAD system were all found to be effective [16,17,18]. The demand for verification was then extended to the methods of the discipline itself in the foreword to the seventh symposium proceedings, where maximal voluntary contraction was questioned as a normalisation reference and the skinfold literature was shown to rest on assumptions that anatomy does not support[19].

## Nine Symposia, Thirty-Six Years: How My Questions Changed

The series was inaugurated in Brussels in 1970, under the auspices of the individual who would subsequently become the subject of this lecture. Following my arrival in Liverpool two decades later, I have had the privilege of attending nine out of the fourteen editions of this symposium that have been held since my arrival, which equates to almost two-thirds of the symposium's existence to date. In retrospect, it is evident that my initial involvement was in the capacity of an electronics technician. The issue that piqued my interest at the time was not, strictly speaking, a theory of swimming; rather, it was the problem of the coach standing at the poolside with an opinion and no evidence to support it. Consequently, I engaged in the construction of various objects. The device under consideration is a camera encased in a plastic housing, the purpose of which is to follow a swimmer along the length of a pool. The data that was collected in Liverpool had been collected the previous autumn [20]. This document is a sample of my doctoral thesis work, which had virtually no publications (it was a different era and we were not yet aware of the importance of international publications). The aforementioned method does not accurately reflect the effort required to measure force in water using the technique of 'fully tethered swimming'. This encompasses the development of software in machine language to read the memory locations where the data from an analogue-to-digital converter was stored, as well as programming the graphical representation of the records. This work was conducted in the late 1980s, a period when the personal computer was in its infancy and had not yet become ubiquitous. Four years later, in Atlanta, a set of photocells and a touchpad that could time each phase of a start and hand the swimmer the number before the next repetition was brought[21]. The documentation from the initial period was consistent with the form of equipment reports, and indeed, they were precisely that. The ambition was modest yet arguably correct: to create a quantifiable entity that had previously only been describable.

The catalyst for this change of opinion was a persistent frustration. As the quality of my instruments improved, it became increasingly apparent that a mere description of the swimmer was insufficient. Despite the ability to measure movement with precision, the underlying reasons for the observed variations in velocity among athletes remained elusive. It was evident that the solution to the enigma was not to be found within the body itself; rather, it resided in the water, from which I was unable to observe. Irrespective of the fact that I was employed at the Sierra Nevada High-Altitude Training Centre at that particular time (As head of the biomechanics laboratory). This experience marked a significant transition, transforming me from a biomechanist to an apprentice in fluid dynamics, a shift that occurred during my journey between Jyväskylä and Saint-Étienne. The present author initiated a programme of research into fish and dolphin behaviour, borrowing their approach to efficiency. The present author began working with people from wind engineering and applied physics, rather than only from the field of sport science [22, 23]. The invited lecture delivered in Porto in 2006 is the clearest marker of this shift: It was argued that the vocabulary employed, which described propulsion as a balance of lift and drag, was erroneous. The wake left behind by a swimmer was proposed as a more appropriate concept with which to replace it.[24] In addition, a comprehensive list was compiled, detailing all the aspects that had not yet been realised. This was preceded by contributions to the ISBS symposia, with a particular focus on the paper 'Vortices and Propulsion'.

The third change was less intellectual and more human. Somewhere along the way I stopped being the person who built the device and became the person who trains the people who build them. From Oslo onwards, most of the work I brought here was not mine alone: it belonged to doctoral students working on starts, on age-group development, on race analysis, on physiology[25, 26]. It also stopped being only about competitive swimming. In Canberra we came with a study of how people walk in water rather than on land, aimed at rehabilitation rather than at performance[27]. That is a different kind of research life, and I did not anticipate how much I would enjoy it.

That Canberra study is also, in hindsight, where we reached the ceiling of the old way of working. It was the most complete analysis we had ever done — three-dimensional, four cameras, corrected for the refraction of the water — and the price of that completeness was every marker placed by hand and every frame digitised by hand, for eight volunteers walking ten metres in a swimming pool that is waist-deep.

Nine years later, in Leipzig, I found myself confronted once again with the question I had first posed in the nineties: how does underwater undulatory swimming actually work?[28] However, on this occasion, the tools at hand finally aligned with the ambitious goals. The purpose of this lecture is to elucidate the events that transpired there. Thirty-three years had elapsed, yet nothing had changed with regard to my curiosity. The most significant change was the shift in the economic burden of data extraction from images. This process, which had historically been the most costly aspect of my laboratory's operations, had become less financially intensive. Porto Alegre constitutes my tenth visit, and it is the first occasion on which I have been able to travel without the aforementioned constraint, the only limitation being my impending retirement.

As is evident from the references, a significant number of Master's and PhD students have contributed to the conference papers. It is only appropriate to express the deepest gratitude for their exceptional work. In 2020, the University of Granada designated the 'Aquatics Lab' at the Faculty of Sports Sciences as a Singular Laboratory, in recognition of the infrastructure it had developed to enable the automatic analysis of competitions in the facility's 25-metre pool. This project was launched in late 2014. Consequently, in 2024, European Aquatics accredited the laboratory as a 'Sport Sciences Lab'.

