ETAPAS

Ethical Technology Adoption in Public Administration Services

Funded by the Horizon 2020 Framework

# Deliverable 2.2

## Code of Conduct for DTs

HORIZON 2020

The Framework Programme for Research & Innovation
Research and Innovation actions (RIA)

# ETAPAS

ETHICAL TECHNOLOGY ADOPTION IN PUBLIC ADMINISTRATION SERVICES

## Grant Agreement Number: 101004594

[H2020-SC6-TRANSFORMATIONS-2020] Socioeconomic and Cultural Transformations in the context of the fourth industrial revolution

***The Framework Programme for Research & Innovation***
***Research and Innovation actions (RIA)***

*Project Title:*

Ethical Technology Adoption in Public Administration Services

# ETAPAS

**Grant Agreement No: 101004594**

**[H2020-SC6-TRANSFORMATIONS-2020] Socioeconomic and Cultural Transformations in the context of the fourth industrial revolution**

**Deliverable**

## D2.2 – Code of Conduct for DTs

| | | | |
|---|---|---|---|
| Deliverable No. | | **D2.2** | |
| Work Package No. | **WP2** | Work Package Title | **Responsible DT adoption: ethical principles and legal framework definition** |
| Task No. | **T2.2** | Task Title | **Ethical principles definition** |
| Lead beneficiary | | **KTH Royal Institute of Technology** | |
| Dissemination level | | **PU** | |
| Nature of Deliverable | | **Report** | |
| Delivery date | | **30th April 2021** | |
| Status | | **F: Final** | |
| File Name: | | **ETAPAS D2.2 – Code of Conduct for DTs.pdf** | |
| Project start date, duration | | **01 November 2020, 36 Months** | |

This project has received funding from the European Union's Horizon 2020 Research and innovation programme under Grant Agreement n°101004594

## Authors List

| **Authors** (Editors) | | | |
|---|---|---|---|
| *Surname* | *Initials* | *Beneficiary Name* | *Contact email* |
| Hansson | SOH | KI | soh@kth.se |
| Fröding | BF | KTH | barbro.froding@abe.kth.se |

## Reviewers List

| **List of Reviewers** (in alphabetic order) | | | | |
|---|---|---|---|---|
| # | *Surname* | *Initials* | *Beneficiary Name* | *Contact email* |
| 1 | Galasso | GG | PwC | giovanna.galasso@pwc.com |
| 2 | Mancini | SM | PwC | sara.mancini@pwc.com |
| 3 | Nastrelli | AN | MEF | antonio.nastrelli@mef.gov.it |
| 4 | Cappuccio | MC | MEF | michele.cappuccio@mef.gov.it |
| 5 | Mazzoli | AM | MEF | angelo.mazzoli@mef.gov.it |
| 6 | Mureddu | FM | Lisbon Council | francesco.mureddu@lisboncouncil.net |
| 7 | Benanti | PB | EAB member | benanti@unigre.it |
| 8 | Ferro | EF | EAB member | enrico.ferro@linksfoundation.com |
| 9 | Mergel | IM | EAB member | ines.mergel@uni-konstanz.de |
| 10 | Neuwirth | RJN | EAB member | rostam.neuwirth@alumni.eui.eu |

| **Document history** | | | |
|---|---|---|---|
| *Version* | *Date* | *Status* | *Modifications made by* |
| 1.0 | 18/02/2022 | Draft | KI |
| 1.1 | 11/03/2021 | Draft | KI |
| 1.2 | 31/03/2021 | Draft | KI |
| 1.3 | 06/04/2021 | Draft | KI |
| 1.4 | 13/04/2021 | Draft | KI |
| 1.5 | 15/04/2021 | Draft | KI |
| 1.6 | 29/04/2021 | Final version of the full deliverable | KI |

