We Need to Talk about AI: The Case for Citizens' Think-Ins for Citizen-Researcher Dialogue and Deliberation

We Need to Talk About AI:
The Case for Citizens' Think-Ins for Citizen-Researcher Dialogue and Deliberation

September 2022

Authors
Emma L. Clarke, ADAPT, Dublin City University
Harshvardhan J. Pandit, ADAPT, Trinity College Dublin
P.J. Wall, ADAPT, Trinity College Dublin

Contributors
Cara N. Greene, ADAPT, Dublin City University
Laura Grehan, ADAPT, Dublin City University
Aaron M. Jensen, Institute for Methods Innovation
Eric A. Jensen, Institute for Methods Innovation
Anne Kearns, ADAPT, Dublin City University
Aphra Kerr, ADAPT, Maynooth University
Dave Lewis, ADAPT, Trinity College Dublin
Gaye Stephens, ADAPT, Trinity College Dublin

About ADAPT

ADAPT is the world-leading SFI Research Centre for AI Driven Digital Content Technology, which brings leading academics, researchers and industry partners together to deliver excellent science, engage the public, develop novel solutions for business across all sectors and enhance Ireland's international reputation.

ADAPT is pioneering new Human Centric AI techniques and technologies including personalisation, natural language processing, data analytics, intelligent machine translation, human-computer interaction, as well as setting the standards for data governance, privacy, and ethics for digital content.

ADAPT delivers an impactful programme of EPE under the 'Engaging in Our Digital World' theme. This programme aims to empower the Irish public to: acquire the skills necessary to engage fully in our rapidly-evolving digital landscape; develop an interest in, knowledge of, and appreciation for emerging AI innovations; and to participate directly in ADAPT research. These aims align closely with the ADAPT vision.

Executive Summary

Artificial Intelligence (AI) has become one of the most important and ubiquitous technologies across the world. On a daily basis, we interact with powerful AI-based technologies through our use of mobile phones, voice assistants, and even our cars. Despite the widespread adoption of AI, questions and concerns exist around the ethical use of these technologies and their potential to reconfigure our personal and working lives.

The AI - Here for Good[1] strategy (2021) places people at the centre of AI-based technological innovation and prioritises the role played by citizens and other members of society in developing understanding and trust in the potential of AI. The strategy also acknowledges the need to establish public trust in AI. However, to date, there have been few attempts to engage the public with the "social and ethical implications of deploying AI systems across the public and private sectors" (Kerr et al, 2020a). There is a clear and growing need for public dialogue initiatives that create opportunities for discussions on STEM technology. The Creating Our Future Expert Report (2022) indicates that new research engagement programmes should be developed based on models of inclusive participatory and deliberative democracy (pg 62).

The Science Foundation Ireland ADAPT Research Centre has developed the Citizens' Think-Ins model of citizen-researcher dialogue. ADAPT's Think-In series to date has focused specifically on AI and the role it increasingly plays in our lives, and its impact on culture and society. External evaluation of the Citizens' Think-Ins series revealed very positive outcomes for both citizens and researchers who participated in the series.

Over the course of 2020 and 2021, ADAPT's Citizens' Think-In series fostered a wide range of discussions on topics related to AI, ethics and privacy. Participants at Citizens' Think-Ins in 2020 and 2021 took the opportunity to share their concerns, raise issues, identify opportunities and potential drawbacks, and to ask questions within the forum.

At a broad level, three major topics emerged through the public discussions. These are:
1. Trust and Privacy
2. Control and Decision Making
3. Governance and Regulation

---

[1] AI - Here for Good strategy: https://www.gov.ie/en/publication/91f74-national-ai-strategy/

This white paper presents an analysis of the various discussions that took place within the Citizens' Think-Ins series. The discussions are presented with specific reference to citizens and civic society, academia, industry, and policymakers, and provide concrete recommendations to each stakeholder group, to draw parallels between their requirements, and to encourage the periodic use of Citizens' Think-Ins as part of a larger deliberative and participatory approach comprising all stakeholders.

Overall, the Citizens' Think-In series offers a glimpse into how stakeholder-focused discussions raise important issues regarding current and future trends in AI. They also emphasise that individuals and groups provide unique perspectives to the challenges associated with particular domains. AI-focused Citizens' Think-In discussions specifically highlighted that the use of automation and technology has broader concerns that go beyond accountability and trustworthiness - participants also had questions regarding the privacy and security of such technologies and their broader impact on culture and society. The Think-In approach provides stakeholders with an alternative mechanism to identify such issues and concerns, as well as providing a platform to trial what methods can be developed to not only resolve the solution legally or technologically, but also to build trust in the process.

Introduction

Context: The Need for Real Debate on the Role of AI in Irish Society

Recently, Artificial Intelligence (AI) has become more important and ubiquitous across the world. We frequently access powerful AI-based technologies in our everyday lives through our use of mobile phones, voice assistants, and even our cars. These technologies can bring great benefit to our personal and professional lives, but their use requires us to share a wide variety of personal data via a complex ecosystem of websites, social media platforms, applications and data brokers. This sharing often takes place without our knowledge or explicit consent, and even when explicit consent is asked for, it is often written in complex legal and technical language which may be difficult for the non-expert to understand fully. This is problematic for many reasons, and raises questions and concerns about the extent to which we can trust this technology and those who set the standards and legal requirements which govern it. Questions and concerns also exist around the ethical use of AI-based technologies and the potential of these technologies to reconfigure our lives.