**Table 2**. Themes of the author's contributions to the Symposium, the BMS editions at which they were presented, and the number of contributions per theme (27 contributions in total).

| Theme of my contributions | BMS edition–year | N |
|---|---|---|
| Underwater undulatory swimming (UUS) | VIII–1998 · IX–2002 · X–2006 · XIV–2023 (×2) | 5 |
| Race components and competition analysis | VIII–1998 · IX–2002 · XI–2010 (×2) · XIII–2018 | 5 |
| Swimming starts (devices, force, specific training, warm-up) | VII–1994 · IX–2002 · XI–2010 · XIV–2023 | 4 |

| Theme of my contributions | BMS edition–year | N |
|---|---|---|
| Qualitative evaluation of technique (error frequency: front crawl, starts, turns) | IX–2002 · X–2006 (×2) | 3 |
| Propulsion and hydrodynamics (fundamentals, flow visualisation, sculling) | X–2006 (Invited Lecture) · XII–2014 | 2 |
| Age-group development and longitudinal follow-up | X–2006 · XIV–2023 | 2 |
| Physiology and environment (altitude, water temperature) | X–2006 · XIII–2018 | 2 |
| Training methods and load monitoring (tethered swimming, added resistance) | VI–1990 · IX–2002 | 2 |
| Intra-cyclic velocity and technical drills | XI–2010 | 1 |
| Aquatic locomotion and health (walking in water) | XII–2014 | 1 |
| TOTAL | | 27 |

Our figures cannot compare with those of the leading contributor to this conference, our friend Joao Vilas-Boas, who has presented more than 60 documents, from the University of Porto (Portugal).

## The New AI-Driven World

It is noteworthy that language models possess a remarkable degree of self-awareness, encompassing the very theories that govern their own development and the application of their so-called 'knowledge'.

Nevertheless, this introspective awareness offers a unique perspective, illuminating the aspects that are not present. At the time of the inception of these systems, there was a paucity of knowledge regarding swimming as a scientific discipline. The capacity to describe, in general terms, the mechanics of the front crawl stroke and the benefits of streamlined body shape for speed was present, but the theoretical underpinnings of the field, such as the propulsive role of vortices, the intra-cyclic fluctuations of velocity, and the energetics of the underwater phase, were largely esoteric. Their fluency masked an almost complete absence of the specialised, hard-won knowledge stored in volumes such as these proceedings.

This paradigm has undergone a shift, a transformation that can be attributed, at least in part, to our own actions and interventions. As researchers digitise their archives, publish open-access articles, share datasets and discuss their findings online, the material enters the corpora on which subsequent models are trained; the swimming science a model can now draw upon is incomparably richer than it was even three years ago. However, it is imperative to maintain a sense of proportion. When measured against the volume of text absorbed in other domains – the billions of lines of software source code, the entire corpus of Wikipedia, vast repositories of general web pages, news and social media, and the immense biomedical and legal literatures – the entire corpus of aquatic sport science is negligible. The model in question has demonstrated a notable lack of knowledge regarding the sport of swimming, a fact that should temper both our expectations and our trust in the model.

A concise overview of terminology is instrumental prior to the proliferation of acronyms. The field of artificial intelligence (AI) encompasses the development of machines capable of performing tasks that are typically considered to require intelligence. Machine learning (ML) is the prevailing branch of this field, whereby systems are not provided with explicit rules, but instead learn patterns directly from data. Deep learning (DL) is a subset of machine learning that is built on many-layered artificial neural networks, and it is the engine behind most of the breakthroughs of the last decade — image recognition, speech and the large language models at the centre of this discussion. The three constitute a set of nested circles: every large language model is a deep-learning system, every deep-learning system is machine learning, and all of it lives under the umbrella of artificial intelligence. It is imperative to bear this hierarchy in mind whenever the acronyms AI, ML and DL are employed subsequently.

The phrase that gives this lecture its title — 'The New AI-Driven World' — is more than a rhetorical flourish. Over little more than a decade, artificial intelligence and machine learning have moved from the margins to the methodological core of sport science, where they are now routinely used to recognise movement patterns, predict performance and estimate injury risk.[29, 30, 31] Swimming, long constrained by Lewillie's 'barrier due to water', is being reshaped by the same forces, as multimodal video and sensor systems begin to dissolve the very obstacle that once isolated our sport from rigorous measurement.[32, 33]

On the positive side, the gains are tangible. In the wider field, deep-learning models now classify sport-specific actions and forecast competitive outcomes with an accuracy unattainable by manual observation alone.[29, 34] In swimming science the benefits are equally concrete: large databases of race results can be mined to reveal age- and stroke-dependent development trends,[35] multimodal video and inertial systems can capture and interpret starts, turns and stroke cycles,[32, 33] and explainable-AI methods such as SHAP (SHapley Additive exPlanations) can rank the demographic, anthropometric and biomechanical variables that drive performance while remaining intelligible to the coach.[36] Used well, such tools compress analysis time, personalise training and, crucially, return interpretable insight rather than an opaque verdict.

