

---

## **Who Speaks for AI?**

**Eric Eaton** (University of Pennsylvania; [eeaton@cis.upenn.edu](mailto:eeaton@cis.upenn.edu))

**Tom Dietterich** (AAAI and Oregon State University; [tgd@cs.orst.edu](mailto:tgd@cs.orst.edu))

**Maria Gini** (University of Minnesota; [gini@cs.umn.edu](mailto:gini@cs.umn.edu))

**Barbara J. Grosz** (Harvard University; [grosz@eecs.harvard.edu](mailto:grosz@eecs.harvard.edu))

**Charles L. Isbell** (Georgia Institute of Technology; [isbell@cc.gatech.edu](mailto:isbell@cc.gatech.edu))

**Subbarao Kambhampati** (Arizona State University; [rao@asu.edu](mailto:rao@asu.edu))

**Michael Littman** (Brown University; [mlittman@cs.brown.edu](mailto:mlittman@cs.brown.edu))

**Francesca Rossi** (University of Padova and IBM T.J. Watson; [frossi@math.unipd.it](mailto:frossi@math.unipd.it))

**Stuart Russell** (University of California Berkeley; [russell@cs.berkeley.edu](mailto:russell@cs.berkeley.edu))

**Peter Stone** (The University of Texas at Austin; [pstone@cs.utexas.edu](mailto:pstone@cs.utexas.edu))

**Toby Walsh** (The University of New South Wales and Data61; [Toby.Walsh@nicta.com.au](mailto:Toby.Walsh@nicta.com.au))

**Michael Wooldridge** (University of Oxford; [mjw@cs.ox.ac.uk](mailto:mjw@cs.ox.ac.uk))

DOI: [10.1145/2847557.2847559](https://doi.org/10.1145/2847557.2847559)

### **Introduction**

These are boom times for AI. Articles celebrating the success of AI research appear frequently in the international press. Every day, millions of people routinely use AI-based systems that the founders of the field would hail as miraculous. And there is a palpable sense of excitement about impending applications of AI technologies.

Yet, while all this is cause for celebration, some researchers have argued that the discipline of AI is fragmented and largely uncoordinated—and that the entire AI community suffers as consequence. We have a number of different organizations at national and international levels representing the AI community (AAAI, IJCAI, ACM SIGAI, ECCAI, PRICAI, KR, etc.), and a myriad of specialized conferences and journals, frequently with very little coordination or communication between them. The results are obvious. Researchers in distinct sub-fields often work in silos, unaware of work that is occurring in other sub-fields of AI, and the development of the field is hindered by fragmentation.

Moreover, in the public arena, the lack of any authoritative voice for AI creates a vacuum, where ill-informed speculation about the potential of AI is rife, and attention-seeking claims in the popular press receive unwarranted attention, with nobody in a position to

speaking for the field and give an authoritative, informed, and balanced response.

During the 2015 International Joint Conference on Artificial Intelligence (IJCAI), Michael Wooldridge chaired a discussion on this topic between six panelists: Maria Gini, Barbara Grosz, Subbarao Kambhampati, Francesca Rossi, Stuart Russell, and Manuela Veloso. The opinions of the panelists varied widely, from encouraging AI organizations to issue press releases and speak on policy issues, to emphasizing that no one can speak for the entire community and instead promoting healthy individual debate, to stating that we should be the scientific voice presenting only facts and not opinions.

To further this debate, AI Matters invited these panelists and several other leaders in the AI community to respond to this question of “Who Speaks for AI?” This article is a collection of their individual responses, curated by Eric Eaton. The contributors were asked to discuss such issues as the fragmentation of the AI community, the role of professional organizations, public outreach and education on AI, and our responsibilities as individual researchers. Collectively, these responses present a variety of perspectives on these issues and how we, as members of the AI community, can address them.

---

## Speaking Expertly, Speaking Well

**Barbara J. Grosz**

Harvard University

As AI researchers, we each speak for AI and we all speak for AI. In speaking, some of us may focus on the social good of AI-based systems for decision-support, others on the risks or dangers of certain deployments of AI capabilities, and others of the harm done when AI capabilities that could improve outcomes are *not* incorporated in systems. Each of us, as AI scientists, has a responsibility to speak on matters on which we have expertise, carefully avoiding exaggerating either benefits or risks of AI technology.

