Report of INESC TEC Ad-Hoc Review Committee on Artificial Intelligence and Robotics

(July 31, 2023)

Table of Contents

1. Introduction

The INESC TEC Ad-Hoc Review Committee on Artificial Intelligence and Robotics (ARCAR) is composed of:

Dr. José Fortes (Chair)	University of Florida, USA.
Dr. Ulises Cortés	Universitat Politècnica de Catalunya, Spain.
Dr. Avi Kak	Purdue University, USA.
Dr. John Leonard	Massachusetts Institute of Technology, USA.
Dr. Petia Radeva	University of Barcelona, Spain.

At the request of INESC TEC, the ARCAR conducted a visit on July 10 & 11, 2023 to INESC TEC Porto, Portugal, to review INESC TEC's domains of Artificial Intelligence and Robotics. During the first day, single-session presentations were made on both domains, on trans-domain (also known as transversal) projects and on INESC TEC as an organization. During the second day of the visit, the ARCAR visited INESC TEC's robotics and bioengineering laboratories and presented its preliminary findings to the INESC TEC Board of Directors and INESC TEC researchers. Dr. Masaru Kitsuregawa from the National Institute of Informatics, Japan, also attended some of the sessions remotely. Dr. Kitsuregawa is a member of the external Scientific Advisory Board of INESC TEC. This report summarizes the ARCAR's findings and recommendations from the said visit.

INESC TEC provided to the ARCAR the following documents prior to the visit:

- Slides of presentations
- INESC TEC Strategic plan 2023-2030
- Gender Equality Plan 2022-2026
- ARCAR Meeting agenda

The ARCAR members held several meetings among themselves, contributed thoughts and text, and reviewed drafts of this report. The final contents of this report have been reviewed and are agreed to by all the ARCAR members who participated in the visit.

This report is divided into three major parts:

- Cross-cutting findings and recommendations.
- Answers to questions from the INESC TEC Board of Directors.
- Findings and recommendations on the Artificial Intelligence domain.
- Findings and recommendations on the Robotics domain.

2. Cross-cutting Findings and Recommendations

2.1 Key Findings

2.1.1 Summary Assessment

The ARCAR was impressed by the excellence of the institutional vision, research, human resources, and infrastructure of INESC TEC. Noteworthy findings include the following:

- The best work being done at INESC TEC is comparable to the best work being carried out anywhere in the world.
- The scope of the research, competencies and experimental facilities is exceptionally broad, enabling interdisciplinary work with practical impact.
- INESC TEC researchers include accomplished senior researchers and outstanding young researchers, which bodes well for sustained long term research excellence.

2.2 Key Recommendations

2.2.1 Benchmarking INESC TEC Performance

The ARCAR recommends INESC TEC to identify and report metrics and associated benchmarks being used to assess performance in relation to internal goals and similar institutions in Portugal and elsewhere. This should be done at the institution and domain levels. For example, the Robotics lab is on a par with many of the best Robotics labs in the world and is likely to be one of very few Robotics labs in Portugal at its level. Qualifying and quantifying these relative comparisons are very effective ways to communicate excellence and assess scientific performance.

2.2.2 Highlighting Research Excellence and Accomplishments

The ARCAR members found it somewhat difficult to calibrate the quality and outcomes of research projects from the presentations. It would have helped to have explicit mentions of publications in highly reputed journals or venues, best paper awards, research prizes, patents, unique prototypes, broadly adopted tools and services, etc. It is also desirable to have numbers of publications and other scientific metrics (overall as well normalized per member) to assess the excellence of the INESC TEC members.

2.2.3 Evidence of Institutional Reviews of Ethical Aspects of INESC TEC Activities

An impressive aspect of INESC TEC is its large portfolio of interdisciplinary projects with social impact. This creates a need for INESC TEC to have policies on ethics and processes to review the compliance of projects with those policies. These policies and processes may already be in place (as revealed at https://repositorio.INESC TEC.pt/server/api/core/bitstreams/b0243350-09a9-4009-9b6d-ba50a7a5578b/content), but they were invisible to the ARCAR committee. When concerns arise, such as “Should INESC TEC conduct work for military agencies?” or “How does INESC TEC work impact the environment and how can any negative impact be mitigated?”, INESC TEC should have answers based on its ethics review processes. One or two slides should be

dedicated to ethics in the institution-level presentation, also including a slide in each scientific domain presentation when specific issues are applicable (e.g., ethical issues in data use in Artificial Intelligence research).