## The ETAPAS Consortium consists of:

| Logo | Organisation acronym and name | County |
|---|---|---|
| | MEF – Ministero dell'Economia e delle Finanze | Italy |
| | PwC – PricewaterhouseCoopers Public Sector Srl | Italy |
| | 2021.AI – 2021.AI APS | Denmark |
| | SINTEF – SINTEF AS | Norway |
| | PROKOM – Sem & Stenersen Prokom AS | Norway |
| | CERTH – Ethniko Kentro Erevnas Kai Technologikis Anaptyxis | Greece |
| | IIT – Fondazione Istituto Italiano di Tecnologia | Italy |
| | LC – The Lisbon Council for Economic Competitiveness ASBL | Belgium |
| | KTH – Kungliga Tekniska Hoegskolan | Sweden |
| | CEA – Commissariat à l'énergie atomique et aux énergies alternatives | France |
| | FDG – Fondazione Don Carlo Gnocchi Onlus | Italy |
| | MUKA – Municipality of Katerini | Greece |
| | UNIGRAZ – University of Graz | Austria |
| | KI – Karolinska Institutet | Sweden |

# Contents

# 1. Introduction

The generic code of conduct is intended to be a starting point for the public organizations in their own work with codes of conduct and other policy documents. The development of the code of conduct was mainly undertaken by the KTH and KI partners. A first draft was written in the first months of the project, based their expertise in the subject, consultations within ETAPAS, and a cursory reading of much of the normative literature on disruptive technologies in the public sector that will be reviewed and referred to in other deliverables within the project.

A workshop for the whole consortium on the 12th of February was devoted to the code of conduct. An updated version and a template for feedback was circulated on the 19th of February. The revised version was shared with the External Advisory Board on the 14th of March. They all answered, and based on their comments a revised version was sent out for a new round of consultations within the project int the first half of April, leading to further revisions of the text. A part of the first external project Workshop, "Disruptive technologies in the Public Sector: how to manage ethical risks and social impacts?" (April 30th), was devoted to the code of conduct. The participants were invited to a public consultation on the code of conduct, to be performed in May 2021.

The code of conduct will be a living document, and it will be revised several times during the remainder of the project period.

# 2. Code of Conduct for disruptive technologies in the public sector

A code of conduct is a statement of principles adopted by an organization to guide its own conduct. These principles can serve both as affirmations of values and as the basis of an internal accountability mechanism for the organization. Codes of conduct usually focus on ethical requirements and recommendations that go beyond what the law requires. In order to stimulate discussions on the effects of disruptive technologies in the public sector, the research project ETAPAS has produced this first draft of a generic code of conduct. We hope that it will be used in public sector organizations as a starting point for developing more detailed policies and standpoints on their own uses of new technologies.

Experience shows that a code of conduct can contribute significantly to an organization's ability to deal with difficult ethical issues, in particular if the code is developed in a process with broad participation within the organization. This generic code is based on a draft that was informed by studies of the research literature, and then further developed in an iterative process, in which comments on each version were used to produce an improved version. Such a process is to be recommended also for the development of other, more specialized codes of conduct. Continued discussions after a code has been adopted, for instance facilitated by an ethical committee, can further contribute to the organization's ethical culture and its ethical competence.

Disruptive technologies are technologies whose introduction disrupts important social structures. Electricity, television, computers, and the Internet are historical examples of disruptive technologies. They have all changed our ways of life, how we work and how we communicate with others. We are currently experiencing the rise of new technologies that may become disruptive, such as artificial intelligence, robots, the Internet of things, new materials, synthetic biology, drones, blockchain technology, quantum computing, virtual and augmented reality, and new medical technologies.

Disruptive technologies can have large impacts in the public sector. This applies for instance to chatbots, tools for detecting tax fraud, decision support for welfare provision, and the use of robots in social care and healthcare. These technologies come with promises and new possibilities, but they can also create new problems in society, as well as exacerbate already existing ones. Preparations are needed to ensure that we make well-informed decisions on whether, and in that case how, we are going to adopt and use these technologies.

The public rightly has high expectations on the public sector, which works for the whole community and is accountable to it. Public officials work with the public's money and on behalf of the citizens. They also have considerable power over individuals. Therefore, it is of particular importance that the public sector's own use of disruptive technologies is in the interest of the public good, leaves no one behind, and is transparent and open to public scrutiny and democratic decisions. This amounts to a double task. The public sector must actively pursue the efficient use of new technologies in order to improve its performance, provide new services, and reduce costs. At the same time, it also has to exert strict controls and continuous evaluation to ensure that no one is treated unfairly or otherwise disadvantaged through the use of these technologies. In these and other respects, the public sector has to be exemplary not only in following the law but also in achieving ethical standards that go beyond what has been codified in the law.