Is there a role for AI societies also to speak for AI? The key role, most particularly where matters of values and ethics are concerned, flows from an observation in James Moor's classic paper on computer ethics: "It is amazing how many times ethical disputes turn not on disagreements about values but on *disagreements about facts*." (Moor, 1985, p. 267, emphasis mine). The members of AI scientific societies have expertise essential to sorting out the facts, expertise AI societies can bring to the public debate. AI societies working together rather than as individual organizations could—and I think should—play a much larger role in bringing such expertise to bear. Absent taking a stand on an issue, they can shed light on choices, sometimes even making the right decision clear without arguing it. A recent example is provided by the book *Enlightening the Constitutional Debate* issued by the Royal Society of Edinburgh and the British Academy in 2014, as the debate was heating up about the vote on Scottish independence. "The purpose of the series was not to influence the referendum process *in a particular way*; the aim was instead to *encourage rich and informed debate* on this hugely important topic." (Royal Society of Edinburgh website, emphasis mine). The encouragement of "informed debate" on "hugely important topics" can have tremendous impact. It is an important role for AI societies, individually and together.

AI societies should, though, be wary of taking stances on political matters. On almost any controversial issue concerning deployment of AI technologies, society members may differ because they may disagree on facts about the

domains in which the AI is being deployed, facts on which they may not be expert. It is irresponsible to represent as *the opinion of a society* an opinion that is held by *only some of its members*. Which members get to decide what the society thinks and on what basis? How many dissenters make an opinion not any more that of the society? There is an alternative: members of AI societies who think political action important on some concern might work with organizations that have the explicit mission of influencing government policy (e.g., in the United States, the Computing Research Association and the Union of Concerned Scientists).

AI societies have another role to play in seeing that informed voices speak for AI: improving public discourse about the field by helping their members learn to speak effectively to the media. There are many messages not getting out, among them concerns of many in the field about the societal impact of AI (dating back at least to the 1970's when Charles Rosen and Nils Nilsson at SRI spoke of the importance of AI researchers thinking about the economic displacement of "industrial automation") as well as serious consideration of risks and ways to mitigate them, and the increasing number of researchers who focus on AI capabilities for systems to work with people rather than replace people. The incentives for journalists are usually not consonant with researchers' goals. They aim for their articles to be read, responded to, "followed", and so for an exciting story. Deep technical understanding is typically not of interest, and thus our instincts as teachers of AI usually do not serve us well in communicating with journalists. It's an acquired skill. More AI societies could give "crash courses" on science communication like Robohub's (robohub.org).

AI has much to contribute to the world. AI researchers want to make a positive difference in people's lives. We have many great stories to tell; we need to tell them well.

## References

- Moor, J.H. (1985). What is Computer Ethics. *Metaphilosophy* 16:4.
- Royal Society of Edinburgh. *Enlightening the Constitutional Debate*. [www.royalsoced.org.uk/1061\\_EnlighteningtheConstitutionalDebate.html](http://www.royalsoced.org.uk/1061_EnlighteningtheConstitutionalDebate.html)

## If an Organization Speaks for AI, Can It Say Anything?

**Stuart Russell**

University of California Berkeley

It is self-evident that anyone and everyone can speak *for* themselves *about* AI. The 3,000 AI and robotics researchers who signed the open letter on autonomous weapons are speaking for themselves. If anyone is to speak *for* AI, it should be a scientific society with broad membership and a democratic constitution. Well-established membership societies such as AAAI, JSAI, RFIA, and AISB could play such a role. (Better still, for issues of global impact, would be a truly international AI society.) Yet there are some who argue that a scientific society should *never* adopt a position on *any* topic; that is, even if an organization can legitimately claim to speak for AI, it should never speak.

Caution in matters of policy is certainly appropriate. General reasons to avoid taking a position include the following:

- The risk of alienating members who disagree with the position adopted.
- Risks to the society's reputation when it strays beyond its expertise.
- The danger of infringing on the policy prerogatives of governments.