2.2.4 Young Researchers as "Products" of the INESC TEC's Scientific Environment

The ARCAR committee was very impressed with the PhD students and other young researchers who gave presentations at sessions and lab tours. Most of them are on par with the best PhD students and postdocs in the world. Their communication skills, technical knowledge, composure, and ability to think on their feet and answer scientific questions are evidence of the sound scientific environment and practices at INESC TEC. This fact should be made explicit and highlighted in future reviews by introducing the grad students and the postdocs to the reviewers and by including them in social events where the reviewers are present.

2.2.5 Integrate Science into Every INESC TEC Activity

INESC TEC's projects that aim at high-level TRL's create opportunities for asking scientific questions whose answers can lead to results and technologies useful for future projects. However, these opportunities should be sought from the beginning of the projects, dedicating people and other resources to such purpose. For example, a PhD student could be tasked with exploring research questions that arise from scaling up, optimizing, generalizing, or otherwise improving the technologies and products generated by the project being started. This "born-scientific" approach would increase the likelihood of science being produced when compared with retrospective efforts to dig out scientific results at the end of the project.

3. Answers to the Board of Directors' Questions

3.1 Quality, merit, relevance and extent of internationalization of the R&D activity, assessed by international standards, considering originality, consistency and rigour, as well as the relevance of the results, its valorization and dissemination, the participation in advanced training and the development and consolidation of scientific careers, as well as the contribution for the promotion of scientific and technological culture.

- The committee was very impressed with the fact that the best of what was presented by INESC TEC is comparable to the best anywhere in the world.
- As a case in point, the research by Prof. Cunha is stellar. His research is stellar in that combining deep learning techniques with ingenious sensing methods allows for critical discrimination that was impossible earlier.
- We would also mention that the CRIS and CRAS laboratories that the SAB visited on 18 of July are truly the crown jewels of INESC TEC.
- However, we saw considerable room for improvement in several aspects of what was presented to the committee.
- The AI domain should employ better metrics and bibliometric tools to compare its scientific merits at national, European and international levels.

- The program also needs to become more selective in choosing the scientific criteria for assessing the level of research engagement by its members and the criteria for evaluating the members.

- While the teams do have some papers in leading conferences and journals, the quality of the publication venue choices is not uniform. In particular, the committee wished to express caution about having too many publications in venues that are considered predatory, including some of the MDPI journal. (As an example, see the paper "M Angeles Oviedo-Garca, "Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI)", Research Evaluation, Volume 30, Issue 3, July 2021, Pages 405419a, https://doi.org/10.1093/reseval/rvab020 ").
- The presentation was found wanting for a description of the activities being pursued for achieving a greater gender balance in the program.

3.2 What is the scientific and technological merit of the research team, especially those with PhD, evidence of international and national recognition and, if applicable to the nature of the R&D activities or to objectives of impact in society, also the technical, cultural or artistic merit available in the team?

- The team certainly possesses scientific and technological merits of their research.
- However, it should improve their presentation, including a comparison with other groups and centres.
- The presentation skipped information related to group awards, organization of international scientific events, and other dissemination and communication activities to measure the impact on society and industry.
- It is highly recommendable to intensify R&D activities that will improve the scientific visibility of INESC TEC like organizing international conferences, workshops and challenges in top-rank conferences as well as organizing special issues in high impact scientific journals. INESC TEC should propose a plan to establish itself as a reference institution at international level and identify which are the strategic lines where it wants to achieve such visibility. Special attention should be paid to high quality scientific programs like ERC and others.

3.3 What are the quality, merit, and relevance of objectives, strategy, plan of activities and organization for the next five years, including contracting new researchers and consolidating scientific careers, with the associated institutional co-responsibility.?

- Answering this question is made difficult by the fact that much of what was presented focused on past research.
- Four of the six projects reviewed in the AI section, had either already terminated or were about to end.
- Nonetheless, assuming that the activities over the next five years would be a continuation of what was presented, and assuming that the program will endeavour to bring up the level of excellence in a larger number of projects to that of the best exemplars of what we saw, we believe the future bodes well for the program.