This discussion is particularly relevant in Ireland because of the number of multinational technology companies who have based their global and European headquarters here. These companies are involved in all stages of AI technology innovation, and can be seen to drive AI innovation on a global scale. Through the National AI Strategy 'AI - Here for Good' (Department of Enterprise, Trade and Employment, 2021), Ireland aims to be an international leader in using AI to benefit both the economy and society more widely by taking a people-centred and ethical approach to its design, development, adoption, and use. This strategy places people at the centre of AI-based technological innovation and prioritises the role played by citizens and other members of society in developing understanding and trust in the potential of AI. With several new EU regulations on the horizon regarding AI, Data Governance, Digital Services (such as the AI Act[2] and the European Data Governance Act[3]), and other topics, Ireland's approach to the increasing technological ecosystem is important given it is home to a large number of companies.

This strategy requires ongoing and meaningful public engagement on the development, governance and use of AI, and indicates that the Government must "prioritise measures to raise awareness about AI and create meaningful opportunities for public participation in discussion around AI development,

---

[2] The AI Act: https://artificialintelligenceact.eu/
[3] European Data Governance Act:
https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

governance and use cases" (Department of Enterprise, Trade and Employment, 2021, pg 19). However, there is a lack of such participatory initiatives as noted in Kerr et al. (2020a)'s review of forty public documents written since 2011 in Ireland, the UK and by the European Commission. They found that Ireland relies solely on reports on data and AI commissioned from private consultancy firms (such as Deloitte 2021, Accenture 2022, PricewaterhouseCoopers (PwC) and the Analytics Institute 2019), while the UK and European Commission also consult public opinion. The authors note that: "Public consultation in Ireland involves consulting with a small number of academics, existing research centres and companies" and "In Ireland there has been no national attempt to engage with the social and ethical implications of deploying AI systems across the public and private sectors".

This points towards the need for public dialogue initiatives to create discussions on these topics and address current gaps through knowledge co-creation and public engagement.

Public understanding, opinions, and concerns are an important part of Ireland's approach to responsible innovation. The most recent Science Foundation Ireland (SFI) Science in Ireland Barometer found that 65% of the Irish public agreed with the statement "people who will be directly affected by scientific research should have a say in how it develops" (Science Foundation Ireland, 2020, pg 5). SFI's 'Shaping Our Future' strategy 2025 outlines a cohesive and collaborative ecosystem, with opportunities for engagement and dialogue, that aims to encourage the public to "think critically about the challenges we face in society, and to use, respond to, take part in, and own research and innovation" (pg 21). One of SFI's current remits is to "undertake public consultations to identify which societal issues are important to the people living in Ireland" (pg 22). The subsequent Creating Our Future campaign[4] demonstrated that the Irish public is keen to have a say in how the research agenda in Ireland might transform lives and create sustainable societies (Expert Committee Report Creating Our Future, 2022).

In the last two decades, discussions about 'scientific citizenship' and the need for engagement between the sciences, technologists and publics have gained greater acceptance within government and institutional policy and practice in Europe (Irwin 2001; Wilsdon and Willis 2004). An OECD report from 2020 recognises that "citizen participation in public decision-making can deliver better policies, strengthen democracy, and build trust." (Innovative Citizen Participation and New Democratic Institutions - Catching the Deliberative Wave Highlights 2020 OECD Work on Open Government, 2020, pg 3).

---

[4] Creating Our Future: https://creatingourfuture.ie/

Building Trust in AI

Importantly, the Expert Committee report for Science Foundation Ireland's Creating Our Future initiative (2022) highlights the need to ensure an inclusive research system in Ireland with fresh initiatives that "communicate the nature of research and the research process as well as accessible research outcomes" (pg 62). This is at the core of the ADAPT public engagement philosophy, and it is within this context that ADAPT has formulated a public engagement strategy for engagement with people on the broad topic of AI-based technologies, the potential for such technologies to reconfigure people's lives, and the ethical implications of designing, developing, deploying and using such technologies. One such initiative took place in February 2018 where ADAPT researchers, along with collaborators from Maynooth University, explored expectations of AI and ethics through a public survey regarding attitudes towards ethics and AI at Science Gallery Dublin (Kerr et al. 2020b). They found a significant divergence between societal expectations regarding shared responsibility and current practices regarding development and use of AI. Their findings reflect how ethics discourses and solutions operate as assurance for investors and the general public rather than as an effective governance tool. This led to the development of broader programmes within ADAPT to further engage the public in discussions related to AI with the aim of creating and providing a forum through which they can voice their opinions, ideas and concerns on how such technologies might affect their lives and society.

Public engagement using dialogue or deliberative processes can be beneficial to a variety of stakeholders (Dialogue in Public Engagement: A Handbook, 2011). These approaches can be used to:

- take the expertise, lived experience, opinions and feelings of other stakeholders into account to develop a deeper understanding why people feel the way they do;
- reveal new opportunities or solutions by addressing the interests, values and needs underlying people's positions;
- develop a research agenda collaboratively;
- receive inputs to research from specific groups of publics;
- improve the trust and reputation of researchers or academics; and
- understand concerns about a particular field of research and potentially meet these concerns

Balaram et al. (2018)'s Artificial Intelligence: Real Public Engagement report found that in general, people believe that decisions made about how technology is used in relation to them are increasingly beyond their control. Additionally, they found that

people believe there is a lack of trust and transparency when it comes to those who are making the decisions. The authors suggest a potential solution that "decisions should be made with the public, not just for the public" and outline various approaches to citizen engagement such as "deliberative and inclusive dialogue between experts and citizens" (pg 9).

Deliberative or participatory processes are an innovative way to foster citizen participation in STEM activities. Participatory or deliberative processes and activities can take a wide variety of formats, such as Citizens' Assemblies[5], Citizens' Juries[6] and Minipublics[7]. As a public engagement tool, ADAPT's Citizens' Think-Ins are opportunities to connect researchers, industry and policy stakeholders and citizens to generate deeper engagement in the research process.

ADAPT's Citizens' Think-Ins are an effective model of citizen-researcher dialogue, discussion and deliberation and can be used to build mutual understanding and trust, and to help address the aims of the 'AI - Here for Good' strategy. A Citizens' Think-In offers participants the opportunity to interact with those at the cutting-edge of research and development (R&D) to consider and discuss some of the societal and ethical issues facing humanity.