It is imperative to acknowledge that this same power carries with it significant hazards. It is evident that numerous high-performing models persist in their state of being 'black boxes', thereby rendering them inaccessible to the inspection of coaches and scientists. This phenomenon has the potential to erode trust and accountability. The dearth of standardised protocols, the paucity of datasets, and the limited external validation continue to undermine the principles of reproducibility and real-world deployment. It is important to note that a model may report a 98% accuracy rating while having acquired no useful knowledge. For instance, if two per cent of athletes are injured, a model that predicts 'no injury' will achieve that outcome precisely. Consequently, the rigorous evaluation that guards against such artefacts remains non-standard practice in this field. In the aquatic environment, these problems are especially acute, due to the scarcity and difficulty of labelling reliable underwater data. Beyond the laboratory, there are subtler dangers. An over-reliance on automated output may result in the deskilling of young researchers. Generative models have the capacity to hallucinate references and fabricate plausible yet false results.[39] Finally, there is the issue of unequal access. It is evident that the issue of computation and data risks is exacerbating the discrepancy between well-resourced and modest programmes. The lesson for our field is not uncritical enthusiasm or reflexive rejection, but disciplined adoption — powerful assistance held firmly within the human-in-the-loop, with the scientist remaining the final judge of every claim.

## Inside the Black Box: How Large Language Models Are Trained

In the event of a proportion of scientific work being delegated to such systems, it is incumbent upon us to undertake a thorough examination of their fundamental nature, and to acknowledge the limitations in the ability of their designers to provide comprehensive explanations. The models that are currently generating our code and summarising our literature have not been programmed with explicit rules in the manner that earlier software was; rather, they have been developed through a training process whose inner workings remain only partially legible.

The initial and most substantial stage is that of pre-training. A neural network constructed on the transformer architecture — the design unveiled in 2017 that enables a model to evaluate the relevance of each word to every other through a mechanism termed self-attention [40] — is exposed to an extensive corpus of text, frequently comprising a substantial proportion of the public internet, in addition to books, articles and code. The objective of the task is ostensibly straightforward: given a sequence of words, predict the subsequent word. When applied to a vast array of examples, this next-token prediction compels the network to encode, within the numerical values of its billions of internal parameters, a substantial and intricate statistical model that elucidates the manner in which language, and, incidentally, a significant proportion of the knowledge expressed in language, is interconnected.

Two features of this process deserve emphasis. First, scale is decisive: capabilities absent in smaller models appear, seemingly abruptly, as the number of parameters and the volume of training data grow — a phenomenon of emergence that even specialists struggle to predict.[41] Second, the resulting artefact is general-purpose; the same foundation model can be adapted to translation, summarisation, coding or question-answering without having been explicitly designed for any of them.

The second stage of the process then transforms the raw predictor into a form that can be utilised as an assistant. Through a process of supervised fine-tuning on curated examples and reinforcement learning from human feedback,

the model is guided towards responses that are perceived as helpful, honest and harmless by people. This alignment is what engenders the sense of cooperation in a chatbot; however, it also introduces an additional layer of opacity, given that the model's conduct is shaped by aggregated human preferences that themselves pose challenges in terms of auditing. It is possible for models to be constructed from other models. Through the process of knowledge distillation, a large and capable 'teacher' model is employed to train a smaller 'student' model, which learns to reproduce the outputs of the teacher. In this manner, a significant proportion of the teacher's competence is transferred into a form that is more economical, expeditious and deployable.[42]

The central issue of the black-box problem can be summarised as follows. The model's knowledge is not stored as discrete, inspectable facts, but rather is distributed across a vast array of weights in a form that is not directly legible to humans. When such a system produces an answer, it cannot, in general, point to the source that justifies it, nor can its developers trace the output back to a specific passage in the training data. The model does not merely seek an answer; rather, it produces the most statistically plausible continuation. This is because a fluent, correct response and a fabricated one are, internally, produced by the same mechanism. Critics have memorably described these systems as 'stochastic parrots' that manipulate linguistic form without any grounded understanding of meaning.[43]

The training data further exacerbates the opacity. The corpora are of such substantial size that they are only imperfectly documented; they inevitably absorb the errors, biases and cultural skew of their sources, and they may contain copyrighted or simply mistaken material that the model will later reproduce with equal fluency. The extent to which a model has or has not been read is largely unknown to the end user.

From the perspective of the sports and swimming scientist, the takeaway message is one of concreteness rather than philosophy. When a model summarises four decades of Symposium proceedings or proposes a training intervention, its answer carries no visible chain of custody: no citation can be verified by inspection, and there is no guarantee that a plausible-sounding number was ever measured. This does not render the tool ineffective; rather, it stipulates the manner in which the tool must be utilised. The value of explainable AI methods, of maintaining human involvement, and of researchers' trained scepticism lies precisely in their capacity to provide external accountability, a capability that black box systems are unable to offer from within.

## AI at the Poolside: What Machine Learning Already Does in Swimming Science

It is recommended that the general considerations outlined above be taken into further consideration, as the concepts discussed have already been demonstrated to have concrete counterparts in the field of aquatic research. Machine learning has the potential to automate tasks that were previously time-consuming for scientists and coaches, such as the analysis of a single race or the management of a season's training load. Furthermore, machine learning can detect patterns that are imperceptible to the naked eye.

The initial focus should be on the analysis of technique. The conventional method of notational analysis of competition, which has historically been performed manually from poolside video, is being superseded by the advent of computer-vision pipelines. These pipelines are capable of detecting the swimmer, segmenting the four phases of the race (start, clean swimming, turn and finish) and computing split times, stroke counts and stroke rate automatically. Convolutional neural networks were demonstrated as early as 2017 to convert continuous video into the simple periodic signals from which stroke cycles can be read [44]. More recent systems utilise above- and below-water cameras to reconstruct a full lane and quantify parameters as specific as the knee angle of an underwater dolphin kick [32]. The most sophisticated of these techniques go beyond a simple bounding box drawn around the athlete to full markerless pose estimation, in which a network locates individual body landmarks frame by frame [45, 48, 49]. General-purpose tools such as DeepLabCut, originally devised to track animal movement from only a few hundred labelled frames, have made this approach broadly accessible [48, 49]. However, the water exacts a price: models trained on dry-land imagery cannot simply be reused but must be retrained almost from scratch on aquatic footage, because reflection, refraction, bubbles and splashes — at the surface It is imperative to note that, above all else, corruption of the visual cues on which detection and pose estimation depend is to be avoided under all circumstances, particularly in the context of underwater environments [46].