## 1. Environmental sustainability

The public sector's use of disruptive technologies must conform to the principles of environmental sustainability. The public sector has a responsibility to ensure that the use of these technologies contributes to, or at the very least does not impede, European and national policies to achieve climate neutrality. It also puts strict demands on the production processes for the hardware and the materials used to produce the new technology. New materials, such as nanomaterials, often have toxic and ecotoxic properties that need to be carefully evaluated in advance of any use that can affect the environment. Biodiversity is a precondition for ecological resilience, and it must always have an important role in the evaluation of environmental effects of new technologies. Production processes must be assessed globally. The public sector should take the lead in ensuring that when importing technological products, we do not thereby export environmental problems.

## 2. Justice, equality and the rule of law

Equal treatment and the rule of law are fundamental values from which public administration should never deviate. New technologies can be harnessed to implement them and to ensure the basic principle that like cases should be treated alike. The introduction of technologies that promote equality of access and opportunity should be strongly prioritized. But technologies can also, inadvertently, have the opposite effect. For instance, algorithms can become biased against minorities or against women if they are trained on data that reflects previous bias and discrimination. This can be the case even if direct use of information about group membership is excluded from the algorithm's input; for instance, home addresses can serve as a proxy for other characteristics that lead to bias. Careful planning is needed to avoid the introduction of bias as an unintended consequence of a new technology. Training data and other inputs that contribute to the shaping of artificial intelligence, algorithmic decisions or machine learning, must be carefully selected and evaluated in order to make sure that discrimination or other undesired effects are not introduced. Automated decision-making that impacts on individuals should only be used when there is reason to be confident that the algorithm does not discriminate against any group of individuals.

## 3. Transparency and explainability

Individuals have the right to know the grounds of decisions affecting them. In many cases they also have the right to appeal to a decision-maker at a higher level, and that right can only be efficiently exerted if the grounds for the original decision are available. Individuals who are affected by a public decision based on automated data processing should have access to clear information that a layperson can understand both on how the decision was made (transparency), and on its justification (explainability). Information must also be presented in an accessible way for disabled and elderly people. Public decisions affecting an individual should be based on criteria that are relevant for the decision. Therefore, the use of erratic and unpredictable decision support systems is not acceptable.

## 4. Responsibility and accountability

The public sector is subject to strict principles of accountability. Ultimately, elected representatives are accountable to the citizens for the activities of public authorities and other entities under their direction. This means that blaming a machine, an algorithm or a decision support system is a particularly poor option in the public sector. In the end, humans will be held responsible for decisions that have effects on individuals or society. There must therefore always be sufficient human oversight and control of automatic decision-making to ensure that human decision-makers can be held accountable. Public servants who oversee work that involves disruptive technologies must have the requisite time and resources to actually fill the responsibilities assigned to them. Adequate procedures must be in place to ensure that individuals who have questions or complaints about a decision can communicate with a responsible person. Public sector organizations should also have procedures for investigating and taking measures against problems arising in their use of disruptive technologies.

## 5. Safety and security

Risks of both intentional and unintentional harm should be carefully evaluated before the introduction of disruptive technologies, and high priority should be assigned to the safety and security of all who are affected. For instance, automation can be harmful to both physical and mental health. Both technological and organizational measures are needed to prevent this from happening. Some disruptive technologies give rise to specific security risks that must be attended to. Adequate protection against hacking and other forms of adversarial intrusions and attacks must be implemented. It is also imperative to prevent the manipulation and misuse of data. Any use of sources that accept data without credible authentication of the source must include a careful investigation of the ways in which the data could have been manipulated. Whenever possible, only authenticated data should be used. This is necessary to uphold the integrity of public decision-making.