These public discussion forums differ from a more traditional panel discussion or public lecture event as they encourage direct participation from the attendees. Through small group discussions, a Think-In provides an opportunity for people from diverse walks of life to deliberate and discuss topical societal issues arising from STEM innovation.

ADAPT has developed guidelines for hosting a Citizens' Think-In which draw upon the Centre's experience of running participatory dialogue events with the public. To view these guidelines, see Doras, Dublin City University's Online Research Access Repository: https://doras.dcu.ie/27334/.

Citizens' Think-Ins as a Method for Public Engagement

A report published in 2010 by the National Informal STEM Education (NISE) network in the U.S. found that forums provide opportunities for thoughtful consideration of the relationship between science, technology and society (NISE Network, 2010). The

---

[5] Citizens' Assembly:
https://www.citizensinformation.ie/en/government_in_ireland/irish_constitution_1/citizens_assembly.html
[6]IPPOSI Citizens' Jury on Genomics:
https://www.ipposi.ie/our-work/policy/health-information/2022-citizens-jury-on-genomics/
[7]Forms of Mini-Publics: An introduction to deliberative innovations in democratic practice:
https://www.newdemocracy.com.au/2017/05/08/forms-of-mini-publics/

NISE network has developed in-depth informal educational experiences that incorporate dialogue and deliberation around societal implications of nanoscale science, engineering, and technology. Their approach to forums has been used to deliberate on a variety of topics including: nanotechnology, fisheries, diseases and genome editing. The NISE Network forum model provided a foundation for developing the ADAPT Citizens' Think-In format for engaging adults with STEM research.

In November 2018, ADAPT hosted a pilot Citizens' Think-In that saw Irish-based adults and researchers engage in dialogue and deliberation around potential privacy and civil liberties implications of AI. The Think-In addressed some of the factors identified by the Science in Ireland Barometer (Science Foundation Ireland, 2015) as limiting science capital in Ireland, namely lack of connection to the role of science in people's lives, and Irish citizens not feeling they have a part to play in the direction of STEM's role in society. External evaluation[8] of the pilot Think-In (conducted by the Celsius Group) revealed that, post-event:

● public participants felt more informed about AI and more comfortable expressing their views about STEM issues
● public participants had a better understanding of the role of STEM in their lives and felt more engaged with STEM research
● researchers had greater familiarity with diverse points of view relating to AI

ADAPT's Citizens' Think-In series has focused specifically on AI technology applications and the role they increasingly play in our lives. The resulting Citizens' Think-In approach is flexible and it is envisaged that the format can be utilised to generate public discussion and deliberation about other STEM technology areas such as the climate and bioeconomy, smart manufacturing and cybersecurity.

The ADAPT Citizens' Think-In Format

Each two-hour ADAPT Centre Citizens' Think-In comprises:

● A short introductory presentation relating to Artificial Intelligence (AI) in society
● Small-group discussion of technology application scenarios, focusing on possible opportunities, impacts, risks and benefits of the emerging technology
● Report-out from each group with thoughts and recommendations

---

[8] Citizens' Think-In: Artificial Intelligence, Privacy and Civil Liberties Evaluation Report:
https://thinkins.adaptcentre.ie/wp-content/uploads/2021/05/Citizens_Thinkin_Evaluation_Report.pdf

- Whole group reflection and deliberation

Citizens' Think-In participants are encouraged to explore their own views, and those of others, and discuss how we as a society should address sometimes controversial issues arising from emerging STEM research and development. Public Think-Ins are generally aimed at adult audiences but may be adapted to fit a particular need. For example, ADAPT researchers and public engagement experts have developed and run Think-Ins with secondary school students and teachers, as well as for Third Level students and staff. Think-Ins can be hosted in person or online (via platforms such as Zoom).

A Citizens' Think-In differs from other participatory processes like a Citizens' Assembly or Jury as the participants are not randomly selected from the electoral register. For a Citizens' Think-In, the call for participants is advertised widely through community networks, public participation networks[9] and partner organisations, and participants self-select to get involved. Attendees do not require previous knowledge of the subject area to take part in a Citizens' Think-In.

Key Objectives of a Citizens' Think-In

Main aim: To facilitate informed multi-directional conversations about issues, opportunities and challenges where science and societal and ethical issues play a key role.

Key objectives:
- To enable the public to learn about, and reflect on, emerging technology and its potential societal impacts
- To provide an opportunity for citizens to share their perspectives and expertise with researchers who are interested in getting public input to inform their work
- To strengthen the public's and the researchers' understanding of diverse points of view related to applications of current and emerging STEM technologies
- To empower participants to participate in public discourse about technologies that raise complex societal issues, by giving them the opportunity to consider different perspectives and sources of information in a deliberative problem-solving environment

---

[9] Public Participation Networks:
https://www.gov.ie/en/organisation-information/a58b8-community-groups-public-participation-networks/

Co-creation with Stakeholders

Co-creation with stakeholders is a core tenet of ADAPT's public engagement programmes, meaning that projects are advanced *with* stakeholders rather than *for* them. Co-creation underpins ADAPT's Citizens' Think-Ins to ensure that the discussion content is of value to the participants and the audience(s) they represent.

DCU in the Community is an example of a community outreach and lifelong learning centre. Its Bridge to Education programme is designed to equip adult learners from North Dublin communities with knowledge, skills and competencies, as well as build their confidence, in order to fully participate and benefit from future further or higher education experiences. Interacting with new technologies and engaging in dialogue and reflection is a crucial part of these programmes. ADAPT collaborated with students from the Bridge to Education programme to identify themes and test discussion scenarios.

