

经合组织数字政府研究

利用人工智能于社会保障

用例、治理和劳动力准备



利用人工智能于社会保障

用例、治理和劳动力准备情况

这项工作于2025年11月29日经公共治理委员会批准解密。

本文件得到了欧盟的资金援助。此处所表达的看法无论如何都不能被视为反映欧盟的官方意见。

本文件，以及其中包含的任何数据和地图，均不妨碍任何领土的地位或主权，也不妨碍国际边界和界限的划定，也不妨碍任何领土、城市或地区的名称。

请引用此出版物：

OECD智能在社会保障中的应用：用例、治理和劳动力准备 ,经合组织数字政府研究,经合组织出版,巴黎, <https://doi.org/10.1787/b52405c1-en> .

ISBN 978-92-64-83504-7 (印刷版)
ISBN 978-92-64-69427-9 (PDF)
ISBN 978-92-64-63541-8 (HTML)

经合组织数字政府研究 ISSN 2413-1954 (印刷) ISSN 2413-1962 (在线)

照片版权：封面 © shurkin_son/Shutterstock.com

关于经合组织出版物的勘误表可在以下网址找到： <https://www.oecd.org/en/publications/support/corrigenda.html> .
© 经合组织 2025



知识共享署名 4.0 国际许可协议 (CC BY 4.0)

本工作根据知识共享署名 4.0 国际许可协议提供。通过使用本工作，您接受受该许可协议条款的约束 (<https://creativecommons.org/licenses/by/4.0/>).

归因 —你必须引用该作品。

翻译 —你必须引用原始作品，识别对原始作品的更改，并添加以下文本： 如果原作与译文之间存在任何差异，则应以原作文本为准。

适应 —你必须引用原始作品并添加以下文本： 这是经济合作与发展组织 (OECD) 原作的一个改编版本。本改编版本中表达的观点和使用的论点不应被视为代表经济合作与发展组织或其成员国官方观点。

第三方材料 —许可证不适用于作品中第三方材料。如使用此类材料，您需负责获得第三方许可并承担任何侵权索赔。未经明确许可，您不得使用经合组织标志、视觉识别或封面图片，也不得暗示经合组织认可您使用该作品。根据本许可证产生的任何争议应按照仲裁法院 (PCA) 仲裁规则2012进行仲裁。仲裁地点为巴黎 (法国)。仲裁员人数为一名。

前言

人工智能（AI）系统可用于改善社会保障福利的获取，确保在恰当的时间将恰当的福利发放给恰当的人。采用AI技术也可能补充更广泛公共部门为提高生产力、提供更以用户为中心和主动的服务所做的努力。

为了探索如何利用人工智能达到此目的，法国的主要社会保险机构（包括 *国家家庭补助金箱* – CNAF，*国家医疗保险基金* – 中国移动，和 *农业互助社* – MSA）和意大利（the *国家社会保障研究所* – 国际劳工组织（INPS））已参与一项合作项目，该项目由欧盟通过技术支持工具资助，由经济合作与发展组织（OECD）在欧洲委员会的合作下实施。

承担此项目的机构目前正在测试人工智能在各种功能中的应用，例如增强内部知识管理、改善与现有和潜在受益人的沟通以及支持数据分析，同时鉴于欧盟人工智能法案对许多用例的高风险认定，仍保持谨慎。

本报告旨在基于经合组织先前开展的差距分析报告中所识别的差距和挑战，提供一套精心策划的国际良好实践，以指导未来的人工智能应用。报告还记录了与国家级公共服务环境相关的具体政策杠杆、工具、实施策略和能力建设举措，尤其关注数据质量、治理和劳动力准备度等方面的问题。它旨在为社会保障部门如何从国家层面的更多样化、更有影响力和更可信赖的人工智能应用中受益并保持一致提供指导。需要进一步的努力来探索社会保障部门人工智能的有效应用，特别是在将这些实践适应社会保护体系的具体运营、法律和伦理环境方面。持续的实验、评估和跨部门合作对于确保人工智能的应用能够带来切实利益，同时保障公平、透明和公众信任至关重要。

最终，这项工作强调了利用AI作为更具包容性、高效和透明的社会保障体系战略推动者的关键机遇。要实现这一潜力，则需要嵌入健全的治理框架，投资于数据和基础设施，并构建公共部门值得信赖的AI所需的机构和人本能力。

致谢

该报告是由经济合作与发展组织（OECD）的两个部门合作的结果：公共治理司（GOV）和就业、劳动及社会事务司（ELS），在GOV司长Elsa Pilichowski和ELS司长Stefano Scarpetta的领导下。该报告由GOV的创新、数字和开放政府部门（INDIGO）在副司长Gillian Dörner的指导下、数字政府单位负责人Marco Daglio的监督下进行领导。该报告由Julian Olsen协调，由Mario Restuccia（第一章）和Julian Olsen（第二章）撰写，Seong Ju Park和Cecilia Emilsson做出了贡献。该报告与ELS的技能与未来准备部门合作产出，在部门负责人Glenda Quintini的指导下，由Magdalena Burtscher（第三章）撰写，Julie Lassébie做出了贡献。

该报告还受益于杰米·贝尔希尔（Jamie Berryhill）的审阅，以及来自INDIGO的里卡多·扎帕塔（Ricardo Zapata）和马可·贝特朗·纳瓦罗（Marco Beltrán Navarro）的输入；来自政府（GOV）的公共管理与预算部门、反腐败与政府诚信部门的阿拉娜·贝克（Alana Baker）和海伦·韦尔斯（Helene Wells），以及来自社会政策部门（ELS）的瓦莉·弗雷（Valerie Frey）和多隆·威克（Doron Wijker）的输入。该报告还获得了安德里亚·乌尔哈默（Andrea Uhrhammer）、马蒂厄·卡恩（Matthieu Cahen）、德博拉·费尔南德斯（Deborah Fernandez）、纳塔莉·科里（Natalie Corry）和詹妮弗·斯坦（Jennifer Stein）的编辑和出版支持。

如果没有欧洲委员会项目官员马西米利亚诺·桑蒂尼的支持和参与，以及为该项目做出贡献的受助机构成员，本报告将不可能完成，包括：

- INPS：瓦莱里娅·博纳沃尔塔、埃马努埃莱·科利尼、贾科莫·格拉西和安东尼奥·奇普里亚尼。
- CNAF：斯特凡·唐内和马哈利亚·斯泰潘诺夫。
- MSA/iMSA：埃马纽埃尔·梅斯里耶，朱利安·让尼，和文森特·泰里瑟。
- CNAM：Fanny Richard、Victor Bret，以及（之前）Seak-Hy Lo。

这份报告也受益于芬兰同行支持的国家研究访问，包括Kela、财政部、社会事务与卫生部、赫尔辛基市、UNA Oy和DigiFinland Oy。

最后，该项目受益于来自奥地利、澳大利亚、加拿大、爱沙尼亚、芬兰、法国、德国、挪威、西班牙、瑞典的公共机构专家进行的15次访谈，他们分享了各自在人工智能方面的方法。

该项目由欧盟通过技术支持工具资助，由经合组织在欧委会合作下实施。

目录

前言	3
致谢	4
缩写和首字母缩略词	7
执行摘要	8
1 社会保障中人工智能的有效应用	10
分析框架	11
经合组织各国的社会保护体系中使用人工智能	12
用于检查means-tested福利资格的AI：加泰罗尼亚能源贫困报告的案例	13
用于文档识别和分类的机器学习：德国联邦就业局的案例	16
人工智能简化理赔流程：芬兰社会保障机构的案例	21
跨领域见解	23
参考文献	26
注意	27
2 解放可信赖人工智能在社会保障中的潜力	28
赋能者——是否为支持人工智能的使用提供了合适的赋能者？	30
护栏——各国如何管理公共部门人工智能的风险？	38
参与度——用户在使用人工智能系统方面是如何被参与的？	43
参考文献	48
注释	49
3 建设人工智能人才队伍	50
理解人工智能对劳动力及技能的影响	51
社会保障领域与人工智能相关的技能需求 通过招聘和战略合作伙伴关系建立内部人工智能能力	55
通过培训发展人工智能相关技能	58
战略人才队伍建设	64
参考文献	69
注释	72
附件 3.A. 经合组织国家中与人工智能相关的培训示例	73

图例

- 图1.1 人工智能在社会保护中的当前和潜在应用 12
 图1.2. 基于（一个）样本招聘启事在VerBIS中提供的建议示例 18
 图2.1 经济合作与发展组织（OECD）政府可信人工智能框架 29
 图2.2. 经合组织公共服务领域数据治理模式框架 35
 图3.1. 按目标群体划分的AI训练示例图示 60

表格

- 表2.1 公共部门人工智能国家战略 31
 表 2.2. 关于采用人工智能系统的投资决策 31
 表 2.3. 用于开发和部署公共部门人工智能系统的专项资金 32
 表2.4. 支持公共机构在人工智能采购方面的支持 33
 表2.5. 数字基础设施的关键组件正在被用于支持人工智能的集成 34
 表2.6. 管理人工智能工具使用的护栏措施 38
 表 2.7. 参与公共部门 AI 政策和用例开发的参与者 44
 表3.1. 目标群体相关的AI技能需求 54

- 附件 表 3.A.1 公共行政中人工智能培训示例 73
 附件 表格 3.A.2. 公共行政中人工智能训练导致学位或资格的实例 74

缩写和首字母缩略词

ADEST	自动化数据提取和结构化工具
人工智能	人工智能。在本出版物中，人工智能系统的定义将如下所示：“人工智能系统是一个基于机器的系统，用于明确或隐含目标，从它接收的输入中推断出如何生成诸如预测、内容、推荐或可影响物理或虚拟环境。不同的AI系统在自主性水平上有所不同。部署后的适应性”。这个定义涵盖了各种类型的AI，其中一些已被政府使用。 经合组织框架 人工智能系统分类 提供了一套用于表征应用程序的工具 在特定环境下部署的人工智能系统。
AOC	加泰罗尼亚开放管理联盟 , 开放管理联盟 加泰罗尼亚
API	应用程序接口
BA	联邦劳动局 德国联邦就业局
BPM	业务流程管理
CNAF	国家家庭补助金箱 法国家庭国家基金 津贴
CNAM	国家医疗保险基金 , 法国国家健康基金会 保险
DGI	经合组织数字政府指数
欧盟	欧洲联盟
开源软件	自由和开源软件
GDPR	通用数据保护条例 ((欧盟) 2016/679号条例)
ITSV	社会保险IT服务 , 奥地利社会保险的IT服务 公司
机器学习	机器学习
OCR	光学字符识别
saas	软件即服务

执行摘要

欧盟各国的社会保障管理机构在服务与福利发放方面面临挑战，包括行政负担高、数据系统碎片化以及资格认定流程复杂等问题。养老金、家庭津贴和残疾福利等关键福利的获取受到信息不足、行政复杂性和数据系统碎片化等因素的阻碍。这不仅影响个人福祉，还削弱了公共投资的效应，并可能降低公众对社会保护体系的信任。