## 6. Privacy

Public administrations have access to large quantities of data, collected for various purposes. The use of such data can have unintended harmful consequences for individuals, not least when information from many sources is combined. The use of such data should be strictly regulated, and efficient measures to prevent and discover unauthorized use should be implemented. In particular, the combination of data from different sources should be subject to strict regulations, and protective measures such as informed consent and anonymization should be implemented to avoid misuse. Best-practice methods for data protection, including deletion mechanisms, should be used, and these methods should be regularly updated. The collection and use of new types of personal sensitive data, such as biological data and data from face recognition technology, should only be considered after broad public consultations.

## 7. Building an ethical culture involving the employees

The employees are the most important assets of public administration. Their participation in the creation of an ethical culture at the workplace is essential for the functionality and credibility of the public sector. New technologies should not be used to introduce more invasive forms of control of workplace behavior, since this can infringe on privacy and be dehumanizing, demoralizing, and destructive to the mental and physical health of workers. Instead, new technologies can and should be used to relieve public servants of routine work and make better use of their competences. If a new technology has impact on the employees and their working conditions, then its introduction must be decided in a participative and co-creative process involving them and their organizations. Public servants should be offered the education and training needed to improve their skills in the ethical use and management of new technologies, with the twofold objective of upskilling and reskilling. They should also be made aware of their rights in relation to their own data, if these may be affected by their employer's use of disruptive technologies.

## 8. Retaining human contacts

Ensuring the wellbeing of all residents must be a priority of the public sector, also in the adoption of new technologies. A common concern with new technologies is that lack of human agency can reduce the social contacts that are crucial for the functioning and the social cohesion of our societies. Contacts with providers of public services, especially social and health-related services, can be important and often indispensable parts of a person's social network. New technologies can and should be used to restructure the provision of public services in ways that improve its contacts with residents. New technologies should not be used in ways that weaken social networks and make residents more isolated. The introduction of alternative human-based modes of contact can in some cases be useful to avoid such negative effects.

## 9. Ethical public-private cooperation

Public-private cooperation is needed to solve many problems in the public sector, not least the development and adaptation of new technologies. The public sector should have enough competences of its own and ability to control and review private sector involvement to ensure that such cooperation works efficiently in the public interest. All public-private cooperation should comply with the ethical principles adopted for the public sector. Private sector performance and accountability should be ensured through appropriate contractual protections.

In its supervision of private sector activities, the public sector should employ adequate technology, for instance for the automatic analysis of large amounts of data, in ways that prevent infringements on privacy. Artificial decision support systems in the private sector that are subject to public supervision must be sufficiently transparent to make sure that efficient supervision can be performed.

## 10. Continuous evaluation and improvement

The social effects of disruptive technologies are difficult, often impossible, to foresee. Due to this uncertainty, the full-scale introduction of such technologies should whenever possible be preceded by carefully evaluated trials that are performed under realistic conditions. Subsequently, when the technology is introduced on a larger scale, evaluation should continue in order to detect (positive and negative) effects that may not have been discovered in the small-scale trials. These evaluations should include the perspectives of residents and employees, particularly groups at risk of exclusion or discrimination from technologies. Research on the social effects of new technologies should also be furthered. Adjustments, improvements and when necessary, replacement of technologies should be made whenever needed.

# 3. Appendix: Some useful texts on ethical principles for disruptive technologies

Article 29 Data Protection Working Group, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, WP251rev.01 (2018)

Article 29 Data Protection Working Group, Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01)

Berryhill, Jamie, Kévin Kok Heang, Rob Clogher, and Keegan McBride (2019) "Hello, World: Artificial intelligence and its use in the public sector." OECD Working Papers on Public Governance No. 36. https://www.oecd.org/governance/innovative-government/working-paper-hello-world-artificial-intelligence-and-its-use-in-the-public-sector.htm

Bird E., Fox-Skelly J., Jenner N., Larbey R., Weitkamp E. and Winfield A., (11 March 2020) "The ethics of artificial intelligence: Issues and initiatives" Panel for the Future of Science and Technology, European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA), ISBN: 978-92-846-5799-5 , Brussels. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf

Brkan, M. (2019). Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond. International journal of law and information technology, 27(2), 91-121.