A parallel line of work exploits inertial measurement units — the accelerometers and gyroscopes now small enough to sit unobtrusively in a cap or on the lower back. Trained on the resulting signals, supervised and deep-learning models classify strokes, recognise the transitions between activities and estimate lap times from a single sensor with

accuracies above ninety per cent,[33, 47] offering an inexpensive, everyday alternative to the instrumented flume or the calibrated camera array.

Where enough data accumulate, the models begin to predict rather than merely describe. Statistical and machine-learning analysis of millions of official race results has exposed age- and stroke-dependent development curves and identified the ages at which training most influences later performance;[35] at the individual level, explainable models combining training history, anthropometry and dryland tests can predict a swimmer's competitive rating and, crucially, rank the variables responsible — telling a coach not only what the model predicts but why.[36] The aforementioned issue of concern, that is to say the 'black box' problem, is naturally addressed by such a predictable outcome.

The same techniques extend to health and durability. By integrating wearable-sensor workload with physiological monitoring, AI models help flag the accumulation of fatigue and the overuse patterns that precede injury — a matter of no small importance in a sport where the shoulder performs tens of thousands of repetitions each week.[30, 31] Personalised, data-driven load management is gradually replacing the traditional, one-size-fits-all training week.

A further possibility, which has emerged almost instantaneously, is that the scientist can now construct the analytical apparatus as readily as utilise it. With the assistance of agentic and 'vibe-coding' assistants, a researcher with only modest programming experience can create a bespoke web or mobile application in days rather than months. Such an application can ingest a training video and return pose-estimated joint angles for a specific movement, overlay kinematic feedback for a swimmer between repetitions, or allow a coach to log and visualise a whole squad's sessions from poolside on a tablet.

The same ease of construction extends to data. The advent of custom scripts and applications has enabled the automation of information retrieval across the internet, encompassing diverse data types such as competition results, historical rankings, meet databases and published datasets. These applications can now consolidate these disparate data sources into a unified, queryable resource. This development has significant implications for data analysis, as it allows for analyses that would previously have necessitated the dedicated expertise of a data team, including large-scale trend detection, cross-event comparison and the multivariable modelling of performance factors. It is evident that the domain, which was previously the exclusive domain of well-funded laboratories, is now becoming increasingly accessible to individual researchers who possess the requisite expertise to pose the pertinent questions.

Finally, and most crucially for practical application, the utility of these analyses is contingent on the ability of the coach to act on them. The contemporary paradigm, therefore, is not predicated on unsophisticated prediction, but rather on explicable, real-time feedback: multimodal systems that amalgamate video, wearable and physiological signals and present them in a format comprehensible to coaches, with the system having been designed in consultation with these professionals themselves [37]. At this juncture, the argument of this section closes its loop – the more capable the model, the more indispensable the human who interprets it, questions it and remains accountable for the subsequent decision.

## How AI Can Enhance the Sport Scientist's Productivity

From the perspective of the sport scientist, the value of these tools is not merely theoretical; it is quantifiable in terms of the hours that are reclaimed and the questions that are answered with greater efficiency. When employed judiciously, artificial intelligence has the potential to streamline the routine tasks associated with research, thereby enabling researchers to focus on tasks that are beyond the capabilities of machines, such as posing the appropriate questions and interpreting results within the complex framework of the sport.

The most immediate benefit is evident in the management of the literature. It is now possible for researchers to access and analyse decades of proceedings, such as those digitised for this Symposium, using retrieval-augmented systems. These systems locate relevant passages, summarise competing findings and surface the historical antecedents of a hypothesis in minutes rather than weeks. This process enables researchers to avoid the rediscovery of results that have already been known for forty years [30].

Within a laboratory setting, agentic and vibe-coding workflows empower scientists with only rudimentary programming knowledge to construct analysis pipelines, to purify and visualise kinematic or physiological datasets, to automate the extraction of stroke-cycle parameters from video[31], and to prototype statistical models — tasks that formerly necessitated the dedicated expertise of a programmer.

In the domain of communication, AI assistants have been shown to expedite the drafting, structuring, and language refinement of manuscripts, grant applications, and conference abstracts. They facilitate cross-language collaboration by producing high-quality translations, which is a clear advantage in a discipline as international as ours.

However, it is important to note that these gains are accompanied by a responsibility to maintain vigilance. It has been demonstrated that models are susceptible to hallucinations, misattribution of references, and the propagation of bias. Consequently, the responsibility for any claim, citation, or figure rests with the researcher and not the machine. It is therefore imperative that artificial intelligence is regarded as a diligent assistant, subject to human oversight and supervision. This approach ensures that it does not become an unquestioned authority, but rather a collaborative partner in the human-in-the-loop process.