当以可靠有效的方式使用时，人工智能（AI）可以帮助公共机构更好地识别需求、个性化沟通、支持后台流程以及更主动高效地提供包括社会保障在内的服务。然而，如果没有适当的治理、控制和监督，也存在AI系统造成危害的风险。因此，社会保障是欧盟AI法案中特别指出的高风险领域之一，尤其是在AI系统旨在评估社会保障资格的情况下。

因此，为了在社会保障领域实现人工智能系统的优势，需要一支人工智能准备就绪的劳动力队伍、健全的治理、适当的保护措施，以及与用户和利益相关者的实质性参与。有了坚实的基础，人工智能系统可以支持更有效和以用户为中心的公共服务。

第一章介绍了公共部门使用人工智能的三个案例研究，探讨了人工智能如何提高效率、改进准确性以及增强公民获取福利的途径。在加泰罗尼亚，一个由人工智能驱动的云平台自动化了确定能源贫困支持资格的过程，通过改进数据共享和增强福利分配给谁的不确定性。在德国，联邦就业机构使用机器学习工具对非结构化格式的职位发布进行分类，加快了职位招聘信息向潜在申请人的传播。在芬兰，国家社会保障机构Kela开发了一个本地托管的人工智能平台，用于自动化处理每年数百万份文件。这些文件是客户附加在各类福利要求上的，现在可以自动分类和处理，从而节省成本、时间和环境。

在社会保障领域释放人工智能的潜力需要围绕三个关键支柱构建强有力的治理 *经合组织政府可信人工智能框架* 赋能因素、监管框架和参与。第二章探讨了国家层面的治理框架和人工智能战略，以及社会保障机构如何借鉴这些努力，确保社会保障领域的人工智能应用有效、可信赖且以人为本。需要关键赋能因素，如基础设施、数据质量和战略投资方法，以促进中央政府和社会保障机构在公共部门使用人工智能系统方面更协调的协作。监管框架，包括伦理监督和遵守欧盟人工智能法案，正在出现但必须在人工智能生命周期中进一步嵌入，以确保安全性和公平性。这在高风险的社会保障应用中尤其相关。德国联邦就业局或芬兰的Kela等机构已经制定了内部指南，但在大多数情况下它们仍然是自愿的。最后，与设计、开发、测试和监督人工智能系统的不同利益相关者进行互动至关重要，以确保它们满足用户需求，没有（有害的）非预期后果，并有助于建立公众信任。虽然与公务员的互动正在改善，但服务使用者、民间社会和

跨境行为者仍然有限——阻碍信任和有效性。需要进一步工作来调整这些方法以满足该行业的特定需求，确保人工智能集成的道德、有效和包容。

最后，在第三章中，报告强调，社会保障机构以及更广泛意义上的公共行政领域采用人工智能，也需要能够开发、实施和使用人工智能工具的劳动力。人工智能的采用将对公共行政领域的劳动力技能需求产生影响，增强人类任务并转变组织流程。因此，劳动力发展——涵盖招聘、外包、培训和组织文化——应成为任何人工智能机构战略的有机组成部分。政府可以使用不同的工具来提高对数字和数据分析人才的招聘和留住，并为不同的人员类型提供量身定制的培训，从普通员工到数字和数据分析专业人员及领导者。外包是解决人工智能技能差距的一种常见做法，然而，建立内部人工智能能力具有一系列优势，包括问责制、数据隐私以及社会保障机构的公共服务使命。培养强烈的创新和学习文化是支持公共部门有效且可信赖的人工智能采用的关键。

认识到人工智能在社会保障方面的益处，不仅需要部署新的AI工具，更需要一种全面的价值驱动方法，这种方法投资于人力、流程和公共信任。通过聚焦于强大的赋能因素、伦理的护栏和有意义的参与，政府可以借助人工智能来现代化服务交付、缩小获取差距，并确保社会保障体系具有韧性、包容性，并为未来做好准备。

1 人工智能在社会保险中的有效应用

本章探讨了人工智能如何应对欧盟各国的社会保障领域的持续挑战。以加泰罗尼亚、德国和芬兰为例，它突出了人工智能工具在支持资格检查、对非结构化数据进行分类以及自动化文档处理方面的应用。这些案例展示了人工智能提高效率、减轻员工和公民负担以及增强福利和服务获取潜力的可能性。然而，它们也揭示了实现这一目标需要关键的基础条件：稳固的数字基础设施、数据互操作性以及清晰的治理框架。

经合组织各国的社会保障机构面临运营挑战，限制其提供社会保障的能力：数据管理碎片化、互操作性低、对弱势群体人群覆盖有限，以及人力和技能方面能力不足（经合组织，2024）^[1]。这些问题因人口变化、经济限制和环境变化等广泛趋势而加剧（经合组织，2024）。^[2]。

在这个背景下，人工智能（AI）已成为支持公共行政在服务设计和交付方面的一种潜在工具。当在更广泛战略中精心实施时，AI系统可以帮助自动化任务，简化流程，减轻行政负担，并改进服务交付（经合组织，2025）^[3]；经济合作与发展组织，2025^[4]。

同时，公共部门使用人工智能也带来了潜在风险，需要建立一个完善的治理框架来确保可信的实施，特别是在社会保障领域（经合组织，2025）。^[4]。这些风险包括个人数据处理的失误、自动化决策缺乏透明度、算法偏见以及强化现有结构性不平等的潜在性。应对这些挑战需要构建可信赖的人工智能采用所需要的基础设施，实施管理风险的护栏，并让利益相关者参与到广泛的观点中。这些治理机制将在第2章中详细阐述。

分析框架

方法论和用例

本章探讨了如何利用人工智能解决社会保护领域的挑战，通过欧洲多个公共机构提供的三个应用案例进行分析。该分析基于文献研究、机构提供的内部文件，以及通过访谈和非正式讨论收集的机构利益相关者的直接意见。

这些访谈由以下支柱指导 *经合组织政府可信人工智能框架*（OECD，2025^[3]），它为评估用例提供了一个结构。这个框架将在第2章中更详细地应用。

用例已根据以下标准进行选择：

- 机构与地域多样性，重点是欧盟成员国（尽可能），鉴于本报告由欧盟通过技术支持工具资助。
- 实施成熟度，优先考虑那些已在生产中且能够评估其影响的人工智能系统。
- 可重复性，确保解决方案可被其他公共部门组织轻松实施。
- 可信度，选择嵌入治理流程中的人工智能实施。
- 问题-解决方案匹配，优先考虑将人工智能作为最佳技术来应对挑战的情况，而不是作为默认选择或实验性选择。

在社会保障领域，成熟的、文档齐全的AI应用案例有限，这限制了选择。然而，所选案例反映了多样的制度环境，可以为其他管理机构提供宝贵的见解。

分析维度

本章对用例的分析借鉴于 *经合组织政府可信人工智能框架* (OECD, 2025^[3]) 和 *经合组织人工智能系统分类框架* (OECD, 2022^[5])。每种用例都从五个维度进行评估：

- **问题定义和制度差距**：人工智能系统旨在解决什么挑战？在能力或服务交付方面，正在解决哪些差距？
- **AI系统描述**：使用何种系统？其功能是什么（例如分类、预测、自动化）？根据经合组织的定义，它如何符合AI系统的标准？
- **实现条件**：哪些政治、立法、技术和组织因素促成了解决方案的部署？
- **结果与关键经验**：取得了哪些成果？组织层面的结果是什么？
- **挑战和局限性**：人工智能解决方案实施后的风险是什么？人工智能实施过程中遇到了哪些挑战？
- **跨行政区域可移植性**：我们能从其他社会保障机构或其他公共管理机构中学到什么教训？解决方案适用于另一种环境的前提条件是什么？

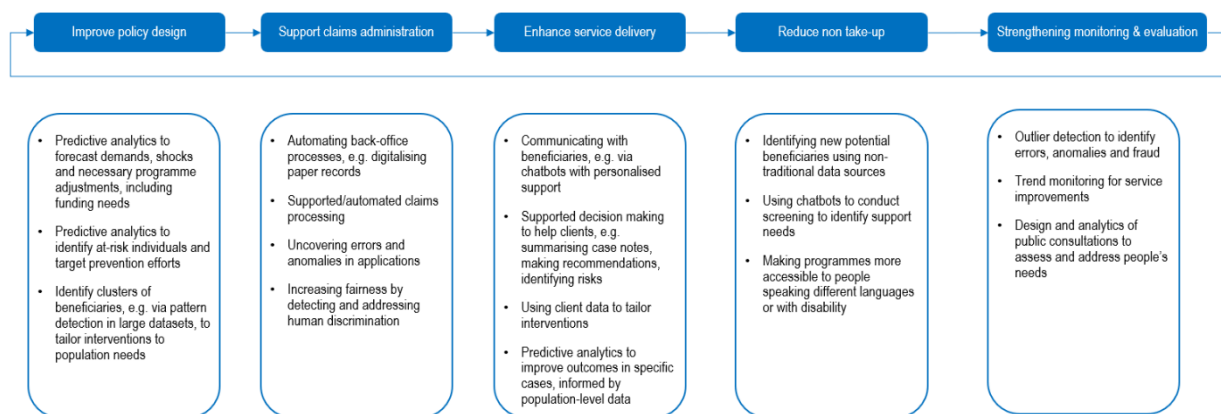
本章提出的应用案例也指导了后面章节的分析。特别是，第2章探讨了人工智能治理框架，第3章探讨了劳动力影响。

经合组织各国的社会保障体系中使用人工智能

人工智能在提升社会保障的覆盖面、有效性及效率方面具有巨大潜力。其潜在应用领域十分广泛：除了减少社会保障计划的未被领取，人工智能还有助于改进政策设计、索赔管理、服务提供以及监控与评估（经合组织，2025）^[4]（见图1.1）。

图1.1 人工智能在社会保障中的当前和潜在应用

人工智能和社会保障的当前及未来可能用例示意图



注意：经合组织关于人工智能在社会保障方面的现状和假设性应用的概述。来源：（经合组织，2025^[4]）。

人工智能已经用于提升客户支持、自动化后台流程和支持欺诈检测。然而，人工智能更高级的应用——例如在预测分析中——尚未普及，并且在经合组织国家中尚未普遍使用人工智能来减少未采用情况（经合组织，2025）。^[4]（参见框1.1）。

一、1. 人工智能与社会保障的未来：人工智能在社会保障体系中的当前和未来应用

经合组织对经合组织各国的社会保护系统中人工智能的当前和潜在用途的评估显示，当今人工智能在社会项目中的应用主要集中于客户支持、后台流程自动化和欺诈检测。一些国家已经部署了人工智能驱动的聊天机器人，以向公民提供信息和支持。像奥地利、加拿大、芬兰和美国这样的国家已经使用人工智能来自动化和简化后台流程，例如通过处理和分类文件、分发和引导收到的邮件和电话，和/或通过预测索赔重新计算可能对索赔人产生影响的地方来优先处理索赔。