- Still, it is recommended for the next SAB meeting to explicitly present such a strategic plan and mission.
- Also, it is recommendable to specify the politics of members' acceptance and membership to INESC TEC. Is there any condition to be a member? What is the evaluation process of the members?!

4. Feedback on Individual Scientific Domains

4.1 Artificial Intelligence (AI)

4.1.1 Vision

4.1.1.1 Summary assessment

Starting with the origin of modern AI as a part of computer science, the Vision statement quickly lists the different application areas that have benefited from the AI approach to solving problems. The Vision statement also highlights the importance of the more traditional symbolic logic-based approaches. This is important for the reasons mentioned in the following subsection. The statement mentions cognitive layers that could sit on top of the neural layers for solving complex problems. Most importantly, the Vision statement discusses the ethical, legal, socioeconomic, and cultural dimensions of AI-based solutions to problems.

4.1.1.2 Feedback on strengths

- Its exceptional strength is the comprehensiveness and quality of writing the Vision statement. It conveys a lot of information without using too many words.

- The fact that the Vision statement highlights the importance of the traditional symbolic-logic based methods is a good thing since, in our opinion, over the long haul, it will be a combination of the traditional symbolic-logic based approaches and the approaches based on the more modern deep-learning based methods that will solve the genuinely complex problems of the future.

- By saying that "*AI solutions must be ethical by design*", the Vision statement touches on an essential reality that underlies modern AI: since all deep-learning-based solutions are data-driven, any *biases* in the construction of the training datasets will be reflected in the solutions provided by the AI based solutions. To the extent such solutions come into widespread use, they could result in subjecting people and/or systems to harm.

4.1.1.3 Feedback on weaknesses

- The strength of the comments in the Vision statement about ethics and trustworthiness turned out to be inversely proportional to how this important subject was dealt with during the presentations made to the committee.

- It was also not clear what exactly was meant by "AI solutions and deployment must be ethical by design." Are we talking about the architectures of the AI

frameworks, or are we talking about how the datasets are constructed for training the networks? Since, in general, ethics deals with our values, our morals, how we relate to others, how we relate to the environment, etc., and since, in general, again, these various dimensions of ethics have no fixed and quantitative answers, it would be impossible to address such issues at the architectural level.

- The Vision statement was so high-level that it would be difficult to derive from it the focus needed at the more nuts-and-bolts level for guiding what is essentially an engineering program of research and development. This shortcoming of the Vision statement could be addressed by extending it a bit and either stating or providing forward pointers to a set of futuristic but more concrete thoughts about the anticipated direction of the AI program.

4.1.2 Challenges

4.1.2.1 Summary assessment

Five Research Challenges (RCs) were presented:

- Build highly valuable and reusable AI resources.
- Exploit models and algorithms for advanced tasks.
- Produce AI models that humans can inspect, understand, learn from, and, at the same time, contribute to.
- Learn models and deploy AI efficiently.
- Enhance perception in dynamic, noisy, and multi-modal environments.

4.1.2.2 Feedback on strengths

- Aligned with their EU projects and with EU strategy.
- We concur with the importance of exploiting the models and algorithms for advanced tasks.
- The importance of why it must be possible to be able to inspect and understand the AI models cannot be overstated.
- Equally compelling is the need to make more efficient how AI applications are deployed considering their excessive data need.

4.1.2.3 Feedback on weaknesses

- Simply stating a challenge is not that challenging. More problematic is delineating a roadmap that gets you to the goalpost. The committee would have very much liked to have seen such roadmaps for all the listed challenges.

- Generally, the more generic a challenge sounds, the less likely that its pursuit would result in anything tangible. Consider the challenge "*Exploit models and algorithms for advanced tasks*." One could ask: Advanced in relation to what? What is the benchmark for judging the degree of advancement of a task? Is the goal to pursue tasks that are more advanced compared with those that have been

addressed in the past in INESC TEC? or is the goal to pursue tasks that are at a higher level of complexity vis-à-vis those being pursued in other labs around the world?

4.1.3 Competencies

4.1.3.1 Summary assessment

The presentation reports sixteen fields of competencies that span many core and applied ideas.