ADAPT researchers in health informatics have established working relationships with various patient and special interest groups. Participants for a co-creation workshop were recruited through these networks. During an online workshop, small group discussions were used to draw from patients' lived expertise and experience to identify themes related to the question of health information. Themes identified in the workshop were subsequently developed into discussion scenarios used at a Citizen's Think-In exploring the question: "What is My Health Information?"

By engaging diverse communities in the co-creation and formative evaluation of the Think-In themes and content, the project partners could ensure that the Think-Ins attracted, resonated with and were beneficial to the intended target audiences.

Citizens' Think-In Themes

In 2020-21, ADAPT hosted 13 public Citizens' Think-Ins to explore various aspects of AI and its benefits and impacts on our lives. The topics discussed in the series include the application of AI in relation to health information, education, smart home technology and shared data spaces. A full list of Citizens Think-In events 2020-21 can be found in the Appendix (pages 27-28). Discussion scenarios were co-created with target audiences for the series and these can be accessed here: https://thinkins.adaptcentre.ie/scenarios/.

Some noteworthy examples of themes include:

ADAPT hosted a Citizens' Think-In to explore the term "Health Information" in March 2021. Outputs from the Citizens' Think-In were provided as input to a subsequent

Citizens' Jury on the theme of Health Information, coordinated by IPPOSI (the Irish Platform for Patient Organisations, Science and Industry)[10].

During Science Week 2021, ADAPT delivered a multi-venue hybrid Citizens' Think-In on the theme 'AI: Friend or Foe?'. This high-profile Think-In was opened by Robert Troy, TD, then Minister of State with Responsibility for Trade Promotion, Digital and Company Regulation, and attracted 105 public participants in discussion hubs in Athlone, Cork, Dublin and online.

ADAPT also partnered with Science Gallery Dublin and the Ars Electronica Festival in 2020 and 2021. Ars Electronica 2020 saw a collaboration between ADAPT and Accenture's Human Insights Lab to co-develop a discussion scenario relating to the use of AI systems to aid decision-making. In 2021, a Citizens' Think-In was co-created with the PROTECT European Training Network (ETN)[11] and stakeholders to address the question: "Who Should I Trust With My Data?".

Insights arising from Think-In Discussions

Over the course of 2020 and 2021, ADAPT's Citizens' Think-Ins fostered a wide range of discussions on topics related to AI, ethics and privacy. Participants shared their concerns, raised issues, tried to identify opportunities and potential drawbacks, and asked questions within the forum. This section presents the findings identified from these discussions based on the salient points of the conversations recorded by scribes present at the event.

At a broad level, three major topics emerged through the discussions, with several other minor topics also being discussed within their periphery. The major topics identified are:
  1. Trust and Privacy
  2. Control and Decision Making
  3. Governance and Regulation

Other minor topics discussed included thematic or topical issues such as security (e.g. the HSE data breach which had recently been in the news), literacy (e.g.

---

[10] IPPOSI: https://www.ipposi.ie/

[11] PROTECT: PROTECT is a multidisciplinary, cross academic-industry and international European Training Network (ETN), funded under the EU's Marie Skłodowska-Curie Actions to train a new generation of Early Stage Researchers (ESRs) as PhD graduates. The network investigates methods to protect the rights and interests of individuals impacted by the continuous large-scale analysis of personal data, while still enabling the economy and society to benefit from rapid innovation in digital applications that use this data and thereby underpin the Digital Single Market.

comprehension of situation and information), vulnerability (e.g. based on gender, age), or were specific about certain domains such as health and medicine, finance, or education.

Trust and Privacy

Trust and **Privacy** were two of the most important and recurrent themes in the discussions across all Think-Ins. In particular, the participants had questions regarding the challenges associated with identifying whom to trust for a given use-case or situation, and whether this is an ongoing process or could be based on other trustworthy characteristics such as the reputation of a company or the accountability of governmental departments and bodies. Similar enquiries were made regarding issues related to privacy - in particular about how to identify the practices and their impacts.

Participants expressed distinct perspectives on privacy through several contexts, such as by associating it with the individual (i.e. self), their families and friends (i.e. groups), within their homes or working environments (i.e. contexts), for issues arising from specific applications such as learning environments or hospitals and clinics (i.e. sensitivity). Participants also discussed the increasing reliance on use of 'smart devices' such as phones, cameras, cars, home automation, and fitness wearables. This led to questions regarding trust and privacy based on how data is being collected, used, and shared through increasingly pervasive uses of technologies.

Participants connected the concepts of trust and privacy to the larger principles of being informed, comprehension of provided information, and accountability of practices. For example, in a discussion on the use of CCTV monitoring devices, participants' questions centred on whether the CCTV camera was visible, had a notice explaining its use, or some indicator for whether the device is active. Similar patterns emerged in other discussions where the understanding of tools and their data practices were connected to their visibility to the participant, the complexity of operations, and reliance on accountability factors such as those provided by the law.

For discussions where trust and privacy were discussed in the context of 'personal data', participants had concerns that included:
- what the data is being used for
- the identities of those collecting, sharing, or receiving the data
- whether such data use provides benefits to participants

For use of data within or by autonomous systems (or AI), participants expressed concerns regarding accuracy (of data, decisions) and the existence of bias within specific contexts (e.g. classroom technology). Such concerns were expressed in a

contextual manner, for example based on existing relationships such as student-teacher or hospital-patient where one entity has an existing implicit trust-based relationship with the other.

Other concerns noted by participants related to what should be considered a "reasonable" expectation of privacy. For example, when comparing public spaces with private ones such as their homes, or whether 'accidentally' recording a neighbour's activities is 'illegal'. Such discussions invariably overlapped with existing theories and concepts from philosophical and sociological domains regarding privacy, society, and 'rights', even if the participants themselves did not discuss them using these concepts. Similarly, participants showed awareness of trust and privacy issues and their implications within the domains of economies (e.g. price/cost of data), technological innovation leading to obfuscation (e.g. invisible tracking and profiling), and limitations on their abilities such as data silos (portability, interoperability) or access to self-repairing capabilities.