虽然目前在经合组织国家尚未普遍，但人工智能在社会保护方面最具潜在回报的应用领域在于预测分析，包括改进覆盖面、减少社会项目中的未领取情况，以及预测需求、支持预防、早期识别并针对客户需求定制干预措施。这类人工智能应用实例已存在，研究人员和政府已试点（且程度较低地实施）人工智能工具来预测无家可归、家庭暴力再犯和工作场所安全事故等不利事件的风险，从而促进预防和早期干预。

政府正在谨慎行事，在识别出适合人工智能解决的政策挑战时，建立监管机制并试点小规模项目，以确保人工智能应用能够安全且规模化地推广。经合组织研究表明，这种谨慎态度也反映了公众的看法，公众对人工智能应用的支持态度冷淡：在参与经合组织“重要风险调查”的27个国家中，不到一半的受访者（40%）认为使用人工智能来处理 and 批准社会项目申请对用户有益。需要更大的公众参与和强有力的治理来建立对人工智能在社会项目中的责任的信任。

来源：（经济合作与发展组织，2025[4]）。

用于检查means-tested福利资格的AI：加泰罗尼亚能源贫困报告的案例

在西班牙的加泰罗尼亚地区，人工智能已成功应用于通过自动化生成低保金资格报告来帮助解决能源贫困问题。这一用例突显了中央数字化基础设施供应商在协调地区内社会保障福利提供方、改进数据共享以及通过自动化帮助减少福利空置方面的作用。

问题定义与制度性差距

能源贫困是加泰罗尼亚的一个紧迫问题，不断上涨的能源成本加剧了这一问题。2015年的一项地区性法律要求能源供应商符合特定资格标准（如低收入或脆弱的生活环境）的无力偿还客户维持最低水平的供应（加泰罗尼亚，2015）^[6]。

这项措施的实施面临着管理收入调查型社会保障的一个普遍挑战：识别符合条件的家庭的困难，因为收入、资产或家庭构成的数据通常分散在不同的系统中，并且当局之间并不总是共享（Pye等人，2017年）^[7]。

评估家庭继续获得供应的资格并为最终实施该措施而进行的程序涉及多个参与方之间的数据交换，以及家庭采取行动同意数据交换和受益接受。当家庭停止支付账单时，能源供应商通知市政社会服务机构，然后该机构在需要时请求额外的同意来收集和处理社会经济数据。

经过人工核实，如果认为该家庭有社会排斥风险（即停止援助），市政府将通知供应商必须维持基本能源供应。此过程包括五次数据交换、三个同意步骤和人工评估，导致高未使用率。

市政当局作为社会保障福利提供者，负责确定家庭是否符合维持能源供应的资格标准，但必要的的数据分散在多个部门，且核实流程复杂且劳动密集。这种复杂性，加上与能源供应商共享信息时严格的隐私要求，给行政带来了巨大的负担。

尽管2015年的法律要求引入数据共享协议，但实施一直面临技术和法律障碍，包括数据保护和社会保障法中明确同意的要求，这适用于拖欠能源账单的家庭。这些限制导致该措施的潜在受益人非采纳率很高。所研究的AI系统正介入资格评估阶段，通过自动化数据收集和验证来减少行政负担，并最终改善福利获取。

AI系统描述

为减少资格评估程序的行政复杂性，加泰罗尼亚开放行政联盟（AOC），一个为该地区提供数字公共基础设施的公共联盟，开发了一个自动化生成能源贫困报告的AI系统。该系统以软件即服务（SaaS）的形式提供给市镇，旨在支持社会工作者核实2015年法律引入的基本能源供应福利的资格（AOC联盟，2023）。^[8]

该系统运行在三个主要平台：*赫斯提亚*，该系统作为地方社会服务信息系统；*奥布塔*负责通过互操作性目录在国家级、区域级和地方实体之间进行安全数据交换；以及一个业务流程管理（BPM）解决方案，用于简化和自动化运营。每月，能源供应商都会识别出现财务困境的客户，并将相关信息传输给市政部门。公务员随后将数据上传到AOC平台，该平台会检索家庭所需的社会经济参数，验证资格，并生成一个预填好的PDF格式报告，再将其发送回能源公司。

关键组件是一个确定性人工智能模型，该模型遵循预定义的一组规则来评估资格。与通过推理提供不确定输出的概率模型不同，确定性模型在给出相同输入时每次都产生相同的结果。这确保了决策逻辑是透明和可重复的，允许受系统影响的人根据需要进行理解和挑战结果（经济合作与发展组织，2022）。^[9] 根据经合组织定义，该系统属于人工智能，因为它基于机器，具有明确的目标（资格验证），并从多个数据输入中推断输出（资格报告），以影响社会保护环境中的决策。

通过自动化数据收集、验证、计算和报告生成，AI系统已经接管了以前由市政工作人员手动执行的任务，通常涉及跨多个数据库的重复检查和手动报告起草。这大大减少了行政工作量和人为错误的风险。虽然由于与受益人同意相关的法律限制，完全自动化目前尚不可行，但该系统代表了对向所有符合条件的居民主动提供福利的第一步。

实现条件

人工智能系统的成功部署得益于一系列使能因素，包括法律框架、制度环境以及现有的技术基础设施。

2015年区域法律为能源供应商与公共管理部门之间的数据共享奠定了法律基础，创建了人工智能系统的应用场景。然而，欧盟通用数据保护条例（GDPR）和加泰罗尼亚社会服务法要求在多个环节获得家庭住户的同意。首先，市政府需要授权生成能源贫困报告，因为将其视为一项社会福利申请。其次，通过访问外部数据来核实申请人社会和经济状况无需获得（默示同意）。最后，根据GDPR，需要获得其他家庭成员的明确同意，才能访问其税务机构收入数据及其他健康敏感信息。这些要求阻止了完全自动化，并让用户参与到流程中（加泰罗尼亚，2017）^[9]。

制度环境和利益相关者的协调是主要推动力，也允许共享IT供应。AOC作为该地区的数字基础设施提供者，能够集中投资，并提供合适的数字技能和人才。AOC提供的免费平台（Via Oberta、Hestia和业务流程管理）使地方政府能够使用安全的共享数据系统，降低了维护和支持成本。

结果与关键经验

自动化能源贫困报告的人工智能系统的主要成果是，通过简化流程，减少了福利未领取现象，并持续保障能源供应。在自动化之前，据估计超过一半符合条件家庭因程序复杂性和提供文件的负担而未能获得支持（Interoperable Europe YouTube频道，2023年）。^[10] 新系统减少了受益人所需的步骤数量，释放了公务员处理更多索赔的能力，使弱势家庭更容易获得支持。

该系统的创新之处在于其通过互操作性平台整合零散数据源，并规模化自动化多步骤资格认定流程的能力。这代表着向主动福利输送迈出了重要一步，并为与其他人工智能模型进行进一步技术演进奠定了基础。

区域标准化是另一项关键成果。通过SaaS提供该系统可确保所有市镇使用相同的资格标准和数据来源，从而减少实施中的差异。这在AOC指出的情况下尤为重要，即由于法规不明确和缺乏指导，市镇以前以不同的方式适用该法律。

挑战与局限性

全面自动化的主要障碍，进而也是符合条件的家庭全面覆盖的主要障碍，是法律方面的。现行立法要求受益人在市政当局能够获取必要数据之前，必须明确同意，有时则需要所有家庭成员同意。为减少流程中的摩擦，AOC已提议转向假定同意模式，并允许选择退出，该模式仅适用于涉及基本社会福利的情况，其目标是在保持的同时扩大覆盖范围。

反对数据处理或福利提供权，根据数据保护和社会服务法律。

与数据质量、标准化和共享相关的技术挑战仍然是一个问题。能源供应商通过电子表格向市政府提交破产数据，并未遵循特定格式。因此，该系统已被设计为能够处理所有可用的各种格式。AOC已提出一个安全的数据交换界面，但所有能源公司都采用它将面临挑战。

人工智能解决方案的实施在透明度方面也存在风险，尤其是对于弱势家庭而言。虽然确定性人工智能模型比概率性模型更容易解释，但数据流和法规要求的复杂性仍然可能使结果难以理解。AOC正在开发算法透明度报告，并探索数字同意机制来解决这些问题（AOC，无日期）^[11]，但公众对自动化流程的认识仍然有限。

跨行政区域可移植性

加泰罗尼亚案例为实行分散化、碎片化社会保障福利提供的社会提供了一些启示，展示了中央协调如何在人工智能实施中成为关键推动力。AOC作为加泰罗尼亚共享的数字基础设施提供者的角色，促成了单个市镇无法独立实施的解决方案的发展。该模式特别适用于社会保障提供者在规模、能力和数字成熟度上存在差异，或多个行为者在福利和服务提供的不同阶段进行干预的系统。

该案例还展示了标准化和数据交换基础设施对于支持人工智能实施的重要性。数据交换平台（Via Oberta）、社会信息系统（Hestia）以及BPM解决方案的使用，使得利用国家数据共享协议构建一个全面且可互操作的系统成为可能。虽然许多国家拥有某种形式的数据互操作基础设施，但加泰罗尼亚的经验突出了跨机构实时、安全访问数据的附加值。

最终，法律框架在塑造公共部门人工智能系统的有效性方面发挥着作用。卡塔卢尼亚目前的明确同意要求限制了系统完全自动化和更广泛覆盖的潜力。AOC提出的转向假定同意模式，并提供便捷的退出选项，仅适用于涉及基本社会保障的情况，可以为其他寻求平衡效率与隐私和伦理标准的国家提供参考。此外，AOC还发布了一份详尽的算法透明度报告。¹ 旨在确保可解释性和可理解性，提高问责制和监督，同时尽量降低风险，建立信任。

用于文档识别和分类的机器学习：德国联邦就业局的案例

人工智能在公共就业服务（PES）中的采用已变得普遍，许多机构已将其整合到核心功能中，如工作匹配和画像（Brioscú等人，2024）^[12]（参见框1.2）。在德国，联邦就业局（联邦劳动局，BA）已部署一个人工智能系统来支持雇主提交的工作岗位广告的分类。该用例说明了机器学习如何应用于通过自动化非结构化数据处理来提高劳动力市场服务的效率。该系统减少了BA员工的手工工作量，同时保留了人工监督，并展示了人工智能如何被集成到现有的工作流程中，以在不增加外部利益相关者负担的情况下改进服务提供。

框1.2. 人工智能在公共就业服务中的应用

在经合组织国家中，有一半的公共就业服务 (PES) 正在利用人工智能来提升对求职者和雇主的政务服务。最广泛的应用是职位匹配，其中利用自然语言处理和机器学习，根据能力来比较求职者资料与职位空缺。这使得公共就业服务能够考虑到可转移技能，并提供更准确和个性化的推荐。

算法系统也被用来支持个性化和主动指导。画像工具利用机器学习，在其注册过程的早期通过分析就业历史和教育等变量来识别有长期失业风险的人。这使PES能够优先配置资源并提供定制化服务，例如培训或咨询。

然而，将人工智能应用于职业服务领域带来了严峻的挑战，因为这些技术对人们获取服务的途径有着至关重要的影响。如果在画像和匹配工具中存在偏见，并且算法复制了历史劳动力市场数据中的模式或依赖地理位置、教育程度等代理变量，那么这种偏见可能会加剧现有的不平等。这可能导致对求职者做出不公平的推荐或错误的分类，特别是那些来自弱势群体的求职者。为了降低这些风险，职业服务系统需要进行系统的偏见测试、采用透明的模型，并有人工监督，以确保自动化决策保持公平。