## DISCUSSION

When Leon Lewillie inaugurated the first of these symposia in 1970, he lamented that, more than any other branch of biomechanics, swimming had been isolated from scientific investigation by the barrier of water. After a period of fifty-six years and fifteen editions, the barrier to entry has not disappeared, but has become more permeable. The medium that was once the source of frustration in measurement is now routinely instrumented, filmed above and below the surface, and simulated numerically. Consequently, what Lewillie could only describe in qualitative terms can today be quantified cycle by cycle. The ongoing existence of this Symposium, along with the archive that preserves it, facilitates the interpretation of that change as a single coherent trajectory rather than a series of disconnected advancements.

Methodologically, the promise of this moment lies in combination rather than in any single tool. Direct measurement and numerical simulation supply the causal account of why a swimmer moves as they do; data-driven models supply scale, speed and the ability to find patterns across thousands of cases. Neither replaces the other, and it is their integration — measurement to ground the models, models to extend the reach of measurement — that yields the strongest conclusions. One such example that has been utilised for a considerable period of time is the capacity to integrate hip speed data, as measured by a 'speedometer', with video recordings of the swimmer. The practice that was previously undertaken using analogue methods for a prolonged period has now been incorporated into a digital system that automatically generates a substantial amount of additional information.

The advent of the New AI-Driven World does not signify an opposition to the prevailing mechanistic understanding; rather, it serves as an augmentation of this comprehension. Machine learning has been demonstrated to accelerate the analysis of technique, the prediction of performance and the delivery of feedback [36, 37]. However, it does not provide an explanation for why a swimmer is fast per se. The optimal union of these two fields is thus the marriage of causal insight, as exemplified by the principles of fluid mechanics, and the pattern-finding power of data-driven models. The former provides a framework for determining the metrics to be measured and the underlying significance of these measurements. The latter facilitates the identification of the most relevant cases from among thousands of cases.

It is imperative that marriage be entered into with a state of mind that is receptive to the possibilities that may arise. The opacity of large models, the scarcity of reliable underwater data, the absence of standardised protocols and the modest size of the aquatic-science corpus all constrain what AI can responsibly deliver today,[30, 34, 38] and the temptation to accept a fluent answer without proof is strong. The utilisation of explicable methods, in conjunction with the maintenance of a human operator in a state of awareness, is not a matter of preference or optionality. Rather, these elements are indispensable prerequisites for the establishment of trust in the reliability of these tools.

In conclusion, the lesson of the archive is applicable in both directions. The digitisation and dissemination of our proceedings is of particular benefit to young researchers, who, in the absence of such resources, would be required to re-derive results that have been established over the course of several decades. Furthermore, these models, which are designed to learn from the content of our publications, are now able to draw upon a more extensive and up-to-date set of data. It is important to note that open repositories such as GitHub have become a valuable asset in this field, facilitating the sharing of code, training data and labelled images, thereby saving researchers time and effort. The quality of AI in this field in the future will depend, to a great extent, on the rigour and openness of what the community chooses to record today. It will also depend on our willingness to bring together and make consistent open-access datasets from different laboratories. Without these datasets, it is not possible to build robust models[38].

## CONCLUSION

The discipline of swimming science has evolved significantly from its original conception as a sport confined to the aquatic environment, as previously described by Lewillie. This field has now become quantitative in nature, with the mechanics of propulsion and the tools of artificial intelligence beginning to converge. The former provides causal understanding, while the latter facilitates scale and speed. The opportunity presented by this juncture is not to opt for one over the other, but rather to integrate them effectively.

The following three commitments are to be observed. Firstly, it is imperative that the energy-and-momentum view of propulsion, validated through the integration of computational fluid dynamics and flow measurement, should evolve into a standard lens rather than a specialist curiosity, as is the case with numerous other examples from our field of expertise. Secondly, the implementation of artificial intelligence should be guided by a set of principles that ensure its integration into existing systems in a manner that is both explainable and traceable. Furthermore, it is imperative to ensure that the application of artificial intelligence remains always subordinate to human judgement, thereby safeguarding the integrity and autonomy of human decision-making processes. Thirdly, it is imperative to cultivate a generation of researchers who possess the capacity to both interrogate a model and interpret a flow field or other fields within our scientific corpus. This can be achieved by integrating coding and machine-learning literacy within the sport-science curriculum[38]. Additionally, it is crucial to ensure that these researchers are well-versed in the historical underpinnings of these disciplines.

To that generation, many of whom will have had no opportunity to watch Mark Spitz swim for Counsilman or read Lewillie's first pages, the purpose of this lecture is clear: to learn the history of the field so as to see further than it has gone. The water continues to resist, the machine is powerful but blind, and the scientist, as in 1970, remains the individual who must ask the right question and stand behind the answer.

## ACKNOWLEDGEMENT AND FUNDING

The author wishes to express their gratitude to the organisers for their kind invitation to participate in the conference as a guest speaker at 'The León Lewillie Opening Lecture, with particular thanks to the president of the organising committee, Dr Flavio Antonio de Souza Castro, and the members of the Steering Group of the Biomechanics and Medicine in Swimming Congress. All my recent research work was supported by the project SWIM III MICIU/AEI/10.13039/501100011033 [PID2022-142147NB-I00] that you be reviewed at https://swim3.ugr.es/en/.