4.1.3.2 Feedback on strengths

- It is important to appreciate the fact that, from the standpoint of AI research, the listed competencies are inter-connected and complementary.
- Going forward, the best AI research is likely to be produced by groups that, through the expertise of the individuals involved, possess deep insights into almost all those competencies. The listed competencies should therefore be a source of great strength for the AI research program.
- The most fundamental problems in AI are at the intersection of several of the topics in the list. To the extent these competencies continue to evolve within INESC TEC, there would be an ever-increasing likelihood that the contributions made by the researchers would address the foundational notions in both core AI and applied AI.

4.1.3.3 Feedback on weaknesses

- The competencies that are listed represent a primarily algorithmic perspective on AI. However, frequently, there is more to AI research than just writing high-level Python code for solving a problem.
- The list does not appear to be cognizant of the fact that AI would not be where it is today had it not been for the advances in software engineering, on the one hand, and hardware (notably GPUs), on the other. It is the constraint-free availability and easy-to-use programmability of platforms like PyTorch and Tensorflow that have played huge roles in AI spreading like wildfire through universities and corporate research labs.
- The most fundamental advances in AI often require extending the PyTorch and/or the TensorFlow platforms. Since these platforms consist of Python wrappers around the actual implementation code that is in C++, any laboratory that wishes to develop concepts in core AI must also possess some expertise in how to extend the AI programming platforms, including how to add new code to the C++ based substrate of the platforms.

4.1.4 Portfolio

4.1.4.1 Summary assessment

The Portfolio includes Projects, Publications, PhD dissertations, Patents, and Awards.

4.1.4.2 Feedback on strengths

- The projects all sound interesting and worthy of research.
- Publications: there are several in sci-indexed journals and international conferences and that is good.
- Patents: two patents achieved.
- International collaborations: Both the European and those with the institutions in the US and the rest of the world are a source of strength.
- PhD thesis: enough.
- Awards: good number but should have been emphasized during the presentations.

4.1.4.3 Feedback on weaknesses

- The quality of the venues chosen for disseminating research is not uniformly high (e.g., using MDPI venues is not recommended)

- It is also not clear as to what extent the researchers follow what has emerged as a common practice for disseminating new knowledge in the leading AI labs in the world. The three steps of this common practice are:

    - When a project has reached the publication stage, you post what would be your submission to a leading conference or journal at the arxiv.org website and make the code available on GitHub. This constitutes the first public disclosure of the new work and also allows to make claims about the priority of the work over similar ideas that others may be publishing at around the same time. This step is necessary because the reviewers at leading venues give reduced credibility to contributions whose implementations are private. A top-notch reviewer is likely to want to run the code to believe the story.

    - At the same time, one submits the paper to a leading conference or a journal. Leading venues do not hold it against the authors if they have posted their submissions at a preprint service like arxiv.org.

    - Should the paper be accepted, the researcher can mention that fact at arxiv.org if you so wish. Additionally, it is highly likely that what is finally accepted does not have all the important details about the project (since all top-notch venues strictly enforce page limits on what they accept), one can have two different versions of the work in the public domain: a somewhat terse presentation at the conference or the journal and a more detailed presentation at arxiv.org. Obviously, should the paper be rejected, the researcher can submit the work to a different venue, or can make do with the public disclosure of the paper at arxiv.org. In either case, s/he should update the paper at arxiv.org taking into account the feedback you received from the reviewers.

- For the projects that have expired or are about to expire (four out of six), the presentation should have focused more on what was achieved and the extent to which the original goals of the projects were met.

- Stress on the scientific achievements within the projects

4.1.5 SWOT

4.1.5.1 Summary assessment

SWOT summarizes separately the strengths of the AI part of the overall program, its perceived weaknesses, the opportunities that lie ahead, and the potential threats to the achievement of the long-term goals.

4.1.5.2 Feedback on strengths

A compact summary of the four important dimensions of AI research in INESC TEC. All the mentioned items are relevant and important in each of the four listings.

4.1.5.3 Feedback on weaknesses

At least one of the weaknesses listed under "W" may be a strength of the AI program at INESC TEC, specifically the item "*The AI domain is spread over many centers*." Cross-disciplinary research is foundational to many key advances in both core AI and applied AI. Often the solutions developed for one application inspire new solutions in other seemingly unrelated applications.