来源：(Brioscú等人，2024[13])

问题定义和制度差距

BA每年接收超过10万个招聘广告，这些广告由雇主以PDF附件或通过电子邮件链接到外部网站等不同格式发送。这些提交在结构和内容上差异很大，难以一致地处理它们。要在BA的内部招聘系统VerBIS中使用，每个广告都必须根据预定义的一组类别进行标准化，包括职位名称、合同类型、所需资格和薪酬等级。

以往，BA员工通过人工进行此分类，平均每个广告耗时超过10分钟。虽然越来越需要雇主以结构化的方式提交招聘广告，这在一定程度上解决了问题，但BA仍保留电子邮件渠道的选项，以降低减少共享职位信息的风险。

AI系统描述

为了自动化分类招聘广告的过程，业务分析师开发了一个名为ADEST（自动数据提取和结构化工具）的机器学习系统，用于从非结构化数据源中提取和分类信息。该工具为VerBIS系统（业务分析师用于职业安置和咨询的内部平台）提供了标准化的招聘广告条目。

BA确保员工保留决策权。案件工作者在VerBIS上看到原始职位广告，就像收到的那样，旁边还显示了由ADEST预填的各种属性，包括每个计算属性的置信度分数（见图1.2）。员工可以接受、修改或拒绝AI的建议，这也有助于改进模型。

从技术上讲，ADEST由几个机器学习模型组成——最值得注意的是光学字符识别（OCR）——这些模型是在大约60万个历史招聘广告上训练的。它由BA的人工智能能力中心开发，并得到外部承包商的支持，并托管在

由 BA 部门的 IT 分部管理的内部基础设施。用户界面设计得直观且透明，具有置信度分数和可追溯性功能，以支持可解释性。

图1.2. 基于 (一个) 样本招聘启事在VerBIS中提供的建议示例

机器学习模型对招聘信息进行分类，并将这些分类展示给用户

Job application form

VerBIS

Position
IT tester
AI model confidence [4 bars]

Contract
Open ended
AI model confidence [3 bars]

Pay scale
Not specified
AI model confidence [3 bars]

来源：德国联邦就业局。

实现条件

ADEST的部署得益于联邦就业局内部的组织能力、技术基础设施和内部治理机制的结合。该解决方案的开发由BA人工智能能力中心领导，并在设计和测试阶段与德国各业务部门密切协调。这种合作确保了该工具能够满足具体的运营需求，并能集成到现有的工作流程中。

在开发过程中已考虑法律合规性和伦理保障措施。BA对所有AI项目应用标准化内部风险评估，根据自动化程度和潜在对用户的影响评估系统（见框1.3）。ADEST由于其人工参与设计和缺乏完全自动化决策被归类为低风险。该系统由BA数据保护办公室、AI治理部门和IT安全部门进行了审查，以确保符合GDPR、AI法案和内部标准。

adeset的实施也受益于有利的制度环境。联邦政府已表示支持在公共行政中使用人工智能，特别是作为一种解决劳动和社会福利服务劳动力短缺的手段（联邦政府，2024）^[14]。这项战略一致性有助于使人工智能计划合法化，并为创新提供了明确授权。

1.3 盒子。加强人工智能可追溯性的 BA 指导原则

为评估与人工智能部署相关的风险因素，BA制定了BA数据伦理指南，该指南包括以下原则：

人在回路中

- 我们确保我们的算法决策系统为客户（自然人和法人）以及我们的员工提供支持。
- 我们保障人类决策的自主性。
- 我们为客户和员工的利益利用数字化的机遇，并整合他们的需求。
- 我们鼓励支持我们员工的有意义的工作。

鲁棒性、安全性和可靠性

- 我们确保我们的算法决策系统的发展和运行符合当前的安全标准，并且用户可以信赖高质量的结果。
- 我们保护所有委托给BA的信息。
- 我们创建稳健的算法决策系统，以将非预期后果降至最低。
- 我们使用客户反馈来持续改进我们的算法决策系统。

隐私和数据质量

- 我们确保个人数据根据数据主体的意愿和法律的数据保护要求进行处理。
- 我们保护个人数据并维护信息自决。
- 我们确保符合数据质量标准，从而为我们的算法决策系统取得最佳可能结果做出贡献。

透明度和可解释性

- 我们确保人与算法决策系统之间的交互是透明的，并且做出的决策是用户友好的、易于理解的和可解释的。
- 我们教育我们的客户和员工关于我们算法决策系统所追求的目标。
- 我们确保在行为艺术领域内，对算法决策系统的开发和使用的透明性。
- 我们展示了算法决策系统的机遇和潜在风险。
- 我们在与算法决策系统互动以及基于概率的结果时，会通知客户和员工。

公平、包容和多元化

- 我们确保我们的算法决策系统考虑到人类的多样性，并关注公平和性别平等。
- 我们确保平等对待原则得到尊重，并且不存在歧视性的不平等对待。
- 我们确保我们的算法决策系统促进性别平等。

- 我们考虑所有相关利益相关者，并将他们纳入设计过程。
- 利用算法决策系统，我们积极促进无障碍和包容。
- 我们系统地建立反馈循环，积极捕捉用户输入，并在发展和完善我们的算法决策系统中将其考虑在内。
- 我们提升算法决策系统的数据伦理能力和敏感度，面向开发人员和员工。

社会效益与公共利益

- 我们平衡经济、环境和社会利益，以确保我们的算法决策系统用于公众利益并为社会创造价值。与公共利益相比，个人和群体的法律保障具有更高的优先级。
- 我们促进社会团结，改善参与劳动力市场和职业教育和培训的机会。
- 我们管理有限的资源以可持续的方式。
- 我们促进满足当今和未来世代需求的发展。
- 我们推广独立和循证劳动力市场研究。

责任和合法性

- 我们致力于根据我们的伦理价值观和法律要求开发和测试算法决策系统，并与独立审查员、学术界和其他组织紧密合作以实现这一点。
- 我们确保在整个算法决策系统生命周期中承担责任。
- 我们记录了与我们的算法决策系统设计和配置相关的决策。

来源：德国联邦就业局。

结果与关键经验

ADEST的引入使得在BA内部工作广告处理效率和一致性得到了可衡量的改进。据内部估计，与之前的传统人工流程相比，分类一则工作广告所需的时间已经减少了超过60%，从而释放了员工的工作能力，用于更复杂的活动。

该系统也有助于提高在VerBIS系统中记录招聘广告的方式的标准化程度。此前，各区域办事处在分类实践方面的做法存在差异，这取决于个人的解释和工作量。ADEST通过对所有传入广告应用相同的分类逻辑提供了一个一致性的基准，从而提高了用于职位匹配和劳动力市场分析的数据质量与可比性。

尽管系统已受到良好反响，但业务分析师承认内部采纳情况在各部门间存在差异。培训和变革管理举措有助于提高采用率，但需要进一步工作将工具融入日常流程并确保一致使用。业务分析师正在监控使用模式并收集反馈以改进系统，并确定可能需要额外培训的领域。

挑战和局限性

ADEST的部署显示出与透明度、可解释性和长期治理相关的一些差距。虽然用户可以评审AI生成的分类，但底层模型仍然不透明。业务分析团队已经在其内网发布了关于数据来源、性能指标和模型结果的文档，但这些材料无法被外部利益相关者访问。此外，目前还没有公开的算法系统登记册，源代码和训练数据未公开，限制了外部审查和问责。

英国人工智能办公室也制定了值得信赖的人工智能内部原则，涵盖透明度、公平性和人类监督。然而，这些原则并非具有约束力，仅适用于该机构内部。如果不与更广泛的经合组织和其他公共服务标准保持一致，那么在机构间实施人工智能治理时就有可能出现碎片化的风险，尤其是在人工智能法案下监管演变的情况下。

最后，虽然ADEST的设计旨在通过适应现有工作流程来尽量减少干扰，但这种做法也限制了进行更根本变革的范围。虽然该系统并非设计用来解决诸如招聘信息的异构性或雇主缺乏结构化数据等上游问题，但它仍然能够提高特定任务的效率。与公共部门中的许多人工智能应用一样，风险在于自动化强化了现有的低效流程，而不是彻底重新设计它们。

跨行政区域可移植性

ADEST系统展示了AI如何在不需要结构变化的情况下，通过针对一项具体且耗时的工作来有效地部署在公共管理中。这种方法对于像BA这样的机构特别相关，这些机构接收大量非结构化数据，并且难以对外部参与者强加标准化。

德国案例也突显了内部采纳和变革管理的重要性。人工智能系统需要持续投入培训与沟通以确保一致使用。这是其他社会保障机构共同面临的挑战，这些机构仍在努力提升员工的AI素养并将AI工具嵌入日常运营，如第三章所述。英国社保部门（BA）的经验表明，可复制性不仅取决于技术可行性，也取决于组织准备状态和持续的制度支持。

人工智能简化管理流程：芬兰社会保障机构的案例

芬兰国家社会保障机构Kela已开发了一个人工智能平台，以改进内部运营并减轻行政负担。该平台自动处理用户提交的文件和消息，从而实现更快、更一致的服务提供。虽然该系统不决定资格或做出福利决定，但它提高了索赔处理的效率，并支持员工管理大量数据。此用例说明了人工智能如何以低风险的方式应用于公共行政中的程序性任务，符合可信人工智能的目标。

问题定义和制度差距

科拉因每年人工处理数百万份文件和客户互动而导致运营效率低下。大多数福利申请是通过在线平台OmaKela提交的，并且在许多情况下，它们需要额外的文件，这些文件不一定由公共机构保留，例如租赁合同或银行对账单，这些文件必须手动审查和分类。同样，与客户的支持消息和电话呼叫通常包含非结构化信息，需要分类或转录，增加了工作量并延迟了处理时间。

除了这些运营挑战外，Kela必须平衡效率与透明度和监督。虽然该机构受益于高度的数字化成熟度，但许多流程需要人工参与。这与德国联邦就业局面临的 *ситуации* 类似，由于缺乏处理工作机会的工具，导致响应能力受限。在Kela的情况下，文件处理与福利发放之间的直接联系增加了程序的压力和监督。

手动流程也会影响用户体验。复杂的步骤、延迟和不一致的沟通可能会阻止弱势个体完成他们的申请，从而导致无人使用。Kela的人工智能平台是为了解决这个问题而开发的，通过自动执行常规任务同时保留人工监督来实现。该系统并非取代决策，而是支持决策，简化工作流程而不损害问责制。

AI系统描述

人工智能平台自动化处理与福利申请相关的文档的接收和处理。它执行一系列任务，如格式转换、文字识别、图像校正、条形码读取和分类。这些步骤在一个工作流中执行，并应用于多种福利类型。

除了处理附件外，该平台还包括转录客户电话和分类支持请求以转移到相应的业务单元的工具。电话转录不完全是自动化的：员工可以选择是否启用它，并根据质量从不同的语音转文本模型中进行选择。