The first issue can be alleviated by a clearly defined, deliberative, democratic process of policy formation. The second issue is mostly a matter of common sense. The third issue is less dangerous than it seems: any organization in civil society may *advocate a policy position*, but typically *makes policy* only with regard to its own membership; neither activity infringes on the role of governments.

There are several legitimate purposes for adopting policies and policy positions. Familiar examples include ensuring adequate levels of research funding, improving opportunities for women and under-represented minorities, and laying out ethical guidelines for the conduct of research. But there are more:

- *Avoiding negative effects*: A scientific society may have special expertise concerning

the possible effects of certain kinds of experiments or applications. For example, molecular biologists in 1975 instituted strong constraints on experiments involving genetic modification of infectious organisms. Failure to self-regulate invites the external imposition of restrictive legislation.

- *Preserving the reputation of the field*: Misconduct in research, misapplication of findings, and negative outcomes for society may all damage the reputation of the field, beyond their intrinsic effects, and it is within a scientific society's legitimate interest to establish policies to prevent such damage.
- *Giving a voice to the community and advocating on behalf of members*: The views of individual scientists are seldom given much heed by the media, compared to those of politicians and celebrities. A scientific society commands attention in ways that individual scientists cannot.

Many major societies have adopted specific policy positions, including the American Medical Association's policy forbidding members from participating in executions<sup>1</sup>; the International Association for Cryptologic Research's policy on mass surveillance<sup>2</sup>; the American Physical Society's policy on nuclear testing<sup>3</sup>; the American Chemical Society's policy on global climate change<sup>4</sup>; the American Society for Biochemistry and Molecular Biology's policy on stem cell research<sup>5</sup>; and the American Psychological Association's 1986 policy against torture<sup>6</sup>, the circumvention of which by a 2005 task force, in the words of the APA Past President and President-Elect, "cast a pall on psychology and psychologists in all countries."

There are urgent issues related to AI, such as automated mass surveillance and autonomous weapons, where one can easily discern the relevance of the legitimate purposes

<sup>1</sup>[www.ama-assn.org/ama/pub/physician-resources/medical-ethics/code-medical-ethics/opinion206.page](http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/code-medical-ethics/opinion206.page)

<sup>2</sup>[www.iacr.org/misc/statement-May2014.html](http://www.iacr.org/misc/statement-May2014.html)

<sup>3</sup>[www.aps.org/policy/statements/97\\_2.cfm](http://www.aps.org/policy/statements/97_2.cfm)

<sup>4</sup>[www.acs.org/content/acs/en/policy/publicpolicies/promote/globalclimatechange.html](http://www.acs.org/content/acs/en/policy/publicpolicies/promote/globalclimatechange.html)

<sup>5</sup><http://www.asbmb.org/uploadedFiles/Advocacy/ASBMB Position Statement on Human Embryonic Stem Cells.pdf>

<sup>6</sup>[www.apa.org/about/policy/chapter-14.aspx](http://www.apa.org/about/policy/chapter-14.aspx)

listed above. For example, misuse of autonomous weapons to cause mass human casualties would be a *negative effect* that would also seriously damage the *reputation of the field*. Regardless of one's position on the issues, it seems *reasonable* for the AI and robotics societies to *consider* adopting a policy; these are not *a priori inadmissible* questions, as some have suggested. The notion that a neutral stance on all matters of policy and on the potential misapplication of scientific knowledge renders a scientific society blameless and "above the fray" is simply incorrect. Inaction is a form of action and those who choose it are morally responsible for its predictable consequences.

---

## Cutting Through the Noise

**Michael Wooldridge**

University of Oxford

I have been an AI researcher for more than a quarter of a century, and over this time I have become used to ill-informed, silly, and sometimes just plain bonkers reports in the media about my discipline. I am used to seeing AI pronounced to be either an impossible folly or something that lies just around the corner. For the most part, I have treated such media reports with the pinch of salt they undoubtedly deserve—something to joke about over coffee. But recently, we have seen media reports of a somewhat different kind: very high profile statements from public figures with impeccable scientific credentials, expressing concern that AI represents an imminent existential risk to the human race. These recent reports are undoubtedly sincere, and because they come from public figures with huge credibility, they have gained a lot of traction. All of a sudden, the language being used to discuss AI in the media has taken on the kind of overtones previously reserved for nuclear weapons.