Control and Decision Making

The discussions involving trust and privacy led to further questions about control over data and its impact on the agency and abilities of participants. Choice was an important factor in these discussions where participants explored what 'control' or 'factors' they could avail of for making decisions about how technologies affect them, particularly regarding the use of sensitive data categories such as health and when used within vulnerable contexts such as hospitals and job applications.

Questions around control and decision making were also prominent when discussing issues related to safety and security. For example, in a discussion about 'smart doorbells' or automated systems that utilise a camera or other similar systems, participants specifically raised issues that went beyond trust and privacy. Some of these concerns included the nature of providing someone (e.g. a stranger, or unwanted person) with access to their homes, risks of such devices being tampered with or hacked, immediate physical potential harm, and the potential for these to take place through weak security or data breaches.

Other control-related factors discussed were concerned with sovereignty or the capacity for self-decision. For example, discussions related to health data contained explorations of whether, and to what extent, patients can make decisions about themselves, and the capacity for medical practitioners to make decisions on behalf of patients based on necessity. Other topics explored situations relevant for 'individual vs group' notions of privacy, such as where friends and family were involved (e.g. social media), or where data affected more than one person (e.g. genetic data). Participants had concerns about how the use of data permitted by one or few

individuals could affect others as a 'group', and how such issues should be identified and addressed.

Another facet of control that was discussed related to identifying and controlling the inferences drawn about people through the use of data and technologies. A discussion about the use of technologies within classroom settings related to how some tools can be utilised to assess whether students were paying attention. Participants discussed not only common issues such as accuracy, necessity, and bias for such uses, but also identified less explored problems - such as those associated with phrenology (e.g. attentiveness or emotion detection) and conformism (e.g. forcing stereotypical behaviour).

A common finding in these discussions was that while some participants were reluctant to 'trust' technologies, most conclusions agreed that having more control would be better but that the amount of effort this required (from participants) to manage may not be practical. From this, a few discussions led to exploration of a "mechanism for indicating (trust/privacy) would help them make informed decisions" and enable control and decision making for their own data and applications.

Governance and Regulation

Participants raised questions regarding rules and regulations for ensuring the safety, security, and accountability of practices in connection with trust, privacy, and control. A common theme amongst these was the role, ability, and extent of governments (or government bodies) to create and utilise appropriate laws for addressing specific concerns. The need for a system of "checks and balances" was reflected across discussions, though participants also specifically elaborated on their desire to perform some tasks themselves, such as management of data and/or technology, with the law supporting them ensuring and providing the capability and means.

Specific instances of existing regulations, such as the General Data Protection Regulation (GDPR), were utilised in discussions with participants expressing a varied amount of familiarity with its contents and uses. Whilst most (if not all) participants were aware of the rights and obligations espoused by GDPR, they discussed issues such as the complexity of understanding how their data is being used, what technologies are being used and who manages them, and how they should understand the 'risks' of such practices.

Discussions around 'simplifying' privacy policies or terms and conditions based on 'common sense' were explored as one potential solution. Some participants also

discussed proactive measures, such as citizens engaging with governing bodies to make policies which companies would be required to follow.

Topics related to automation, including the development and use of AI, also mirrored similar concerns and calls for regulation. In this, citizens expressed lesser or no familiarity with the technologies involved, the complexity of understanding what (if any) rules should be in practice, but still asking for "strong measures" in terms of legislating and regulating to ensure problematic uses are regulated and to ensure citizens are made aware of such measures.

Analysis of Participant Perspectives

Citizens as a Stakeholder Group

Stakeholders, whether as individuals or a group, are an integral part of policy making - especially where it concerns developments that affect them at personal, professional, and societal levels. The Citizens' Think-Ins provide a glimpse into how such stakeholder-focused discussions raise important issues regarding current and future trends. They also emphasise that individuals and groups provide unique perspectives to the challenges associated with a particular domain, and that their discussions are a key to identifying and ensuring socio-technical solutions are designed correctly from the onset.

More specifically, the discussions lead towards a research agenda not only for participation in the policy-making stage, but also its subsequent stages regarding policy effectiveness and policy enforcement. This requires citizens to be informed, or at least made aware, of the policies and developments that take place within specific domains, sectors, areas, or environments - which currently can only be achieved in a limited capacity through conventional means. Such communications must leverage social activities and media to provide information where it is contextually relevant. Some examples of these in which participants expressed interest or willingness to participate include: literacy campaigns (e.g. related to data, privacy, AI), workshops (e.g. assessing technologies), forums for discussion (e.g. Citizens' Think-Ins and juries).

Along with the establishment of an ongoing communication process involving citizens, there is also an immediate need to establish feedback mechanisms between stakeholders. Mechanisms such as the Citizens' Think-Ins can provide citizens with a way to provide their voice and perspectives to the government to specific policy-based areas of governing bodies and to focused research groups within academic environments. They can also establish a dialogue with the companies that create and/or provide technologies. This requires facilitating citizen panels and

co-creation groups of relevant stakeholders within the larger Think-In framework in order to stimulate and support discussion which leads to informed policy making.

Effective and Proactive Regulations

Regulations provide obligations and establish requirements for how information models should be utilised. In practice, this only creates a baseline for what a situation should ideally look like. The use of legal language, differences between legal concepts and what is presented to citizens in actuality, and common use of deceptive patterns such as obfuscation and nudging can lead to scenarios where such laws are either not considered sufficient or not adequately applied. This was evident in the Think-Ins, where participants specifically called for better rule-making, its effective enforcement, and proactive measures that provide them with the required safety and security that they expect from their data uses and technological applications.