该平台运行在由Kela完全管理的本地云基础设施上，使用模块化组件。它依赖于Tesseract（用于OCR）、Whisper（用于语音识别）和PyTorch（用于分类）等开源软件。这些工具通过Kubernetes进行容器化和编排，这意味着应用程序可以在不同环境中一致地运行，并具有自动管理功能。这允许实现可扩展性和弹性。

实现条件

科拉的AI实施战略一直基于长期以来的制度选择，这些选择优先考虑自治和主权。在私有云基础设施上使用开源技术开发和运营AI平台的决定，确保了对系统的完全控制，减少了对外部供应商的依赖，并允许随着时间的推移进行增量式适应。

开源软件的使用也支持招聘和知识共享。通过与广泛使用的技术保持一致并为开源社区做出贡献，Kela可以吸引具有相关专业知识的优秀人才，并受益于更广泛的生态系统支持。

除了技术实施之外，Kela已与其它公共机构紧密合作，制定其人工智能治理战略。国家监管机构通过发布有关负责任人工智能使用的指导方针提供了帮助。这种协作方式得到了公共行政中人工智能的强大政治支持，在明确的监管界限内创造了有利于实验的环境。

结果与关键经验

该人工智能平台为Kela的内部运营带来了可衡量的改进。2024年，它处理了超过1600万份附件，据估计节省了约38个人工年的员工时间。通过自动化常规任务，该系统减轻了人工工作量，提高了处理速度，并为与受益人进行更复杂的互动释放了能力。

效率提升也转化为基础设施节省。平台调整大小和转换文档的能力减少了240TB的存储需求，降低了运营成本，并减少了基础设施的环境足迹。

尽管系统并未直接处理福利资格或未参保问题，但它为未来改进奠定了基础。通过简化后台流程，Kela能够更好地探索更主动的服务模式，包括自动联络和实时资格检查——前提是法律和伦理保障措施仍然到位。

挑战和局限性

语言处理工具（如语音识别和消息分类）的有效性受限于芬兰语高质量训练数据的存在，以及芬兰北部所说的萨米语数据的近乎缺失。这会影响准确性，并需要工作人员保持参与以审查输出。

保持对系统的完全控制也伴随着权衡。Kela 对内部开发和本地基础设施的承诺需要持续投入技能和资源。随着组织向混合云模型转型，它需要平衡灵活性与其对主权和数据安全的重视。

跨行政区域可移植性

凯拉的案例说明了人工智能如何在社会保障机构中引入，同时又不完全自动化决策。在一个行政行为影响人们生活的领域，法律和制度框架对决策施加了严格的限制。通过专注于文档分类和信息处理，凯拉展示了人工智能如何在保障公平性和问责制这些安全措施的前提下提高效率。

在内部构建和维护平台，使用开源技术，对于技术能力强的机构来说尤其相关。这种方法减少了对外部供应商的依赖，降低了长期成本，并使该机构能够根据需求的变化来调整系统。

同时，该案例突显了制度准备的重要性。Kela的成功不仅仅依赖于基础设施：它需要清晰的策略、稳定的治理以及重视技术和运营团队之间协作的文化。这些条件在责任感下放的更碎片化系统中并不总是存在，在这些系统中，数字化转型是由采购周期驱动的，而非长期规划。

跨领域见解

本章介绍的三种用例展示了人工智能如何支持社会保障机构解决运营效率低下、改善服务交付，并为更积极的事务提供奠定基础。然而，成功的实施也取决于制度环境、法律框架和运营优先级。

常见使能因素：明确的使用案例、共享的基础设施和制度支持

所有三个案例都源于一个明确界定的问题，即在采纳人工智能的收益超过风险。在加泰罗尼亚，各市政当局能力上的异质性和数据验证的重复性任务凸显了集中式解决方案的必要性，该方案得益于能源贫困法。在德国和芬兰，自动化需求来自业务单元，而人工智能治理结构

组织层面提供了清晰的指导方针。在每个案例中，AI都不是默认选择，而是作为最合适的工具被选中。

内部技能和技术基础设施是共享的使能因素。每个机构都基于现有的数字系统：加泰罗尼亚的共享互操作性平台、德国的内部IT基础设施以及芬兰的本地的云。AOC、BA和Kela都得益于内部技能和专业知识。这些基础使得AI系统能够集成到现有的系统和业务流程中，而不是需要开发定制的解决方案，或依赖外部承包商。

最终，组织和制度支持至关重要。无论是通过AOC在加泰罗尼亚的中心作用，英国广播公司内部团队之间的协调，还是在Kela内部的技术能力，每个机构都确保人工智能发展符合更广泛的战略和政策目标。

周期性挑战：透明度、偏见和技能

在所有三种情况下，缺乏算法透明度是主要限制。虽然每个系统都设计了安全措施，例如加泰罗尼亚的确定性逻辑或德国和芬兰的人机回路验证，但所有机构都没有维护人工智能系统的公共登记册或发布代码或详细文档。这限制了问责制和外部审查，使得全面理解或挑战结果变得困难。

偏差和不公平也是关注的焦点，特别是在依赖历史数据或非结构化输入的系统。尽管所有用例都不涉及完全自动化的决策，但这三家机构都认识到，如果人工智能系统没有得到仔细的监控，就会有加剧现有不平等的风险。在没有清晰的文档或可解释性机制的情况下，这些风险更难被检测和缓解。

人工智能应用面临的另一个挑战是在所有相关语言中缺乏高质量的训练数据。诸如语音识别或消息分类的语言处理工具需要大量数据集，而芬兰语的数据集稀缺，萨米语的数据集几乎不存在，这降低了准确性并需要人工审核。在加泰罗尼亚，由于加泰罗尼亚语和西班牙语的并存，开发稳健语言模型的难度也类似，存在问题。

技能差距是一个共同限制，尽管处理方式不同。德国和芬兰投资于内部能力建设，而加泰罗尼亚则依赖一个中央数字解决方案供应商来支持市政服务交付。在所有情况下，员工培训和组织变革对于确保采纳和监督至关重要。这一主题将在第三章中进一步探讨。

关键发现

人工智能有潜力提高效率，减轻工作人员和公民的负担，并提升社会保障以及其他公共服务领域的福利和服务获取能力。人工智能工具可通过自动执行资格审核或促进文件处理等方式，帮助减少申请人和公共部门员工的行政负担，并改善服务获取情况。这体现在欧盟的三个应用案例中：

- **加泰罗尼亚：利用人工智能审核means-tested benefits的资格**

在西班牙的加泰罗尼亚地区，开发了一个AI系统，通过自动化资格评估来促进解决能源贫困问题的法律的实施。AI被用于一个基于云的平台，根据来自多个公共来源的综合数据生成能源贫困报告。该解决方案显著减轻了市政当局和公民的行政负担。然而，完全自动化受到明确同意的法律要求的限制，数据共享流程仍然依赖于私营能源公司的手动输入。加泰罗尼亚的经验突出了数据互操作性、共享基础设施以及利益相关者之间协调作用在实现可扩展的AI应用中的关键重要性。

- **德国：用于文档识别和分类的机器学习**

德国联邦就业局（BA）实施了一款机器学习工具ADEST，用于对非结构化招聘数据进行分类，以便将其整合到其中央就业系统中。该人工智能系统通过处理超过10万条每年的招聘信息，显著减少了人工工作量，向公务员展示结构化数据以及置信度分数，以供验证。BA的方法说明了人工智能如何在损害信任和监督的情况下提高服务响应性。

- **芬兰：AI将简化索赔处理**

芬兰的社会保障机构Kela运营一个内部AI平台，以自动化处理文件和支持请求。该平台每年自动化处理超过1600万份文件，支持通过其在线门户提交的申请。该系统执行格式转换、分类、OCR和文档调整大小等任务——每年据估计可节省约38人年工作量，并节省超过240 TB的存储空间。芬兰的做法表明，维护内部基础设施和FOSS工具对于保持技术主权、管理成本和建立机构专业知识具有长期价值。

从此，社会保障部门有三个关键点：

- **赋能者很重要：** 这些案例突显了在数字基础设施、数据治理和内部制度能力方面进行基础投资的重要性。集中的数字平台可以通过降低成本，加速采用并为分散的行政机构降低进入门槛。

- **护栏是必要的** 战略、有效、负责且值得信赖的人工智能部署需要人类监督、数据保护、透明度和公平性。德国的模式注重设计伦理人工智能，而芬兰的基础设施反映了战略优先事项。

- **以用户为中心的参与至关重要** 所有三个例子都强调人工智能系统必须针对行政工作流程和用户需求进行定制。这是因为有效的解决方案需要最终用户——无论是内部还是外部——积极参与设计和迭代。

总体而言，这些例子表明，人工智能在谨慎实施时，能够降低获取门槛，提高效率，并构建更具包容性和响应性的社会保障体系。

参考文献

- AOC (n.d.), *算法透明度：能源贫困报告的自动化*, [11]
<https://www.aoc.cat/en/ia-transparencia-pobresa-energetica/> (于2025年7月21日访问)。
- AOC联盟 (2023) *自动化能源贫困报告的试点，以避免供应切割*, <https://www.aoc.cat/en/blog/2023/pilot-pobresa-energetica/> (于2025年6月23日) [8]
- 布罗斯科, A 等人 (2024), “公共就业服务的新曙光：服务交付在人工智能时代”, *经合组织人工智能论文*, 经合组织出版, 第19号, 巴黎, <https://doi.org/10.1787/5dc3eb8e-en>. [12]
- 布罗斯科, A 等人 (2024), “公共就业服务的新曙光：服务交付在人工智能时代”, *经合组织人工智能论文*, 经合组织出版, 第19号, 巴黎, <https://doi.org/10.1787/5dc3eb8e-en>. [13]
- 联邦政府 (2024), *答复 20/12831 - 关于对小问题询问的答复-文件 20/12516 - 根据第二部社会保险法典的就业服务中心和工作市场促进*, <https://dserver.bundestag.de/btd/20/128/2012831.pdf> (2025年8月1日访问) [14]
- 加泰罗尼亚 (2017) *LLEI 12/2007, d'11 d'octubre, de serveis socials*, [9]
<https://portaljuridic.gencat.cat/ca/document-del-pjur/?documentId=415692> (于2025年7月21日访问)。
- 加泰罗尼亚 (2015) *LLEI 24/2015, 2015年7月29日颁布, 关于采取紧急措施应对紧急情况的法律在住房领域和能源贫困*. [6]
- 互操作性欧洲 YouTube 频道 (2023) *SEMIC 2023：可信人工智能平行论坛* [10]
用于公共部门互操作性, https://www.youtube.com/watch?v=Tg_vVoDo6wc&t=1016s (于2025年7月21日访问)。
- OECD (2025), “人工智能与OECD国家社会保护的将来”, *经合组织人工智能文献*, 第42号, 经合组织出版, 巴黎, <https://doi.org/10.1787/7b245f7e-en>. [4]
- OECD (2025), *人工智能治国：核心领域的现状与未来方向*, *政府职能*, 经合组织出版商, 巴黎, <https://doi.org/10.1787/795de142-en>. [3]
- OECD (2024), *巨型趋势与社会保障的将来*, 经合组织出版商, 巴黎, <https://doi.org/10.1787/6c9202e8-en>. [2]
- OECD (2024), *现代化社会保障接入：策略、技术和数据*, *经合组织国家的发展*, 经合组织出版商, 巴黎, <https://doi.org/10.1787/af31746d-en>. [1]
- 经济合作与发展组织 (2022年), “经济合作与发展组织人工智能系统分类框架”, *经合组织数字经济* [5]
论文, 第323号, 经济合作与发展组织出版社, 巴黎, <https://doi.org/10.1787/cb6d9eca-en>.
- 皮尔, S. 等 (2017), “欧盟的能源贫困：政策与措施分析”, *政策与措施分析*, 在 *欧洲能源转型：政策制定的洞察*, <https://doi.org/10.1016/B978-0-12-809806-6.00030-4>. [7]

注意

¹ <https://www.aoc.cat/en/ia-transparencia-pobresa-energetica/>.