As an AI researcher, I am deeply concerned about this, for many reasons. It isn't that I believe the singularity is imminent, or that AI represents an imminent existential risk to humanity. My concern is that these public statements, and the associated public debate, are completely uninformed by those who really know what current AI is capable of, and where it is going: the AI research community. My

personal view is that part of my role as a professional scientist (and one that receives public funding for his research) is to inform public debate about my discipline, as the occasion and necessity arises. Of course, I could try to do this as a private individual, but it seems to me that a much more natural avenue to do this would be through our various professional societies: AAI, IJCAI, ECCAI, PRICAI, ACM, IEEE, and so on. And yet, as far as one can tell, little or no effort has been made by these societies to engage with the recent debate in any way: the silence has been deafening.

So, what can or should such societies do? I believe that a professional AI society can and should act as an authoritative and responsible voice on matters concerning AI, actively informing public debate, governments, and policy makers on the potential benefits and risks of AI. A professional AI society could and should be the first point of contact for the media on matters relating to AI, giving measured professional responses to reports such as those that we have seen recently. I acknowledge that there are difficult and potentially controversial value judgements to be made here, and great care needs to be taken in deciding which issues are within the bailiwick of a professional society, and which issues are down to personal ethical judgements.

Some AI researchers, for example, would clearly find it unacceptable for a professional society to issue a statement condemning the development of autonomous weapons—they would view such a statement as essentially political in nature, and would argue that such statements are not the business of a scientific society. I accept there is some strength to this argument, but I don't see that this precludes an AI society taking a more active role in taking control of the agenda in the AI debate, and making measured, professional statements about what AI can and cannot do, and highlighting areas that governments and policy-makers need to be aware of. Such statements can be made without adopting a political standpoint.

If we fail to do this, and in particular, if we fail to step up to the current public debate about AI, then I honestly believe we are not fulfilling our obligations as responsible and professional publicly funded scientists.

---

## The Role of AAAI

### Tom Dietterich

AAAI President and Oregon State University

The Association for the Advancement of Artificial Intelligence is the largest AI professional membership organization. It is frequently contacted by the news media, and generally the elected officers of the organization (President, Past President, and President-Elect) answer media questions. If the questions concern a narrow topic, we refer them to leading researchers with the relevant expertise. In addition to reacting to reporters' questions, the officers have taken the initiative this year. Eric Horvitz (a previous AAAI President) and I wrote a Viewpoint article for the *Communications of the ACM* (Dietterich & Horvitz, 2015) in which we sought to put the fears of long-term AI outcomes into perspective and raise a set of more immediate concerns that need to be addressed by researchers and companies deploying AI technology.

I believe that a membership organization such as AAAI has the most legitimacy to speak on behalf of the AI community. AAAI is an international organization with members from many countries. We are working hard to attract new members, to encourage them to participate in the organization, and to earn their support. AAAI has recently formed two committees to address policy issues. The Ethics Committee is responsible for exploring the issues surrounding the ethical use of AI technology, while the Government Relations committee focuses on encouraging government funding of AI research.

On contentious issues—such as the proper role for AI technologies in military systems—where there is no consensus across the AI community, AAAI has a responsibility to organize debates, workshops, written forums, and so forth to help the community understand the issues. I would like to see the AAAI Ethics Committee summarize the perspectives raised in those events in the form of white papers or AI Magazine articles that fairly present the full range of positions of the community.

Of course prominent researchers will speak out on important issues. It is not the role

of any organization to attempt to dominate the conversation. Instead, AAAI should help the broader AI community (including sub-disciplines that run separate conferences, such as ICML, NIPS, CVPR, ACL, ICAPS, KR, CP, etc.) engage with the issues and consider them carefully. When we speak publicly, our positions should have a technical basis and reflect careful analysis and discussion. In this way, we can become a trusted source of authoritative information about the capabilities of AI technology and its appropriate application.

### References

Dietterich, T. and Horvitz, E. (2015). Viewpoint — Rise of Concerns about AI: Reflections and Directions. *Communications of the ACM* 58 (10): 38–40.