This discussion indicates the need for better legal mechanisms to deal with common problems through creation and enforcement of regulations, as well as better socio-technical mechanisms such as easier management and reporting mechanisms. The Citizen's Think-Ins also provided insight into the potential for developing technological solutions that empower users to track and manage their preferences related to data, AI (e.g. smart devices), or other technological contexts. This reflects an important ethical and philosophical consideration regarding the 'agency of the user', as well as a critical research direction for ensuring responsible innovation within a rapidly progressing field such as AI.

Contextual Information Models

The expression "the right information in the right place at the right time" describes a framework for information models (Fischer, 2012) that aids in facilitating and ensuring responsible and ethical developments based on mechanisms that rely on the participants' agencies - such as consent. Participants in the Think-Ins mentioned their concerns about assessing information for clarification of their issues regarding trust, privacy, ability to control, and risks. This calls for better models that can provide information without overloading their users with complexity. Participants reacted positively towards trialling novel approaches such as the simplification of text within policies, use of visual mechanisms for comprehension of complex information, mechanisms that personalise content (e.g. express whether a personal concern is applicable), and to establish a contextual dialogue between agencies - such as between a doctor and a patient.

These topics provide opportunities that are critical and important to the functioning and success of information-based processes, such as the use of privacy notices, or in ensuring that citizens are aware of their rights and abilities. They are also necessary as a first step towards encouraging people to accept and utilise technologies by giving them the means and mechanisms to comprehend, at their own comfort and

capability levels, what they are using, how it functions, and how they (and their personal data) are affected by it.

Given that the Think-Ins were primarily focused around the two topics of data governance and AI, the interpretation of contextual information models for these provides ample research exploration opportunities for researchers. For example, the explainability of AI or automated decisions is an important topic that has several issues identified and discussed within the Think-Ins. By adapting the contextual information associated with these based on the citizen's perspectives, new avenues for solutions can be identified. This is an important aspect of both research as well as product development, since user studies are typically focused on trialling created technologies (or an approach) with the citizens, instead of engaging them at an earlier stage to understand what their *needs* and *issues* are.

Think-In Outcomes

External evaluation carried out by the Institute for Methods Innovation demonstrates that participants in ADAPT's Citizens' Think-Ins found the experience to be very positive overall. A total of 338 participants attended a Citizens Think-In in 2020-21. Of those, 83 participants completed the pre-event survey and 35 completed the post-event survey.

Of the respondents, 94% indicated that they enjoyed the Citizens' Think-In they attended. Overall, 83% of respondents agreed (46%), strongly agreed (37%) or somewhat agreed (6%) that they were actively able to participate in the event and 90% felt comfortable asking questions at the event. 89% of respondents felt the workshop was a good use of their time. 71% of respondents agreed (38%), strongly agreed (24%) or somewhat agreed (9%) that their contribution to the event was valued. Overall, 86% of the participants said they were satisfied with the event. These results indicate that the Citizens' Think-In format creates a positive environment for dialogue and deliberation. After attending a Citizens' Think-In, the majority of public participants felt that AI research is fascinating (87%), interesting (93%), useful (85%) and beneficial (75%).

Attitudes towards AI researchers saw a positive shift on the untrustworthy - trustworthy scale (12% untrustworthy in the pre-event survey, dropping to 0% untrustworthy in the post-event survey). This indicates that the interactions with researchers increased respondents' trust in the researchers. The evaluation also demonstrated there is less trust among participants when it comes to AI research itself. 29% of respondents found that AI research is untrustworthy according to the post-event survey results. Although there is more work to do in terms of building trust in AI research, the positive shift in perceptions of trust towards researchers

after attending a Citizens' Think-In demonstrates that this method of public engagement could be effectively used as a way to increase trust.

The evaluation found that being involved in a Citizens Think-In was a very positive experience for researchers. 89% felt that a Citizens' Think-In was ideal for engaging with members of the public. The researchers who responded found the feedback they received from the public was relevant (100%), useful (100%), valuable (100%) and inspiring (89%).

The full results of the evaluation can be found at: https://bit.ly/ThinkInEvaluation22

Conclusions

Recommendations for Stakeholders

Based on the discussions that took place within the Think-Ins and their analysis, we present our conclusions based on the Quadruple Helix four-stakeholder model consisting of citizens, academia, industry, and policy makers. Through these, we aim to provide concrete recommendations to each group, to draw parallels between their requirements, and to encourage the periodic use of Citizens' Think-Ins as part of a larger deliberative and participatory approach, comprising all stakeholders.

Citizens

1. Think-ins are an excellent activity for engaging citizens in knowledge generation, for identifying opportunities for guided research, for utilising identified issues and perspectives for evidence-based explorations, and for understanding current notions of social, cultural, political, and ethical perspectives.
2. Think-ins provide an invaluable window for citizens to understand the broader ethical implications of research, technologies, and related policies - which is typically under-investigated and contains specific groups of stakeholders who may be at a risk of being disconnected from societal concerns and needs.
3. Think-ins facilitate awareness and a culture of discussion through which society can be provided an opportunity to influence research, progress, and policy making. They can enable citizens to provide their perspectives and questions, which can be collectively addressed and resolved.
4. The divergence of how people think, and the nuances between different concepts such as privacy, security, transparency, accountability and ethics is evident when engaging with the public through forums such as Think-Ins. Understanding of such differences is crucial to ensuring that citizens' voices do not lead to misinterpretations of their concerns.

5. These Think-Ins showed that people, in general, are aware of the *value* of their data and enquire about what *benefits* they get in return. This is a topic that should be further explored to identify better mechanisms and dialogues with/for citizens and what kind of information they would like to receive about it.

Academia

1. Evaluation found that participating in a Citizens' Think-In is a useful way for researchers to gain meaningful insights into public views on their areas of research.