2

挖掘值得信赖的AI在社会安全领域的潜力

本章探讨了如何确保人工智能的可信和有效使用，借鉴了从国家层面收集的数据，以帮助比较各国的趋势。在可能的情况下，从社会保障部门额外补充了案例，以帮助突出国家和组织方法之间的相互作用，为人工智能的有效和可信使用创建共同治理框架。它关注三个成功采用公共部门的支柱：支持者、护栏和参与。支持者是指支持人工智能实施所需的设施、数据、技能和治理。护栏涵盖法律、道德和监督措施，以帮助管理风险并将人工智能的使用与公共价值观保持一致。参与强调员工、用户和利益相关者在设计和部署人工智能系统方面的重要性。尽管出现了有希望的倡议——特别是在试点项目中——但在战略、协调和包容性参与方面仍然存在差距。本章呼吁社会保障机构借鉴这些实践，以采取更一致、跨政府的方法，确保社会保障中的人工智能是透明的、公平的，并以用户的需求为中心。

为释放人工智能在提升公共服务质量和降低成本方面的潜力，需要建立一套强有力的治理框架。换言之，仅仅拥有采用特定人工智能工具的专业团队是不够的，还需要在国家层面和组织层面建立相应的结构，以治理人工智能，确保其使用有效且可靠。

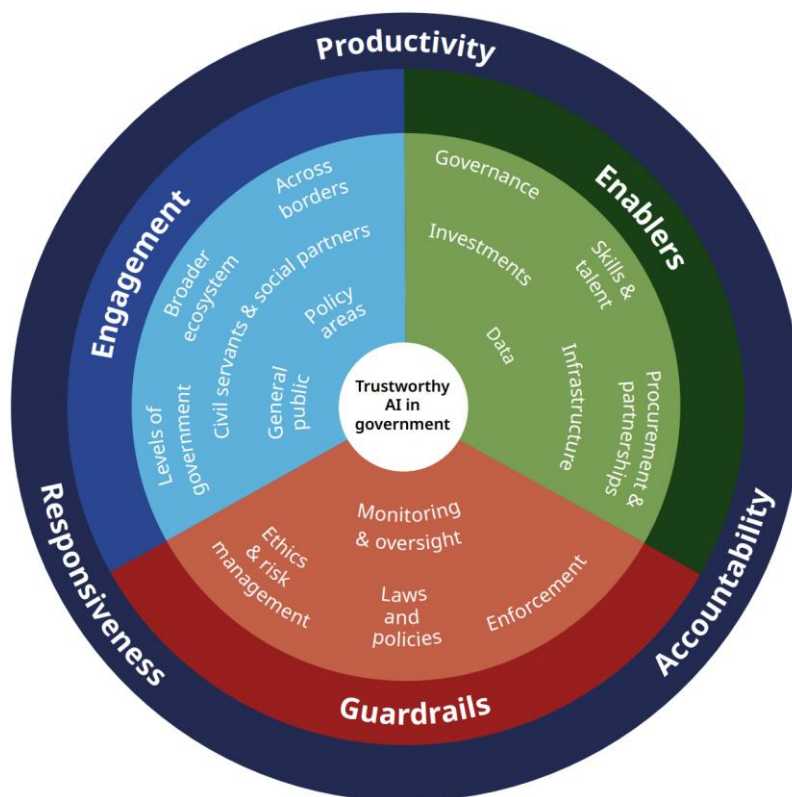
借鉴欧盟人工智能法案和 经合组织政府可信人工智能框架

(OECD, 2025^[1]), 本章重点阐述了公共部门成功采用AI的三大支柱：促进因素、监管措施和参与。促进因素指的是支持AI实施所需的设施、数据、技能和治理。监管措施涵盖法律、伦理和监督措施，以帮助管理风险并将AI使用与公共价值保持一致。参与强调了让员工、用户和利益相关者在设计和部署AI系统方面参与其中的重要性。此外，这些支柱有助于创建一个高效、问责且能响应其服务的人民和商业需求的AI赋能公共行政体系。

这些支柱在社会保障领域也尤其关键，因为人工智能系统被用于与福利和服务提供相关的决策中。因此，社会保障组织需要在组织内部构建使能因素、护栏和参与方法，但同样需要与中央政府机构、多边系统和其他部门合作，以确保人工智能的一致且有效运用，同时满足监管要求。

图2.1 经济合作与发展组织 (OECD) 政府可信人工智能框架

战略、有效、负责和值得信赖的人工智能实施的三个支柱



来源：(经合组织，2025[1])

为了帮助深入了解欧盟及以外这些治理框架的发展，本章使用了为人工智能元素2025经济合作与发展组织数字政府指数（经济合作与发展组织，即将出版）捕获的初步数据。^[2] 数据已于2025年收集，涵盖了2023-24年期间。对人工智能附件的回应包括22个欧盟国家和18个非欧盟国家。所有国家要么是经合组织成员，要么是加入国。¹

该数据来源于各个国家的国家级层面，有助于比较不同国家的趋势。在可能的情况下，还从社会保障领域补充了更多示例，以帮助突出国家和组织方法之间的相互影响，从而为人工智能的有效和可信应用建立共同治理框架。

赋能者—— 是否为支持人工智能的使用提供了合适的赋能者？

赋能因素对于政府有效实施人工智能至关重要。它们建立了让人工智能由专业公共部门专业人员能够有效、负责任且自信地进行设计和部署的条件。关键赋能因素包括强大的人工智能治理（包括人工智能和数据战略方法）、人工智能能力（包括基础设施、技能和人才）以及合作与伙伴关系。所有这些都需要中央权威或人工智能协调机构在国家层面付出努力，同时也需要在社会保障机构内部的组织层面付出努力，以确保它们配备了所需资源，能够有效地支持福利发放。虽然本章侧重于关键的治理和技术赋能因素，但与劳动力相关的重要赋能因素在第三章中进行了涵盖。

本项目前期分析（经济合作与发展组织，未发表）^[3]表明法国和意大利的社会保障领域人工智能应用仍处于早期阶段，相关机构正通过测试、试点和实验探索机会并管理风险。尽管进展正在有机地取得——由IT或统计团队推动并通过内部咨询支持——但治理、基础设施、数据访问和数字技能等关键赋能因素仍与技术同步发展。然而，需要更清晰、机构范围的人工智能战略，以整合人工智能的战略方向，支持基础设施升级，改善数据共享，并扩大有前景的人工智能应用案例。此外，通过国家层面的更大合作和机构间人工智能工具的更好共享，人工智能探索将获得进一步赋能的明确机遇。

指导社会保障机构内人工智能使用的策略

战略文件是推动人工智能实施、建立长期目标以及规划投资的基础。与国家战略相补充，社会保障领域的组织可以制定机构或行业战略，以引导人工智能的采用，并针对其用户的具体需求。这些战略可以提供必要的分量，以帮助克服现代化基础设施以增加计算能力、改善数据访问和共享以及为扩大试点项目提供明确授权等关键挑战。

许多国家政府都有一个国家人工智能战略（在参与的欧盟成员国中为91%，在参与的欧盟非成员国中为78%，参见表2.1）。任何机构层面的补充战略都需要与国家战略的总体愿景和框架保持一致，以实现公共部门人工智能的协同使用，从而最大限度地发挥其影响。

表2.1 公共部门人工智能国家战略

基于参与欧盟成员国与非欧盟国家（经合组织成员国和伙伴）的国家级数据

	%欧盟	%非欧盟
是的，目前到位。	91% (20/22)	78% (14/18)
策略目前正在开发中。	0% (0/22)	6% (1/18)

注意：此数据基于22个欧盟成员国和18个非欧盟经合组织成员国及伙伴国的初步数据。来源：为2025年经合组织数字政府指数收集的数据。

例如，芬兰提供了一个国家层面和组织层面的AI努力的策略一致性的强有力范例。其国家AI战略通过资金、基础设施和劳动力发展支持公共部门的创新。此外，芬兰的机构——如Kela和DigiFinland——将芬兰描述为一个“以数据驱动国家”，其数据治理已设计了几十年。同时，国家社会保障机构Kela已将AI和自动化嵌入其组织战略中以提高服务交付和内部效率。通过将国家目标与机构层面的实施保持一致，芬兰确保了政策、投资和执行之间的一致性——最大限度地发挥AI的影响力，以构建更具响应性、安全性和以公民为中心的社会保障体系。

根据治理体系和自治程度，在地方层面可能也需要制定战略。例如，赫尔辛基市在一种明确的伦理人工智能框架下运作，该框架通过与在其服务提供背景中促进透明度、公平性和人工监督，来补充芬兰的国家战略。类似地，西班牙的加泰罗尼亚地区拥有一项人工智能战略，该战略制定了地方投资计划，并强调了考虑地方语言的需要。

人工智能的战略投资

在人工智能项目上的战略投资——包括专门资金和明确的采购指南——对于政府部门采用和持续实施人工智能至关重要。它使机构能够开发高质量解决方案、吸引人才并建立必要的设施。如果没有对人工智能项目的专门资金，努力可能会陷入碎片化、资源不足或实施不善的风险。有针对性的投资可以加速创新、降低风险，并使政府能够提供更智能、更高效的公共服务。最后，明确的采购指南也有助于确保人工智能工具符合伦理、安全，并与公共价值观保持一致。

表 2.2. 关于采用人工智能系统的投资决策

基于参与欧盟成员国与非欧盟国家（经合组织成员和伙伴）的国家级数据

	%欧盟	%非欧盟
基于潜在效率或成本节约的证据（基于案例研究）	41% (9/22)	44% (8/18)
为响应用户需求或要求（基于调查或反馈分析）	23% (5/22)	22% (4/18)
政治或行政指令（基于人工智能战略，或高级规划）	68% (15/22)	50% (9/18)
公共部门人工智能没有具体的投资决策标准	27% (6/22)	28% (5/18)