---

## Partnering with the Computing Research Association

### Charles L. Isbell

Georgia Institute of Technology

### Michael Littman

Brown University

AI is a field that addresses a big, fundamental problem with core societal and philosophical implications. A natural consequence is that the field fragments over time, giving rise to dozens of subfields and their conferences, which in turn further fragment the field, sometimes in oddly semi-religious ways.

This fragmentation is bad for the community because we become too siloed to work on the larger problem that inspired us in the first place. It is bad for policy and outreach because it becomes difficult for us to speak to policymakers and the public about the issues of the field; hence, when Hollywood stirs up public sentiment about Skynet destroying us all, there's no scientific voice offering a clear riposte.

One might argue that what our field needs is an equivalent of the Computing Research Association (CRA). CRA's mission is "to enhance innovation by joining with industry, government and academia to strengthen research and advanced education in computing." In particular, "[they] enhance public and policymaker

understanding of the importance of computing to make the case for federal investment in research.”

Actually, that mission sounds like it might cover our needs already. In fact, “CRA’s membership includes more than 200 North American organizations active in computing research: academic departments of computer science and computer engineering, laboratories and centers [...], and affiliated professional societies (AAAI, ACM, CACS/AIC, IEEE Computer Society, SIAM, USENIX).”

That’s right, AAAI.

So, AI already has a solid and established organization that acts as an advocacy center, at least in the United States. We might ask if the outreach and advocacy needs of AI are so distinct that we need another organization separate from CRA. We believe the answer is no; rather, the leadership of AAAI should continue to work with CRA. More to the point, AAAI should act as a unifying force for the researchers who spend most of their time at NIPS, or AAMAS, or CVPR, instead of or in addition to AAAI. Insofar as the field of AI needs organizational leadership, AAAI should simply provide it, working to best leverage its relationship with CRA (which has been quite effective in its advocacy role). Acting as a voice of the community is simply a matter of doing so. If the voice is effective, the community will benefit and follow.

---

## Leading the AI Community

### Toby Walsh

University of New South Wales and Data61

If you had a blank sheet of paper, you would not create multiple AI societies, rival conferences, .... So what should we put on that blank sheet of paper?

First, we need an affordable membership society that appeals to the world AI community and that is growing like our conferences. This means an annual membership fee under \$100 (and less for developing nations), services like job boards, mentoring, local events, etc.

Second, we need AI conferences better run and coordinated. Moving AAAI and IJCAI six months apart is a great opportunity to exploit.

But why do we treat each conference as an one-off? We learn so much each time we run a conference: conflicts, good reviewers for papers, etc. We throw all this away. Conferences also need to be more affordable. It’s not enough to keep the same registration fee. We should be cheaper.

Third, we need a magazine pitched to a general audience that promotes AI like CACM promotes computer science. The public has a real appetite to understand AI. The AI magazine has served a specialist audience well for many years, but we need something now to appeal to a wider audience. It needs topical news stories (like CACM), position pieces (like CACM) addressing societal challenges and survey articles (like CACM). If you haven’t guessed, I’m impressed by how the dynamic editor at CACM has built a mouthpiece for ACM. Great job, Moshe. AI needs something similar.

Fourth, we need to engage the public more. There are immense economic, legal, and ethical challenges facing society as AI and related technologies become available. I agree with those that say science should not be political. But that doesn’t mean AAAI should avoid the political stage. We need to be on the stage, providing sound and impartial advice.

Of course, we may agree on some desirable end goals but that ignores the challenge of navigating there. AAAI must lower its costs, and streamline. For instance, the AAAI conference moves around with no real purpose. The local AI community is barely engaged, and each venue poses new challenges. Here’s a better idea. We pick two venues, one on the East Coast and one on the West Coast, and alternate between them. AAAI can invest effort once in finding two venues, locking in some excellent rates, and then we polish the conference experience each time.

All in all, AAAI needs to do more good for AI. IJCAI and the AI journal annually invest hundreds of thousands of dollars in the AI community. AAAI, with much more in the bank, needs to lift its game here. AAAI also needs to be a force for good by coordinating the vast volunteer effort available within the AI community. AAAI is a society of its members, and it is these members that could best promote AI.