## REFERENCES

1. Lewillie L. Introduction. In: First International Symposium on Biomechanics and Swimming, Waterpolo and Diving; 1970 Sep 14-16; Brussels, Belgium. Brussels: Universite Libre de Bruxelles, Laboratoire de l'effort; 1970.
2. Counsilman JE. Foreword. In: First International Symposium on Biomechanics and Swimming, Waterpolo and Diving; 1970 Sep 14-16; Brussels, Belgium. Brussels: Universite Libre de Bruxelles, Laboratoire de l'effort; 1970.
3. Lewillie L. Research in swimming: historical and scientific aspects. In: Hollander AP, Huijing PA, de Groot G, editors. Biomechanics and Medicine in Swimming. Champaign, IL: Human Kinetics; 1983. p. 7-16.
4. Lewillie L. Quantitative comparison of the electromyogram of the swimmer. In: First International Symposium on Biomechanics in Swimming, Waterpolo and Diving. Brussels, Belgium: Universite Libre de Bruxelles; 1971. p. 155-159.
5. Clarys JP, Lewillie L. The description of wrist and shoulder motion of different waterpolo shots using a simple light-trace technique. In: First International Symposium on Biomechanics in Swimming, Waterpolo and Diving. Brussels, Belgium: Universite Libre de Bruxelles; 1971. p. 249-255.
6. Lewillie L. Welcome from the Symposium President. In: Swimming II. Baltimore: University Park Press; 1975. p. 3-4.
7. Clarys JP, Jiskoot J. Total resistance of selected body positions in the front crawl. In: Swimming II. Baltimore: University Park Press; 1975. p. 110-117.
8. Jiskoot J, Clarys JP. Body resistance on and under the water surface. In: Swimming II. Baltimore: University Park Press; 1975. p. 105-109.
9. Clarys JP. Human morphology and hydrodynamics. In: Swimming III. Baltimore: University Park Press; 1979. p. 3-41.
10. Clarys JP, Scafoglieri A, Tresignie J, Cattrysse E. Dubious use or misuse of scientific information in commerce and policy making, e.g. the swimsuit case. In: XIIth International Symposium for Biomechanics and Medicine in Swimming. Canberra, Australia; 2014. p. 100-106.
11. Clarys JP. The Brussels swimming EMG project. In: Swimming Science V. Champaign, IL: Human Kinetics; 1988. p. 157-172.
12. Piette G, Clarys JP. Telemetric EMG of the front crawl movement. In: Swimming III. Baltimore: University Park Press; 1979. p. 153-159.