4.1.6 Recommendations

Since, overall, our assessment of the AI program was very positive, our best recommendation would be to stay the course while you are raising the standards to which the individual researchers are held regarding the dissemination of the new knowledge produced by them. As mentioned elsewhere in this report, the best of what we saw compares well with the best research projects anywhere in the world. Engineering research, by its very definition, is founded on laboratories. The more excellent the capabilities provided by a laboratory, the greater the possibility that the research coming out of that lab will be world-class.

We were very impressed by the laboratories we visited. Those laboratories can be the means to produce work of unparalleled quality by INESC TEC researchers.

4.2 Robotics (ROB)

4.2.1 Vision

4.2.1.1 Summary assessment

See below Feedback on strengths and weaknesses.

4.2.1.2 Feedback on strengths

The vision segment for Robotics identified key topics that provide ample opportunities for new technology development and societal impact.

4.2.1.3 Feedback on weaknesses

The actual vision statement reads: "*Robotics became more intelligent, autonomous, and useful in a wide area of applications. This new paradigm poses new challenges and problems to be solved that require new scientific approaches.*" Is it possible to revise this to be a succinct "vision statement" sentence that can be a sort of "guide star" for the team's efforts? For example, consider: "*Creating robust, long-lived autonomous systems that can safely interact with the physical world to amplify human capabilities in challenging real-world environments"* or perhaps with something to say why, for example for the Health of the Planet.

The team can do a better job in articulating, from a "robotics science" perspective, what is most unique in their vision for the efforts of INESC TEC in comparison to what is happening in the world in competing robotics labs. Where are the specific opportunities where INESC TEC is uniquely poised to achieve breakthroughs in the scientific disciplines that underlie robotics, while also helping society?

4.2.2 Challenges

4.2.2.1 Summary assessment:

See below Feedback on strengths and weaknesses.

4.2.2.2 Feedback on strengths:

The four articulated main challenges are: (1) Autonomy, (2) Physical Interaction, (3) Human-Robot Collaboration, and (4) Resilient Robotic Systems – this is a strong list, that compares well with what one would expect to see in a leading robotics research laboratory. These challenges provide ample opportunities to provide real-world impacts in science and technology, and in producing future leaders in robotics for Portugal and beyond.

4.2.2.3 Feedback on weaknesses:

The team can better articulate the ways in which it seeks to contribute to fundamental disciplines of robotics (such as perception, task and motion planning, navigation, control, manipulation, and human-robot interaction). How is the team seeking to make fundamental advances in these areas that can benefit and impact the research progress of other academic researchers worldwide? This might come via not only new algorithms, but also potentially unique data sets, software, and simulation tools.

4.2.3 Competencies

4.2.3.1 Summary assessment

See below Feedback on strengths and weaknesses.

4.2.3.2 Feedback on strengths

Both CRAS and CRIS have strong capabilities in the "systems" aspects of robotics. The technical expertise required to design, build, and operate the systems that we have seen from both labs is immense. Both teams deserve strong praise for their field impressive robotics capabilities. This involves not just creating innovative hardware, but also the underlying software and data processing.

4.2.3.3 Feedback on weaknesses

The lab tours and discussions with students and staff in the labs on day 2 of the review were more compelling than the PowerPoint presentations on day 1. We understand that the team was directed to not include videos that showed their systems operating in real-world environments, but doing so might have been helpful to convey the complexity and difficulty of the challenges being considered. However, critically these deployments need to be connected to first-tier publications using the data from these systems – which aspects of the team's core competencies are most novel?

It has been said that "he or she who has more than three priorities has none". The list of competencies is long – in which areas do you seek to truly advance the state-of-the-art and have a worldwide impact, vs. in which areas do you just need to operate at the pre-existing state-of-the-art? Trying to push for scientific advances in so many areas at once might be too diffuse vs. a more focused approach on key areas (such as "navigation and mapping in complex environments", "resilient, sustainable software architectures for lifelong autonomy", "physical interaction in complex environments", …).

4.2.4 Portfolio

4.2.4.1 Summary assessment:

See below Feedback on strengths and weaknesses.

4.2.4.2 Feedback on strengths:

The breadth of the robotics portfolio is impressive, and competitive with efforts at leading institutions worldwide, such as CMU's Field Robotics Center, the Australian Field Robotics Center, and the Woods Hole Oceanographic Institution.