2. Theories of privacy that connect the legal obligations and enforcement to citizen expectations need to be developed, discussed, and refined. The current models where privacy is taken at either extreme points (e.g. law vs societal) are not sufficient to ensure citizen trust, given that legal, commercial, and academic practices are all considered too complex to grasp for the general public. Instead, better mechanisms that connect these domains to answer the specific issues based on common perspectives may provide a more approachable way for discussing privacy and AI (or other STEM topics) with the public.

3. Academia should identify or further develop current mechanisms for information provision and comprehension. Such mechanisms should provide a better way for citizens to understand data practices, the use of automation or AI, and assist both policy makers and technology creators/providers in conveying different levels of technical and domain-specific knowledge to the general public.

4. While most research has focused on the individual concerns (e.g. self managed preferences), collective models and instruments for identifying and answering questions related to topics such as privacy and accountability can lessen the burden on stakeholders. These topics can provide a key area to explore various ways through which societal concerns can be identified and utilised as a group.

Policymakers

1. Current laws have limitations for tackling the important issues expressed by citizens through these Think-Ins. However, creating new laws also requires time and effort, and has additional considerations that may severely burden regulators and authorities. A way out of this circular problem could be to establish granular mechanisms that are readily accessible to citizens, and that reduce the complex legal processes and knowledge to something that individuals (or small groups) can manage. This can be achieved through creation of smaller governance bodies, or further empowerment of citizen

representative groups (e.g. consumer protection), or providing citizens with the ability to directly participate within policy making at various levels through novel formats such as Citizens' Juries, Assemblies, and Think-Ins.

2. The Citizens' Think-Ins surrounding AI specifically highlighted that use of automation and technology has broader concerns that go beyond accountability and trustworthiness. Citizens also had questions regarding the privacy and security of such technologies, which were framed within the same context as any other tool or device they used. While the proposed European AI regulation does address issues related to AI, and the existing obligations provided by GDPR address those associated with data privacy and security, the specific inter-relationship between the AI Act and the GDPR must be further elaborated to identify and pre-empt issues and enforcement challenges. Think-ins provide stakeholders with an alternative mechanism to identify such issues, as well as trial what methods can be developed to not only resolve the solution legally or technologically, but also to build trust in the process.

3. There is a growing body of work that relates privacy and morals to monetary value - for example by enabling data to be 'sold'. The Citizens' Think-Ins show that such concepts are poorly understood, and that their implications as well as issues are non-apparent given their novelty and niche knowledge. Therefore, future regulatory proposals such as the European Data Governance Act and Data Spaces[12] (e.g. for Health) should first explore whether citizens have the necessary data literacy to understand how they will function, assess their risks, and thus have the capability to make informed decisions. In addition, these should also assist in identifying and regulating certain practices based on what 'value' can be exchanged for which data, and how citizens can be assured that such transactions are legal, ethical, and not exploitative.

Industry

1. Companies need to innovate in terms of design and practices related to how they communicate about privacy and trustworthiness; conventional mechanisms are not sufficient, and so new methods should be developed that directly address citizens' concerns rather than merely satisfying complex legal obligations for information provision.

2. There needs to be reliable and comprehensible mechanisms for conveying that technology is accurate, unbiased, and secure - this has to go beyond statistical values (such as 98% accurate facial recognition) and relate to the contextual use of that technology within a use-case (such as "we correctly

---

[12] Common European Data Spaces:
http://dataspaces.info/common-european-data-spaces/#page-content

recognise faces at your door in 99% of instances").

3. There is an immediate need for human involvement and conflict resolution mechanisms which are explored and applied within the AI by-design; for example, what happens when facial recognition does not work in certain cases, or when there is a false-positive.

4. Multidisciplinarity is key - what works in one situation may not necessarily work in other contexts. Consequently, industry must involve diverse stakeholders in the development and application of AI.

5. An understanding of the social, cultural and political context in which the AI technologies are being developed and implemented is key. The mechanisms that directly address citizens' concerns cannot be universal; they will need to be socially and culturally contextual and be designed for the specific cohort of people who are impacted by the technology. This includes notions of privacy and trustworthiness which are likely to have different meanings for different groups of people.

Next Steps

The next phase of the Citizens' Think-In project is underway and is innovating by engaging a 'Quadruple Helix' (QH) of stakeholders from academia, civil society, industry and government (Carayannis and Campbell, 2012). The QH approach brings together a range of project partners, each with a unique perspective to enhance dialogue on evolving societally-relevant topics. The QH is a multi-stakeholder approach which sees government, academia, industry, and civil society as key actors to promote a democratic approach to innovation. Hasche et al. (2020) defines the QH as "a network of relationships" in which the various actors "interact in value-creating processes to transform various inputs into valuable outputs for themselves and others".

In the 2022 series, known as '#DiscussAI Think-Ins', wider stakeholder engagement will result in meaningful conversations about the role of AI in society and its impacts in different social and cultural contexts, with greater potential to influence emerging research and public policy.

A co-creation approach continues to be at the heart of every stage of the Think-In process: from addressing priorities for deliberation, topic content development, Think-In refinement and co-evaluating the process, so that all stakeholders (including citizens) have an opportunity to define what "success" looks like for them.

If you are interested in participating in or hosting your own Citizens' Think-In, please contact emma.clarke@adaptcentre.ie

To find out more: https://thinkins.adaptcentre.ie/

Works referenced

Balaram, B., Greenham, T., & Leonard, J. (2018). Engaging citizens in the ethical use of AI for automated decision-making. [online]
Available at:
https://www.thersa.org/globalassets/pdfs/reports/rsa_artificial-intelligence---real-public-engagement.pdf [Accessed 24 Aug. 2022].

Bradley, C. and Wingfield, R., 2020. National Artificial Intelligence Strategies and Human Rights: A Review. Global Partners Digital and Stanford Global Digital Policy Incubator. [online]
Available at:
https://www.gp-digital.org/wp-content/uploads/2020/04/National-Artifical-Intelligence-Strategies-and-Human-Rights%E2%80%94A-Review_April2020.pdf [Accessed 24 Aug. 2022].