---

## Promoting Synergies Between Conferences and Organizations

**Francesca Rossi**

University of Padova and IBM T.J. Watson

AI is shaping our present and future in ways that were unimaginable only a short time ago. However, it is indeed a fragmented scientific discipline. This fragmentation creates unnecessary boundaries, and does not allow science to advance at full speed.

While the existence of specialized conferences and other events is needed for significant deep research activity, it is absolutely necessary that the AI sub-disciplines have general venues where to interact among each other. The envisioned Federated AI conference, which should include all major AI conferences in a flexible co-location scheme and should take place for the first time in 2018, is one such event. Tools to search for useful research results in other sub-disciplines should also be put into place, to allow researchers to work in a virtual global research environment, where building on top of existing results is easy and redundant work is eliminated.

Broad AI conferences such as AAAI and IJCAI should focus on building communication and synergies among AI sub-disciplines, and should present innovative results in a way that all AI researchers can appreciate, rather than providing deep technical details about specific results such as in any specialized conference. Such conferences should also consistently provide a venue for discussing issues related to the impact of AI on the society as well as its ethical implications. This trend has already started in a very significant way with AAAI and IJCAI 2015, and should steadily continue and grow.

Academic and professional organizations related to AI should be at least linked, if not coordinated, to work in a synergic way. A lightweight umbrella organization, with representatives from all disciplines and associations, should be put into place in order to facilitate communication and interaction, as well as initiatives that go across disciplines and associations. A membership organization such as AAAI is a good format, although the whole

structure and bylaws should be reconsidered to assure scientific and geographical diversity.

Individual AI researchers should be present in the international media with their personal views about AI and its use. However, while the main role of an AI associations should be to foster discussion and information dissemination among its members about scientific advances as well as societal impact, it should also disseminate statements that refer to the responsible use of AI, after consultation with its members. The recently formed AAAI committee on AI and ethics is an example of an initiative that should both foster discussion on these themes and also propose policies and positions to be taken by the executive council of the association.

---

## Embracing Specialization

**Peter Stone**

The University of Texas at Austin

Much has been made of the “fragmentation” of AI over the past years and decades. The field began with a vision of studying the nature of computational intelligence, and quickly researchers began focussing their attention on specific, specialized aspects of intelligence, such as planning, learning, motion, vision, language, etc.

Though many have proposed “remedies” for this fragmentation, it’s not clear that it’s such a bad thing. Specialization has allowed for much deeper understanding and more sophisticated technical advances than would have been achievable by generalists.

Rather than fixing the fragmentation, I suggest that we ought to acknowledge that there will always be people who consider themselves primarily vision researchers or planning researchers (for example), and who consider the specialized conferences and journals as their premier publication venues. That’s not something that we ought to try to change.

Instead, we should encourage a sense that what I’m referring to as “specialized” research can simultaneously contribute to such venues *and* to the more general AI endeavor. AAAI, IJCAI, AIJ, and JAIR should continue to encourage papers and articles from the various

sub-disciplines as they already do. But there are additional steps we can take.

First, I propose that we create a journal modeled after Science or Nature that publishes a greater number of short papers that can be seen as the go-to place to keep up with the field in an efficient manner, publishing summaries of the “best of the best” from the specialized disciplines.

Second, it would be a very positive step to periodically have a confederated conference, much like what happens in the theory community, in which AAAI or IJCAI is collocated with conferences such as ICML, CVPR, ICAPS, AAMAS, etc. so that there are more opportunities for cross-pollination.

With regards to professional organizations, I believe it would be healthiest for the field to have a single membership organization that encompasses all of the specialized areas. If accepted by the community, I believe AAAI is well-poised to fill this role. But if necessary, a new organization could be formed in its place.

I believe that one key role of such an organization should be to keep a pulse on the views of its members with regards to current hot topics being reported in the press, perhaps through surveys distributed to conference attendees (or fellows/senior members). It would be very useful to have a spokesperson or committee serve as an authoritative voice of AI experts that can serve as the natural place for the press and policy members to check in, and to quickly comment on misstatements that are potentially damaging to the field.