There are many highlights, including the operation in the flooded mine, the deployment of agricultural robots in unique settings, and the long-term underwater autonomy.

The Transversal projects were engaging and show great potential for societal impacts.

4.2.4.3 Feedback on weaknesses:

The team should better showcase selected projects that achieved recognition via publications and awards from the leading conferences and journals in the field.

4.2.5 Projects

4.2.5.1 Summary assessment

See below for Feedback on strengths and weaknesses.

4.2.5.2 Feedback on strengths

The team has clearly been successful in securing major funding to pursue a wide range of important and innovative activities, while educating future leaders in critical technologies that can help Portugal and the world. The range of topics being explored is impressive, and nearly every topic raised has merit for contributing to scientific progress in robotics. The team is working on some "grand challenge problems" that can guide such progress (but can articulate those grand challenges better.)

4.2.5.3 Feedback on weaknesses

The teams can do a better job in "telling its story", bridging the gap between the systems efforts and underlying publishable contributions to robotics. The team can better establish how its efforts compare worldwide by stressing which contributions have achieved publication at the leading peer-reviewed international robotics venues; these include IEEE ICRA, RSS, IEEE RAL, IEEE TRO, IJRR, JFR, IEEE JOE, and IEEE IROS. The team should also consider co-organizing workshops at ICRA, RSS and/or IROS on their specialty topics, to help connect with other leading efforts in these areas and build the brand of INESC TEC on the world stage in academic robotics.

The team should also consider creating short inspirational videos that can be shared with young people, peer institutions, and international media, to better get the word out about the outstanding efforts underway – especially for the grand challenge problems. The team should also consider if there are opportunities to excite more women and more people from underrepresented backgrounds to pursue graduate study in Robotics with INESC TEC.

4.2.6 SWOT

4.2.6.1 Summary assessment

See below for Feedback on strengths and weaknesses.

4.2.6.2 Feedback on strengths

The teams have an excellent infrastructure for field robotics research, with strong capabilities to build first-of-a-kind systems to perform challenging missions in complex environments. The teams have achieved major levels of highly competitive European funding. International collaboration and reputation are strong in their specialty domains.

4.2.6.3 Feedback on weaknesses

The teams should seek to highlight and enhance their publications in the first-tier international robotics conferences and journals (listed above). This can help the teams to secure more basic research funding. It would be great to increase exposure in the larger robotics community – for example, what would it take to be invited to give keynote talks at leading conferences such as ICRA or IROS. Help make sure that the leaders in the field know about the strong work taking place at INESC TEC.

The teams should articulate specific efforts to increase diversity – they are probably doing some of this already, but it would be great to see outreach to undergraduate interns from diverse backgrounds (to try to recruit them for graduate school), or potentially to have a short summer school with hands-on activities, or something of this nature.

4.2.7 Recommendations

In summary, a few key points are listed here:

- The lab tours and interactions with students were a highlight of our visit.
- Do not hide your light under a bushel – improve your storytelling.
- Highlight how your strong "systems" work also translates into fundamental scientific contributions.
- Focus where you can have the greatest impact – where you are addressing important open problems?
- Create Venn Diagrams showing the expertise of your team and placement among international competitors.
- Create a "scientific roadmap" for your robotics efforts – prioritizing areas of key strength.
- Pursue publications in the leading peer-reviewed Robotics conferences and journals.
- Consider internships, summer school(s), and/or outreach events to increase the diversity of your team.
- At the institutional level, consider metrics that can elevate the top peer-reviewed refereed conferences can be considered as impactful as journal articles.
- Keep up the great work!

5. Acknowledgements

The ARCAR thanks the INESC TEC researchers, administrators, and support staff for the opportunity to assess and learn about INESC TEC scientific activities, the materials and presentations made available, the excellent logistic support of the ARCAR activities and travel arrangements, and the warm hospitality in Porto.

6. Signature Page

The Members of the Ad-Hoc Review Committee on Artificial Intelligence and Robotics:

Dr. José Fortes (Chair) _____ Date _____

Dr. Ulises Cortés _____ Date _____

Dr. Avi Kak _____ Date _____

Dr. John Leonard _____ Date _____

Dr. Petia Radeva _____ Date _____