Brkan, M., & Bonnet, G. (2020). Legal and technical feasibility of the GDPR's quest for explanation of algorithmic decisions: of black boxes, white boxes and Fata Morganas. European Journal of Risk Regulation, 11(1), 18-50.

Central Digital and Data Office and Office for Artificial Intelligence of the UK Government, 18 October 2019, "A guide to using artificial intelligence in the public sector - Guidance on building and using artificial intelligence in the public sector." https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector

Engstrom, David Freeman, Daniel E. Ho, Catherine M. Sharkey, and Mariano-Florentino Cuéllar (2020) "Government by algorithm: Artificial intelligence in federal administrative agencies." Administrative Conference of the United States.
https://www.sipotra.it/wp-content/uploads/2020/10/Government-by-Algorithm-Artificial-Intelligence-in-Federal-Administrative-Agencies.pdf

Floridi, Luciano, et al AI4People's Ethical Framework for a Good AI Society. https://www.eismd.eu/wp-content/uploads/2019/11/AI4People's-Ethical-Framework-for-a-Good-AI-Society_compressed.pdf

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". AI magazine, 38(3), 50-57.
file:///C:/Users/babbi/Downloads/2741-Article%20Text-5397-1-10-20170921%20(1).pdf

High-Level Expert Group on Artificial Intelligence (2019) "Ethics Guidelines for trustworthy AI." European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Kaminski, M. E. (2019). The right to explanation, explained. Berkeley Tech. LJ, 34, 189. https://core.ac.uk/download/pdf/222814754.pdf

Koops, B. J., Di Carlo, A., Nocco, L., Casamassima, V., & Stradella, E. (2013). Robotic technologies and fundamental rights: Robotics challenging the European constitutional framework. International Journal of Technoethics (IJT), 4(2), 15-35.

Leitner, Christine, and Christian M. Stiefmueller. "Disruptive technologies and the public sector: the changing dynamics of governance." A. Baimenov, P. Liverakos (eds.)Public Service Excellence in the 21st Century. Palgrave Macmillan, Singapore, 2019. 237-274.]

Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute. https://doi.org/10.5281/zenodo.3240529

Madiega, Tambiama (2019) "EU guidelines on ethics in artificial intelligence: Context and implementation." European Parliamentary Research Service.
https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf

Misuraca, G., and van Noordt, C.., Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19540-5, doi:10.2760/039619, JRC120399
https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120399/jrc120399_misuraca-ai-watch_public-services_30062020_def.pdf

OECD, AI group of experts (AIGO), 2019, "OECD AI Principles overview". https://oecd.ai/ai-principles

Renda A., (February 2019), "Artificial Intelligence Ethics, governance and policy challenges - Report of a CEPS Task Force", CEPS, ISBN 978-94-6138-716-5, Brussels. https://www.ceps.eu/wp-content/uploads/2019/02/AI_TFR.pdf

Ronzhyn, Alexander, Maria A. Wimmer, Vera Spitzer, Gabriela Viale Pereira, and Charalampos Alexopoulos. "Using disruptive technologies in government: Identification of research and training needs." In International Conference on Electronic Government, pp. 276-287. Springer, Cham, 2019.

Rulebook Workgroup in Sitra's IHAN program, "Rulebook for a Fair Data Economy - Rulebook Template for Data Networks", Version 1.2 en. https://media.sitra.fi/2020/06/11123656/rulebook-for-a-fair-data-economy.pdf

The World Economic Forum, December 2020, "Ethics by Design: An organizational approach to responsible use of technology" http://www3.weforum.org/docs/WEF_Ethics_by_Design_2020.pdf

Ubaldi, B., et al. (2019), "State of the art in the use of emerging technologies in the public sector", OECD Working Papers on Public Governance, No. 31, OECD Publishing, Paris. https://www.oecd-

ilibrary.org/docserver/932780bc-en.pdf?expires=1616170024&id=id&accname=guest&checksum=0E595EDF852F41299C2A72C1D0D064D8

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. International Data Privacy Law, 7(2), 76-99.

Winfield, Alan (2019) "Ethical Standards in Robotics and AI." Nature Electronics 2, 46-48. https://www.nature.com/articles/s41928-019-0213-6