13. Clarys JP. A review of EMG in swimming: explanation of facts and/or feedback information. In: Biomechanics and Medicine in Swimming. Champaign, IL: Human Kinetics; 1983. p. 123-135.
14. Kemper HCG, Verschuur R, Clarys JP, Jiskoot J. Total efficiency and swimming drag in swimming the front crawl. In: Biomechanics and Medicine in Swimming. Champaign, IL: Human Kinetics; 1983. p. 199-206.
15. Olbrecht J, Clarys JP. EMG of specific strength training exercises for the front crawl. In: Biomechanics and Medicine in Swimming. Champaign, IL: Human Kinetics; 1983. p. 136-141.
16. Bollens E, Annemans L, Vaes W, Clarys JP. Peripheral EMG comparison between fully tethered and free front crawl swimming. In: Swimming Science V. Champaign, IL: Human Kinetics; 1988. p. 173-181.
17. Cabri JMH, Annemans L, Clarys JP, Bollens E, Publie J. The relation of stroke frequency, force and EMG in front crawl tethered swimming. In: Swimming Science V. Champaign, IL: Human Kinetics; 1988. p. 183-189.
18. Clarys JP, Toussaint HM, Bollens E, Vaes W, Huijing PA, de Groot G, et al. Muscular specificity and intensity in swimming against a mechanical resistance — surface EMG in MAD and free swimming. In: Swimming Science V. Champaign, IL: Human Kinetics; 1988. p. 191-199.
19. Clarys JP. The historical perspective of swimming science. Foreword. In: Biomechanics and Medicine in Swimming VII. London, United Kingdom; 1996. p. XI-XXXIV.
20. Arellano R, Pardillo S. An evaluation of changes in the crawl-stroke technique during training periods in a swimming season. In: MacLaren D, Reilly T, Lees A, editors. Swimming Science VI. Liverpool, United Kingdom; 1992. p. 143-149.
21. Arellano R, Moreno FJ, Martínez M, Oña A. A device for quantitative measurement of starting time in swimming. In: Troup JP, Hollander AP, Strass D, Trappe SW, Cappaert JM, Trappe TA, editors. Biomechanics and Medicine in Swimming VII. Atlanta, United States; 1996. p. 195-200.
22. Arellano R, Gavilán A, García F. A comparison of the underwater undulatory swimming technique in two different body positions. In: Keskinen KL, Komi PV, Hollander AP, editors. Biomechanics and Medicine in Swimming VIII. Jyväskylä, Finland; 1999. p. 25-28.
23. Arellano R, Pardillo S, Gavilán A. Usefulness of the Strouhal number in evaluating human underwater undulatory swimming. In: Chatard JC, editor. Biomechanics and Medicine in Swimming IX. Saint-Étienne, France; 2002. p. 33-38.
24. Arellano R, Terrés-Nicoli JM, Redondo JM. Fundamental hydrodynamics of swimming propulsion (invited contribution). In: Vilas-Boas JP, Alves F, Marques A, editors. Biomechanics and Medicine in Swimming X. Porto, Portugal; 2006. p. 15-20.
25. de la Fuente B, Arellano R. Effect of specific training on swimming start performance. In: Kjendlie PL, Stallman RK, Cabri J, editors. Biomechanics and Medicine in Swimming XI. Oslo, Norway; 2010. p. 249-251.
26. Morales E, Arellano R, Femia P, Mercadé E. Regression analysis model applied to age-group swimmers: study of stroke rate, stroke length and stroke index. In: Kjendlie PL, Stallman RK, Cabri J, editors. Biomechanics and Medicine in Swimming XI. Oslo, Norway; 2010. p. 127-129.
27. Cadenas C, Arellano R, López-Contreras G, Taladriz S. Spatiotemporal characteristics of walking on land and in water: reliability of measurements and analysis of differences. In: Mason B, editor. Biomechanics and Medicine in Swimming XII. Canberra, Australia; 2014. p. 543-546.
28. Arellano R, Ruiz-Navarro JJ. Underwater undulatory swimming: linking kinematics, wave analysis and flow visualization. In: Staub I, editor. Biomechanics and Medicine in Swimming XIV. Leipzig, Germany; 2023.
29. Cust EE, Sweeting AJ, Ball K, Robertson S. Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance. J Sports Sci. 2019;37(5):568-600. doi:10.1080/02640414.2018.1521769
30. Kumar R, Bogia P, Haq AU, Singh V, Reddy TO. Leveraging artificial intelligence and machine learning in sport sciences: a systematic literature review of applications, outcomes, and future directions. Sport Sci Health. 2025;21(4):2429-2443. doi:10.1007/s11332-025-01482-y
31. Musat CL, Mereuta C, Nechita A, Tutunaru D, Voipan AE, Voipan D, et al. Diagnostic applications of AI in sports: a comprehensive review of injury risk prediction methods. Diagnostics. 2024;14(22):2516. doi:10.3390/diagnostics14222516
32. Chern YR, Chen YH, Lin FS, Lin HC, Chen GT, Chu CP, et al. A butterfly stroke swimming recording and performance analysis system based on computer vision and machine learning. Measurement. 2025;251:117171. doi:10.1016/j.measurement.2025.117171
33. Delhaye E, Bouvet A, Nicolas G, Vilas-Boas JP, Bideau B, Bideau N. Automatic swimming activity recognition and lap time assessment based on a single IMU: a deep learning approach. Sensors. 2022;22(15):5786. doi:10.3390/s22155786
34. Reis FJJ, Alaiti RK, Vallio CS, Hespanhol L. Artificial intelligence and machine-learning approaches in sports: concepts, applications, challenges, and future perspectives. Braz J Phys Ther. 2024;28(3):101083. doi:10.1016/j.bjpt.2024.101083
35. Xie J, et al. Machine learning of swimming data via wisdom of crowd and regression analysis. Math Biosci Eng. 2017;14(2):511-527. doi:10.3934/mbe.2017031
36. Carvalho DD, Goethel MF, Silva AJ, Vilas-Boas JP, Pyne DB, Fernandes RJ. Swimming performance interpreted through explainable artificial intelligence (XAI) — practical tests and training variables modelling. Appl Sci. 2024;14(12):5218. doi:10.3390/app14125218

37. Camilleri V, Camilleri RY, Fialovszky M, Pace D, Seychell D, Montebello M. Towards explainable multimodal sensing for swimming analysis: early findings from the SWIM-360 project. Sensors. 2025;25(22):7047. doi:10.3390/s25227047
38. Richter C, O'Reilly M, Delahunt E. Machine learning in sports science: challenges and opportunities. Sports Biomech. 2024;23(8):961-967. doi:10.1080/14763141.2021.1910334
39. Lendvai GF. ChatGPT in academic writing: a scientometric analysis of literature published between 2022 and 2023. J Empir Res Hum Res Ethics. 2025. doi:10.1177/15562646251350203
40. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. In: Advances in Neural Information Processing Systems 30. 2017. p. 5998-6008. doi:10.48550/arXiv.1706.03762
41. Bommasani R, et al. On the opportunities and risks of foundation models. arXiv:2108.07258. 2021. doi:10.48550/arXiv.2108.07258
42. Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. arXiv:1503.02531. 2015. doi:10.48550/arXiv.1503.02531
43. Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. 2021. p. 610-623. doi:10.1145/3442188.3445922
44. Victor B, He Z, Morgan S, Miniutti D. Continuous video to simple signals for swimming stroke detection with convolutional neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017. doi:10.48550/arXiv.1705.09894
45. Mathis A, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci. 2018;21(9):1281-1289. doi:10.1038/s41593-018-0209-y
46. Giulietti N, Caputo A, Chiariotti P, Castellini P. SwimmerNET: underwater 2D swimmer pose estimation exploiting fully convolutional neural networks. Sensors. 2023;23(4):2364. doi:10.3390/s23042364
47. Aram S, Thompson IM, Hudson DA, Banks J. Investigation of supervised machine learning for IMU-based swimming activity recognition. In: ISBS Proceedings Archive. 2025;43(1). Available from: https://commons.nmu.edu/isbs/vol43/iss1/7
48. Ruiz-Navarro JJ, Cano-Adamuz M, Andersen JT, Cuenca-Fernández F, López-Contreras G, Vanrenterghem J, et al. Understanding the effects of training on underwater undulatory swimming performance and kinematics. Sports Biomech. 2024;23(6):772-787. doi:10.1080/14763141.2021.1891276
49. Ruiz-Navarro JJ, López-Belmonte Ó, Cuenca-Fernández F, Gay A, Arellano R. The effects of eccentric training on undulatory underwater swimming performance and kinematics in competitive swimmers. J Hum Kinet. 2024;93:53-68. doi:10.5114/jhk/175824

## A Glossary of Key AI and Large-Language-Model Terms

*I. Foundations: Language Models and Their Building Blocks*

**Prompt —** the natural-language instruction or query supplied to a model to elicit a desired response; the primary interface through which a user directs model behaviour.