注意：此数据基于22个欧盟成员国和18个非欧盟经合组织成员国及伙伴国的初步数据。来源：为2025年经合组织数字政府指数收集的数据。

当审视驱动公共部门人工智能系统投资决策的因素时，如表2.2所示，27%的欧盟成员国报告尚未采用任何特定标准，而非欧盟成员国则为28%。这一数据凸显了欧盟成员国采用

人工智能投资的战略方法，这同样反映在之前对法国和意大利社会保障机构的分析中（经合组织，未发表）^[3]。尽管当时已有朝着长期规划发展的努力，但这些努力仍处于初级阶段，因为现有的投资集中于使用内部能力进行试点，以及对应从私营部门获取哪些能力进行了一些探索。似乎没有哪种方法在组织层面或国家层面显示出强有力的协调或监督。

欧盟成员国的公共部门人工智能投资受到多种不同的方式管理，从去中心化框架到高度集中的监督。例如，在 **奥地利** 各部委根据国家的中央人工智能战略独立定义其投资方针。相反，**保加利亚** 通过电子政务部集中决策权，该部门对所有的信息和通信技术（ICT）和人工智能（AI）相关支出行使财务和战略控制权。在 **法国**，数字事务部际指导方向（DINUM）协调投资，以确保决策既以问题为导向——需要明确、可衡量的效益——又与人工智能战略中概述的国家目标战略一致。**匈牙利** 强调成本效益和减少官僚主义，在政治优先事项的指导下，并由其人工智能联盟支持。虽然方法各异，但共性包括更加重视可证明的价值、与国家数字战略保持一致，以及将人工智能融入更广泛公共部门现代化努力中。

通过采取对人工智能的战略投资方法，也有机会探索共同的方法来确保和主权解决方案，以及促进资源共享和共享专业知识。这样，更战略性的欧洲投资可以由此支持欧盟范围内更可持续和包容的人工智能采用，特别是当其成员国面临预算限制时。

人工智能专项基金

国家对人工智能的专门拨款——无论是国家层面还是组织层面——确保了对基础设施、内部人才和研究的持续投入，从而开发并部署与公共价值相一致的、值得信赖、高影响力的AI系统。它还可以促进跨部门战略规划和协调，减少碎片化，提高AI驱动公共服务的可扩展性和有效性。虽然国家层面的拨款可以帮助克服组织层面的预算压力，并确保政府内部更一致的方法，但将组织预算的一部分用于创新解决方案可以是一种有效的方式，用于在社会保障机构中探索人工智能。

表2.3中的数据显示，参与欧盟和非欧盟成员国的AI计划报告的可获得资金率几乎相等且很高。然而，值得注意的是，这两个国家集中的可用资金也属于更广泛的数字和ICT提案资金的一部分。虽然这意味着AI更全面地融入了数字和ICT投资的总体投资决策中，但争夺通用资金的竞争可能会限制在AI采用早期阶段所需的实验和发展。

表 2.3. 用于开发和部署公共部门人工智能系统的专项资金

基于参与欧盟成员国与非欧盟国家（经合组织成员和伙伴）的国家级数据

	%欧盟	%非欧盟
是否总体	86% (19/22)	78% (14/18)
是的，有专门用于人工智能的资金。	45% (10/22)	33% (6/18)
是的，作为对数字和ICT提案的更广泛资金支持的一部分。	41% (9/22)	44% (8/18)
不，没有资金可以用来测试和发展人工智能系统。	14% (3/22)	22% (4/18)

注意：此数据基于22个欧盟成员国和18个非欧盟经合组织成员国及伙伴国的初步数据。来源：为2025年经合组织数字政府指数收集的数据。

欧盟公共部门采用人工智能的资金模式在结构和集中化方面存在显著差异。例如，**奥地利** 依赖于去中心化方法，每个部门管理其自身的与人工智能相关的预算。**葡萄牙** 同样将AI资金整合到更广泛的数字化转型计划中，使资源与国家战略目标保持一致。**法国** 相比之下，它通过公共行动转型基金（FTAP）等集中式机制来引导定向投资，该基金支持跨部门倡议，如ALLiaNCE孵化器。在组织层面，赫尔辛基市在**芬兰**，资助了“试验加速器”，这是一个内部项目，用于探索和发展人工智能试点（另见第3章）。

为人工智能提供专项资金也将为补充围绕吸引和发展人工智能人才和技能（如第3章中讨论的）所做的努力提供一个机会。为了确保此类投资的效益能够持续一段时间，保留这些新开发能力的相关补充性努力也可能很重要。

支持人工智能采购

当组织不完全内部开发人工智能系统时，采购指南对于确保所获取的系统符合严格的伦理、法律和技术标准至关重要，特别是在敏感或高风险环境中。内部支持工具，例如指南、标准条款或供应商名单，也增强了机构评估、管理和监控人工智能解决方案的能力，促进了透明度、问责制和长期公共价值。数据显示，人们越来越重视采购需要发挥的作用，50%的参与欧盟成员国和56%的非欧盟成员国支持在公共部门采购人工智能。

表2.4。支持公共机构在人工智能采购方面的支持

基于参与欧盟成员国与非欧盟国家（经合组织成员和伙伴）的国家级数据

	%欧盟	%非欧盟
是否总体	50% (11/22)	56% (10/18)

注意：此数据基于22个欧盟成员国和18个非欧盟经合组织成员国及伙伴国的初步数据。来源：为2025年经合组织数字政府指数收集的数据。

然而，采购方法多种多样。例如，**立陶宛** 强调通过法律框架、预商业化采购模式和中采购支持来推动创新驱动采购。**法国** 拥有结构化的采购策略，优先考虑安全、主权和成熟的AI解决方案，并将启动一项全国性征召活动，以识别生成式AI工具，以便快速向公众部署以支持此目标。**瑞典** 提供负责任的AI采购国家指导，包括针对生成式AI的具体建议。尽管**捷克** 和**匈牙利** 它们正处于过渡阶段，正在更新策略并准备制度支持机制，以符合欧盟人工智能法案的要求（有关更多信息，请参阅“护栏”部分）。

根据国家背景，这些采购方法可能适用于社会保障机构，以简化采购流程、实现物有所值，并实现更具创新性的采购流程。在一个国家背景下不存在这些中央安排的情况下，组织可能需要考虑为采购团队配备必要的AI素养水平，以便能够驾驭AI系统的采购，以及利用共享资源，如**欧盟的人工智能模型合同条款**。²

可扩展的基础设施

可扩展的基础设施确保人工智能解决方案能在需求增长时获取计算能力，以实现更快开发、更好的决策，并进而提供更高效的社会服务。

在组织层面，许多社会保障机构在考虑本地、云和共享计算基础设施与数据中心之间进行权衡。本地基础设施可以提供对敏感数据的更大控制权，并可能支持严格的合规需求，但由于维护成本更高且可扩展性较低。云解决方案可以提供更大的灵活性，更容易访问先进的计算能力，并有可能节省成本，尽管它们可能会引发数据主权和供应商锁定方面的担忧。多个参与方之间的共享基础设施也可以是减少重复并降低各机构成本的一种方式，但这需要强有力的治理、协调和控制来确保数据隐私和安全。考虑到基础设施和数据，表2.5概述了用于支持公共部门AI集成的关键组件——无论是作为主要基础设施解决方案，还是作为结合它们的混合模型的一部分。

表2.5 数字基础设施的关键组件正在被用于支持人工智能的集成

基于参与欧盟成员国与非欧盟国家（经合组织成员和伙伴）的国家级数据

	%欧盟	%非欧盟
云计算平台	59% (13/22)	78% (14/18)
数据中心和存储系统，包括计算能力	68% (15/22)	50% (9/18)
硬件加速器（例如：GPU）	41% (9/22)	33% (6/18)

注意：此数据基于22个欧盟成员国和18个非欧盟经合组织成员国及伙伴国的初步数据。来源：为2025年经合组织数字政府指数收集的数据。

来自欧盟各成员国的案例反映了在安全、互操作性和数字主权方面的不同国家优先事项。在国家层面，**法国** 通过其国家云战略鼓励可信云采用，优先考虑经SecNumCloud认证的SaaS解决方案。**比利时** 和 **斯洛文尼亚** 提供集中式政府云平台以简化服务交付。**爱尔兰** 为企业提供云采购指导，以确保跨机构的 securely and efficiently 采用。**意大利** 在其公共部门创新框架内，逐一探讨云应用的案例。**奥地利**，社会保险机构的IT服务，ITSV，在其本地运行AI模型，使用应用程序容器以增加安全性。Kela也做出了类似的选型 **芬兰**，在本地运行其 AI 平台。Kela 的平台也在应用程序容器上运行，使用免费和开源软件，以确保安全性、稳定性和易于复制。这些努力反映了一种共同和战略性的转变，朝着主权计算能力发展，以确保安全和可扩展的 AI 系统。

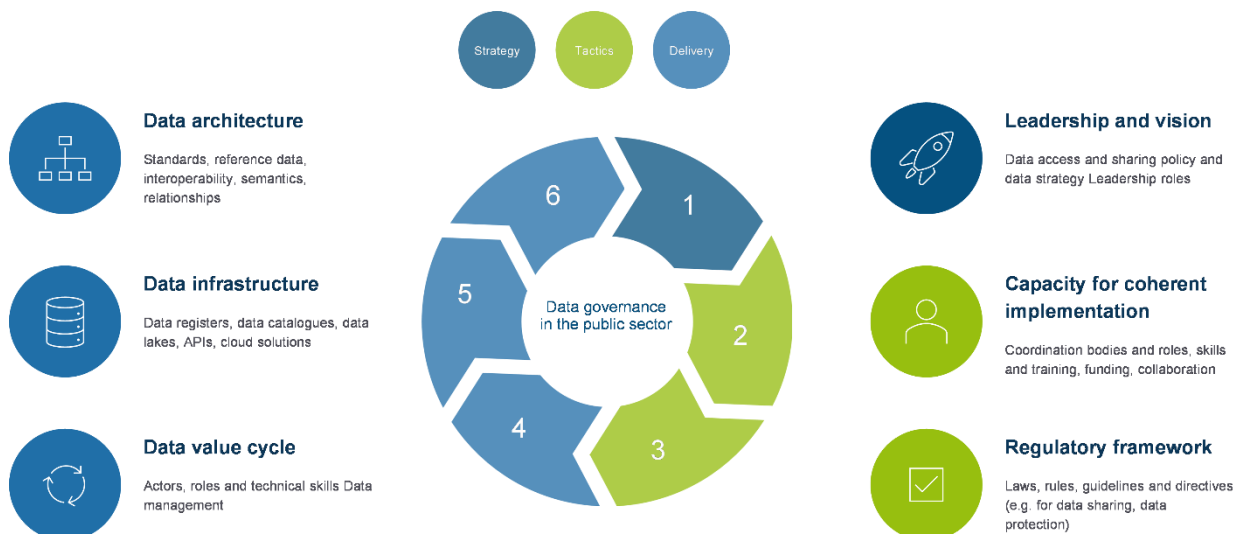
人工智能的数据治理

获取高质量、准确且相关的数据是公共部门有效并可信地使用人工智能的基础。人工智能系统依靠数据来学习、进行预测并支持决策；而用于训练人工智能模型的数据中的任何缺陷或偏见都将影响其结果。在社会 security sector，这一点尤为重要，因为人工智能工具可能会影响有关福利资格、欺诈检测或个性化服务交付的决策。所有这些领域都是选择包含在模型中的数据可能导致对个人造成严重甚至有害后果的地方。