Since such a committee will never be able to speak for all individuals, the person or people in charge should keep a balanced view, that includes publishing minority opinions. Even if not perfect, such a voice for the field will surely be better than letting anyone with an opinion about AI grab the headlines and shape public opinion.

---

## **Don't think of a Killer Robot: Framing the Public Debate and Perception of AI**

**Subbarao Kambhampati**

Arizona State University

AI is finally getting significant media exposure and generating public debate. It is hard to watch any media these days without coming across some mention of the progress and impact of AI. This is of course good news; it has certainly been a while since the community had this kind of popular reception. With a few exceptions, however, we seem to be mostly reacting rather than framing this coverage and debate. One worrying consequence of this has been that a significant amount of public debate and coverage is being framed in terms of “stopping evil AI.” The whiplash-inducing speed with which the popular press went from “AI has been a disappointment” to “AI is going to take over” is a spectacle to behold in and of itself.

Of course, some of this doomsday preoccupation is caused by the concerns expressed by people outside AI, as well as some over-the-top reporting. But we in the AI community too seem to be complicit to a certain extent in lending prominence to this preoccupation. A recent example is the extent to which the coverage of AI surrounding IJCAI 2015 has been about autonomous weapons.

I think it is important that the AI community take a more active role in framing the public discourse on AI. We need to focus on emphasizing the societal benefits of AI rather than dwell merely on the possible safety and security issues. Organizations such as AAAI certainly have a role to play here, by educating the public as well as policy makers about the benefits of AI. As a step in this direction, for IJCAI 2016 next year, we will have “Human-Aware AI” as the theme of the conference. While some variation on stopping the risks of AI may well have rhymed better with the current media preoccupation, we deliberately decided on this more general theme, so as to highlight the many positive interactions AI systems will continue to have in cohabiting with humans be they in the form of human-robot teams, human-in-the-loop decision support systems, and human-machine collaborative scenarios.



The issue of “de-fragmenting” the AI community is an important one. That the myriad energetic sub-communities of AI are thriving is a cause for celebration no doubt, but we all believe that the enterprise of AI is bigger than the sum of its sub-communities. We can foster this by highlighting integrated systems, and papers that cross sub-community boundaries, and supporting vigorous tutorial exchange across sub-communities. Flagship conferences such as AAAI and IJCAI certainly have a role to play here, by providing a unified rather than fragmented/federated forums for the field. This is one reason why we are avoiding horizontal sub-tracks at IJCAI 2016.

Coming to the role of professional societies such as AAAI in regards the policy positions, in my view the primary role of a professional society is not so much to take positions on behalf of its membership, but rather to facilitate informed debate among its members, law makers, as well as the public about all sides of an issue of relevance to the field. As the old Fred Friendly saying goes, the mandate is “...not to make up your mind, but to open it. To make the agony of decision making so intense that you can escape only by thinking.” This can, and should, be done through policy forums in the professional magazines (such as AI Magazine) and conferences.

---

## Broadening the Discussion: Ethical Decision Making

**Maria Gini**

University of Minnesota

With the increasing popularity of AI, the media are interested in us. Do we need a consistent voice to address their questions? Perhaps for some issues, but we still have much to do before we can fully understand how AI will change our lives that I think it is not yet time to articulate a concise vision.

AI has reached the point where we know how to build systems that make decisions. We know how to build systems of agents with bounded rationality. Such systems will become more commonly available and will be able to do many tasks currently done by people. The dream of AI is becoming a reality.

How do those systems make decisions? Most

often the choices are based on some form of utility function that computes utility values for each choice. How are the utilities computed? Some human has to design a utility function that balances the factors that affect the decision. After that, a choice is typically cast as an optimization problem (utility maximization or cost minimization).

All good so far. But now think about the design of utility functions for tasks as complex as autonomous driving. One critical task is to decide what to do when something goes wrong. Here is the classical example: your autonomous car is coming to an intersection and detects a car on the cross street about to cross in front against the light. As an imminent collision is predicted, your car has to decide to run into the other car or to swerve into a light pole on the side of the road. The decision algorithm will quickly quantify the cost of killing the occupants of the other car against the value of the life of its own occupants. The issue here is not *if* the car can have ethical values, the issue is *who* determines how this decision will be made and *what ethical values* are incorporated into the decision algorithms.