**Token —** the elementary unit of text — a character sequence, word fragment or word — into which input is segmented and upon which a model computes.

**Context window —** the maximum span of tokens a model can attend to at once, bounding how much prior text it may consider when generating a response.

**Large Language Model (LLM) —** a neural network trained on vast text corpora to predict successive tokens, thereby acquiring broad linguistic and reasoning capabilities.

**Inference —** the stage at which a trained model is run on new input to produce outputs, as distinct from the training stage.

**Temperature —** a sampling parameter controlling the randomness of a model's output; lower values yield more deterministic, higher values more varied text.

**Multimodal —** descriptive of a model able to process and integrate more than one type of data, such as text, images and audio.

*II. Directing and Grounding Models: Prompting and Reliability*

**Prompt engineering —** the systematic practice of designing and refining prompts to obtain accurate, reliable and reproducible outputs from a model.

**Chain-of-thought —** a prompting approach that elicits explicit intermediate reasoning steps, improving performance on complex tasks.

**Hallucination —** the generation of plausible-sounding yet factually incorrect or fabricated content by a model.

**Guardrails —** the constraints and safety mechanisms that restrict a model's outputs to acceptable, safe and policy-compliant behaviour.

**Retrieval-Augmented Generation (RAG) —** a technique that grounds a model's output in external documents retrieved at inference time, improving factual accuracy.

**Fine-tuning —** the further training of a pre-trained model on a narrower dataset to specialise its behaviour for a particular task or domain.

**Benchmark —** a standardised task or dataset used to measure and compare model performance objectively.

*III. From Models to Agents: Autonomy, Tools and Oversight*

**Agent —** an autonomous software system that uses a model to plan, decide and act — invoking tools and executing multi-step tasks with limited human intervention.

**Loop —** the iterative cycle in which an agent repeatedly observes, reasons, acts and evaluates until a task is completed or a stopping condition is met.

**Harness —** the surrounding software scaffolding that connects a model to tools, memory and control logic, enabling it to operate as a functioning agent.

**Model Context Protocol (MCP) —** an open standard that connects AI models to external tools and data sources through a uniform interface, extending an agent's reach beyond its training data.

**Orchestration —** the coordination of multiple models, agents and tools into a single coherent workflow directed toward a common objective.

**Human-in-the-loop —** a design pattern that preserves human review and approval at key decision points, keeping automated systems accountable and correctable.

**Claude Code —** an agentic command-line coding tool developed by Anthropic that enables a Claude model to read, write and execute code autonomously within a developer's environment.

*IV. AI-Assisted Software Development: Practices and Workflows*

**AI tooling —** the ecosystem of software instruments — editors, assistants, orchestration frameworks and evaluation suites — that support the development and deployment of AI-based applications.

**Framework —** a reusable, structured software foundation offering conventions and pre-built components upon which applications, and increasingly AI agents, are built.

**Task-specialised model —** a model selected or configured for a particular class of task, such as planning or code generation, reflecting the practice of matching model capabilities to problem type.

**Vibe coding —** an informal, exploratory style of software creation in which the developer steers a model through natural language, accepting its suggestions with minimal manual review.

**Structured vibe coding —** a disciplined variant of vibe coding in which the developer retains explicit control through defined specifications, review checkpoints and testing, thereby tempering the model's autonomy with human oversight.

**Vibe coding cleanup —** the corrective phase in which a developer reviews, debugs and refactors AI-generated code, resolving the errors and inconsistencies that unsupervised generation tends to introduce.

**Agentic coding —** a programming paradigm in which an AI agent autonomously writes, tests and revises code across successive iterations toward a stated goal.

**Smoke test —** a quick, preliminary check that verifies whether the most basic functions of a system work before more detailed testing is undertaken.

**System-level prompt engineering —** advanced prompt design aimed at orchestrating complete systems rather than isolated responses, coordinating multiple prompts, tools and agents toward an integrated goal.

*V. Specification-Driven Engineering: Building Real Products*

**Spec-Driven Design —** the discipline of producing highly detailed and complete specifications — requirements, interfaces and acceptance criteria — before implementation, so that documented design intent, rather than ad-hoc code, governs the build.

**Spec-Driven Development (SDD) —** a methodology in which a precise written specification is authored first and thereafter serves as the executable source of truth from which an AI agent generates, tests and maintains the working code of a real project.

**Spec Kit —** an open toolkit that operationalises spec-driven development by transforming structured specifications into the prompts and scaffolding that guide AI coding agents.

**OpenSpec —** an open specification framework for defining and sharing machine-readable requirements that AI agents can consume as authoritative build instructions.

**Product-oriented architecture —** a modern software architecture designed around real, deployable products — favouring modularity, maintainability and scalability — rather than experimental prototypes.