Carayannis, E.G. and Campbell, D.F. (2012). Mode 3 knowledge production in quadruple helix innovation systems. In Mode 3 knowledge production in quadruple Helix innovation systems (pp. 1-63). Springer, New York, NY.

Cunningham, Y. (2019). Citizens' Think-In: "Artificial Intelligence, Privacy and Civil Liberties" Evaluation Report. Celsius Group, Dublin City University.
Available at:
https://thinkins.adaptcentre.ie/wp-content/uploads/2021/05/Citizens_Thinkin_Evaluation_Report.pdf [Accessed 24 Aug. 2022].

Department of Enterprise, Trade and Employment. (2021). AI - Here for Good: National Artificial Intelligence Strategy for Ireland. [online]
Available at: https://www.gov.ie/en/publication/91f74-national-ai-strategy/
[Accessed 24 Aug. 2022].

Dialogue in Public Engagement: A Handbook. (2011). [online]
Available at:
https://edinburghbeltane.files.wordpress.com/2011/11/dialogue-handbook-final.pdf
[Accessed 24 Aug. 2022].

Expert Committee Report Creating Our Future. (2022). [online]
Available at:
https://creatingourfuture.ie/wp-content/uploads/2022/07/Creating-Our-Future-EXPERT-Report.pdf [Accessed 25 Aug. 2022].

Fischer, G. (2012). Context-aware systems: the'right'information, at the'right'time, in the'right'place, in the'right'way, to the'right'person. In Proceedings of the international working conference on advanced visual interfaces (pp. 287-294).

Hasche, N., Höglund, L. and Linton, G. (2020). Quadruple helix as a network of relationships: creating value within a Swedish regional innovation system, Journal of Small Business & Entrepreneurship, 32:6, 523-544

Innovative Citizen Participation and New Democratic Institutions - Catching the Deliberative Wave Highlights 2020 OECD Work on Open Government. (2020). [online] Available at: https://www.oecd.org/gov/open-government/innovative-citizen-participation-new-democratic-institutions-catching-the-deliberative-wave-highlights.pdf [Accessed 24 Aug. 2022].

Irwin (2001). Constructing the scientific citizen: Science and democracy in the biosciences, Public Understanding of Science, 10(1), 1-18.

Kerr, A., Barry, M. and Kelleher, J.D.. (2020a). Ethical Artificial Intelligence. [online] Media Literacy Ireland.
Available at:
https://www.medialiteracyireland.ie/news/we-expect-ethical-artificial-intelligence-who-will-deliver-it [Accessed 24 Aug. 2022].

Kerr, A., Barry, M. and Kelleher, J.D. (2020b). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. Big Data & Society, 7(1), 2053951720915939.

NISE Network (2010). Forums | NISE Network. [online]
Available at: https://www.nisenet.org/forums [Accessed 24 Aug. 2022].

Science Foundation Ireland. (2020). Science in Ireland Barometer 2020. [online]
Available at: https://www.sfi.ie/engagement/barometer/ [Accessed 24 Aug. 2022].

Science Foundation Ireland. (2015). Science in Ireland Barometer 2015. [online]
Available at: https://www.sfi.ie/resources/SFI-Science-in-Ireland-Barometer.pdf [Accessed 24 Aug. 2022].

Wilsdon, J. & Willis, R. (2004). See-through science: why public engagement needs to move upstream. Project Report. Demos, London.

Acknowledgement of Contributions

Appendix

Citizens' Think-Ins events 2020 - 2021

Event Name	Date	Place	Topics	No. of Participants	No. of Researchers
STE(A)M in Junior Cycle 2020 workshop: *How might Artificial Intelligence transform society and what are the possible consequences of this?*	1 Feb. 2020	Drumcondra Education Support Centre, Dublin 9	1. Artificial Agents 2. Driverless Cars (MIT Media Lab's Moral Machine) 3. Social Credit System	12	5
STE(A)M in Junior Cycle 2020 workshop: *How might Artificial Intelligence transform society and what are the possible consequences of this?*	29 Feb. 2020	Navan Education Centre, Co. Meath	1. Artificial Agents 2. Driverless Cars (MIT Media Lab's Moral Machine) 3. Social Credit System	12	5
Citizens' Think-In on Artificial Intelligence: *What are the ethical issues we face in data storage and dissemination in the digital age?*	9 Sept. 2020	Online - Ars Electronica 2020	1. Artificial Agents 2. Driverless Cars (MIT Media Lab's Moral Machine) 3. Social Credit System 4. Medical Agents and Trust	27	13
Citizens' Think-In on Artificial Intelligence in Society	27 Nov. 2020	Online - European Researchers Night	Digital Home Assistants	3	5

STE(A)M in Junior Cycle 2021 workshop: *How might Artificial Intelligence transform society and what are the possible consequences of this?*	3 Mar. 2021	Online	1. Driverless Cars (MIT Media Lab's Moral Machine) 2. Social Credit System	18	7
Citizens' Think-In: *"What is My Health Information?"*	4 Mar. 2021	Online	Health information	29	13
Citizens' Think-In: *"Who Should I Trust With My Data?"*	9 Jun. 2021	Online	Data Governance	14	6
Citizens' Think-In: *"Who Should I Trust With My Data?"*	9 Sept. 2021	Online - Ars Electronica 2021	Data Governance	23	9
Citizens' Think-In: *"What is My Health Information?"*	7 Oct. 2021	Hybrid - online and the Guinness Enterprise Centre	Health information	15	8
Citizens' Think-In: "AI - Friend or Foe?"	11 Nov. 2021	Dublin, Athlone, Cork	Technology in the home	44	28
Citizens' Think-In: "AI - Friend or Foe?"	11 Nov. 2021	Online	Technology in the home	61	
Total (13)				**338**	**109**

Engaging Content
Engaging People