Those decisions should not be the purview of the AI researchers, but of society at large. As more critical decision-making tasks are delegated to AI systems, the society at large will need to take an ever-increasing role in shaping the behavior of such systems. We need to broaden the conversation on these issues beyond AI researchers. AI organizations and conferences have an essential duty to open these discussions by fostering contributions and exchanges with other disciplines.



**Eric Eaton** is a faculty member at the University of Pennsylvania in the Department of Computer and Information Science, and in the General Robotics, Automation, Sensing, and Perception (GRASP) lab. His research is in machine learning and AI, with ap-

plications to robotics, environmental sustainability, and medicine.



on applications to ecology and ecosystem management.

**Tom Dietterich** is President of the Association for the Advancement of Artificial Intelligence. He serves on the Board of the International Machine Learning Society and on the Advisory Board of the NIPS Foundation. He studies machine learning



navigation in dense crowds. She is an AAAI Fellow, past chair of SIGAI, and is on the editorial board of numerous journals, including Artificial Intelligence, Autonomous Agents and Multi-Agent Systems, Robotics and Autonomous Systems.

**Maria Gini** is a Professor in the Department of Computer Science and Engineering at the University of Minnesota. She studies decision making for autonomous agents in a variety of contexts, from distributed task allocation for robots, to robotic exploration, teamwork for search and rescue, and



She received the ACM/AAAI Allen Newell Award in 2009 and the IJCAI Award for Research Excellence in 2015 for her research contributions to natural language processing and multi-agent systems.

**Barbara J. Grosz** is a Higgins Professor of Natural Sciences at Harvard University. She is a member of the National Academy of Engineering, the American Philosophical Society, the Royal Society of Edinburgh, and the American Academy of Arts and Sciences and a



computing education, and co-created Tech's MOOC-supported MSCS.

**Charles Isbell** is a Professor and Senior Associate Dean of Georgia Tech's College of Computing. He received his B.S. from Tech, and his Ph.D. from MIT. Charles' research interests are varied, but really he does interactive AI. Charles also pursues reform in



of the Association for the Advancement of Artificial Intelligence (AAAI).

**Subbarao Kambhampati** is a professor of Computer Science at Arizona State University. He is the program chair for the International Joint Conference on Artificial Intelligence (IJCAI) 2016 and the president-elect



taught at Duke, Princeton, Rutgers, Brown and several online MOOCs, including the latest course co-taught with Charles Isbell with whom he won a Shakey award in 2014. Littman has served on the editorial boards for JMLR and JAIR, and is an AAAI Fellow.

**Michael L. Littman** primarily studies reinforcement learning, but has done work in machine learning, game theory, networking, solving partially observable Markov decision processes, and computer solving of analogy problems. He has



Editor in Chief of JAIR. She is an ECCAI and an AAAI fellow, and co-chairs the AAAI committee on AI and ethics.

**Francesca Rossi** is a professor of computer science at the University of Padova, Italy, currently on leave at the IBM T.J. Watson Research Center. She has been president of ACP and of IJCAI and is Associate



Stuart Russell is a Professor of Computer Science at the University of California Berkeley. He is a fellow of AAAI, ACM, and AAAS, winner of the IJCAI Computers and Thought Award, holder of the Chaire Blaise Pascal in Paris from 2012 to 2014, and author (with Peter Norvig) of *Artificial*

*Intelligence: A Modern Approach*.



**Peter Stone** is the David Bruton, Jr. Centennial Professor of Computer Science at the University of Texas at Austin. He received his Ph.D. in Computer Science in 1998 from Carnegie Mellon University. Professor Stone's research interests in AI include

machine learning (especially reinforcement learning), multiagent systems, and robotics.



**Toby Walsh** is a Professor of Artificial Intelligence at the University of New South Wales and Data61 (formerly NICTA). He is the Chair of AAAI's International Committee, and an AAAI Fellow.



**Michael Wooldridge** is Head of the Department of Computer Science at the University of Oxford, where he has been a Professor since 2012. His research interests are in multi-agent systems, and he has published more than 350 articles on this subject. He is President

of the International Joint Conference on Artificial Intelligence (IJCAI) from 2015–17.