数据治理框架为数据使用和共享提供了连贯的愿景，确保在各级政府中制定清晰的数据使用、处理和共享策略。同时，数据治理可以帮助公共部门组织负责任地处理数据，通过保护隐私和数据安全，并满足监管要求，从而为值得信赖的人工智能系统奠定基础。

如图2.2所示，由六个元素组成的 **经济合作与发展组织公共部门数据治理模式框架**（OECD，2019^[4]）所有这些都影响着数据如何被用于人工智能。为确保数据能够成为人工智能采用的推动者，需要更新数据战略以使其与更广泛的人工智能目标保持一致，提供高质量的 结构化数据，构建互操作性解决方案，促进数据共享。

图2.2. 经合组织公共服务领域数据治理模式框架



来源：(经合组织，2019[4])。

在欧盟和非欧盟成员国中，数据治理框架和数据集成工具被用于支持在国家层面将人工智能集成到政府中。然而，这些工具的采用尚未普及，表明需要进一步进步。

通过确保训练数据不重复社会结构不平等，可以提升人工智能模型的公平性。此外，这必须与数据隐私和安全方面的强大保障相平衡。社会保障数据通常包含敏感个人信息，将其用于人工智能必须遵守严格的法律和伦理标准。这包括使用数据训练人工智能模型、数据处理、匿名化和访问控制的明确规程，以维护公众信任、保护公民权利，同时确保算法透明度，促进创新和改进服务交付。图 2.2 展示了用于支持人工智能在参与欧盟和非欧盟成员国间整合所使用的的数据治理的一些关键组成部分。

将AI嵌入数据战略

更新数据政策以与更广泛的AI目标保持一致需要将AI考量嵌入到数据价值链中。例如，数据收集应确保数据具有代表性，并且不会引入对基于其训练的任何AI模型的偏见。这种融合确保数据治理可以成为AI治理的基础层。

领导力在协调数据和人工智能战略中发挥作用，确保对可信人工智能的数据再利用是共享愿景的一部分。在 **法国**，数据与AI政策的融合由数据、算法和源代码总管理员领导。这一新角色取代了首席数据官，确保数据与AI交汇领域的强大授权和协调。

数据策略的协调实施被集中 **立陶宛** 通过一个由政府管理的数据空间，其基础是强大的法律和数据处理标准。在 **欧盟**，数据治理法案和数据法案为数据访问、共享和再利用提供了一个统一框架。这些模式共同有助于通过创建一个安全、互操作和透明的环境来促进数据的共享和使用。

数据对齐能力和人工智能实施能力

数据治理机构也应与负责人工智能实施的机构保持一致。例如，数据管理员角色应融入人工智能项目团队，人工智能伦理委员会应包括数据治理专家。这种一致性确保了在数据访问、质量和再利用方面的决策都充分理解其对人工智能结果的潜在影响。

数据质量与相关性

高质量数据是人工智能训练和使用的先决条件。数据质量的不同维度对应用于数据的模型类型（Mohammed等，2025）具有不同影响。^[5] 因此，在许多国家，与掌握和处理数据的相关方进行合作是国家级数据战略的关键部分。此外，用于训练模型的数据应与预期结果相关。因此，数据质量不仅取决于技术标准或格式，还取决于将影响模型输出的定性因素。

为解决数据质量的实际问题，已制定相关指南和标准，特别是在开放数据领域。**爱尔兰**，由公共支出改革部开放数据单位起草的开放数据发布指南（Lee, 2021）^[6] 鼓励发布者使用国际标准来提高数据集之间的互操作性，例如使用 ISO 639 语言代码或 NUTS 地理分类。**意大利**的数字代理机构，AgID，建议遵循四个数据质量特征：准确性（句法和语义），一致性，完整性，及时性（AgID，2023）^[7]。

最后，国家统计局**芬兰**已草拟了一个数据质量框架，旨在帮助用户评估公共管理数据集的可重用性，并帮助数据生产者提高数据质量。该框架是芬兰公共管理人工智能战略的一部分。该框架的标准分为三类：信息描述现实的程度（正确性、准确性、一致性、时效性、完整性）、信息描述方式（可追溯性、可理解性）以及信息可用性（可移植性、用户权利、及时性）（芬兰统计局，2022）^[8]。

语义一致性对于人工智能系统正确地跨各种来源解释数据至关重要。这包括遵循元数据标准，采用开放且可互操作格式，并使用共享词汇表和本体来描述数据元素，如前文所述 *经合组织关于加强数据获取和共享的建议书*（OECD, 2017）^[9]。诸如 *schema.data.gouv.fr* 在法国或 *schema.gov.it* 在意大利说明了政府可以通过提供标准化的数据共享和发布模式来促进语义一致性，最终提高数据的互操作性和可重用性。

尽管如此，虽然数据质量指南非常严格，但它们的实施仍然基本上是自愿的。根据2023年经合组织OURdata指数，不到一半由国家发布的高价值数据集包含完整的元数据或可通过API访问，这是AI系统中可发现性和可重用的两个基本条件。这表明，即使对于最关键的数据集，质量和技术可访问性方面的差距仍在继续限制它们在高级应用中的潜力。（经合组织，2023）^[10]。随着社会保障机构考察如何利用现有数据来识别和解决未领取情况，如果数据集不可访问或无法发现，他们的努力将受到限制。

数据集成与互操作性

数据集成和互操作性框架是政府中人工智能的关键使能因素，因为它们允许公共部门组织安全高效地共享和复用数据。这些模型依赖于标准化协议和通用数据结构来确保跨系统的兼容性。互操作性框架还定义了数据交换的规则，包括认证、授权和日志记录，这在处理敏感信息时至关重要。

互操作性平台通常包括 API 网关和数据目录。例如，**加泰罗尼亚语**

互操作性模型为加泰罗尼亚政府及地方当局之间的数据交换提供了治理和技术框架。通过 **奥布塔** 服务，管理者可以通过遵守互操作性框架协议（CMI），访问互操作数据和文档目录。该模型还连接到西班牙的国家互操作性平台，**数据中介平台**。

其他国家也存在类似的方法。在 **芬兰**，数字与人口数据服务机构管理下的基于X-Road的数据交换层能够实现组织间安全、标准化的数据传输。接入平台需要严格的技术和管理合规，包括接口合同和安全协议。可用服务列在公共API目录中，确保透明度和可发现性。**意大利的 国家数据数字平台**（PDND）遵循类似的逻辑，作为一个国家互操作性枢纽，执行标准API和元数据模式，以促进跨部门的数据共享。

用于训练AI模型的资源（例如公开数据集）

人工智能系统可以对其训练时所使用的类似的数据进行预测（经合组织，2022年）^[11]。用于训练人工智能模型的资源，例如公共数据集，因此是政府中人工智能的关键使能因素，因为它们提供了开发准确、相关和有效的人工智能系统所需的基础数据。获取高质量、有代表性的数据集确保人工智能系统更好地与公共部门需求保持一致，并能提供明智的结果。

因此，使用真实数据训练模型是确保其准确性的最佳方法。这样做最好的方式是在组织内部进行，由具备相关数据的专业技能和知识的团队来完成，并且不会将机密信息泄露给第三方。例如，**奥地利的** ITSv已经开发了一款可以处理医疗账单的产品，无论其格式如何。该产品的AI模型是由过去手动处理账单的工人进行训练的。通常，该软件会要求用户手动填写数据，而不是用AI预先填写，以便比较两个结果并测试准确性。**德国的** ADEST在部署环境中也使用与其分析相同类型的数据进行了训练，而员工对预测的手动校正也在不断改进模型。

在 **葡萄牙**，与社会保障的受益人关系模型（Visão 360和E-clic）已实现自动化，允许用户请求信息与澄清，提交投诉，跟踪请求状态，以及查阅与社会保障的过往互动。目前，上述请求的答复由工作人员处理，但机器正在与工作人员进行人工智能学习过程，以在不久的将来自动提供其能够提供的回复类型，并使用实际参考。

其综合社会保障风险管理平台也使用人工智能作为其预防和控制内欺诈和不规范行为的战略的核心部分。通过人工智能驱动的机器学习，该系统交叉引用多个内部和外部数据源，构建受益人和公司的风险画像，评估行为指标，识别和检测风险模式/可疑异常，并分配风险等级。根据风险等级，由服务机构提出适当的行动方案。它起预防性和预测性作用。最后，葡萄牙社会保障权利模拟器也使用人工智能根据公民的历史和具体条件来预测未来场景和社会保障权利。

此外，在开放数据上训练模型对于确保人工智能透明度至关重要。许多欧盟国家确实维持着国家开放数据门户网站，为人工智能训练和模型开发提供基础资源，而国家门户网站集中在了欧洲的一个数据.europa.eu上。例如，**法国的** 联盟计划还策划了高质量、特定领域的 datasets，以支持公共部门的人工智能发展。

这些努力展现了共同致力于使基于正确数据的可信赖、准确且有效的AI系统得以实现。

测试和实验

对人工智能产品进行小规模常规持续测试有助于及早发现问题并让最终用户参与设计，从而确保成功过渡到部署阶段。为此，经合组织人工智能原则建议政府“考虑使用实验来提供一个受控环境，以便测试和扩展人工智能系统”（经合组织，2021）^[12]）。在**加泰罗尼亚**，AOC在推广该解决方案之前，先在三个市镇测试了能源贫困报告的自动化。**德国**联邦就业局也参与了ADEST的开发，让全国各地的就业机构的代表小组运行了一个试点。最后，在**芬兰**，凯拉的人工智能项目是探索性的：与其寻求保证这些项目的成果或投资回报率（ROI），该组织反而将其思维方式转向实现“学习回报”。这意味着每个项目——无论其是否成功实施——都对组织有价值，或可为未来项目提供经验教训，或可作为可构建或再用于其他项目的技术组件。

护栏——各国如何管理公共部门人工智能的风险？

防护措施是关键政策工具，用于应对人工智能的风险，以支持政府在人工智能方面的负责任和可信赖使用。它们包括具有约束力和非约束力的措施——例如法律、法规、标准、监督机构以及透明度流程——这些措施有助于管理风险，并确保人工智能与法律、道德和社会价值观保持一致。在社会安全领域，防护措施对于保护用户权利和维护公众信任尤为重要。当与适当的促进因素相结合时，它们有助于为公共部门中人工智能的开发和部署创建一个安全且负责任的环境。

在法国和意大利的社会保障领域，人们致力于负责任地使用人工智能，并遵守国家和欧盟的法律框架。早期的努力——例如伦理宪章、审计机制和结构化的人工智能框架——前景光明，但目前仍有许多安全措施处于开发阶段。随着人工智能系统的不断演进，并可能根据欧盟人工智能法案被归类为高风险，机构需要加强围绕透明度、偏见和监督的安全措施，将中央开发的安全措施适应社会保障的具体环境。与上一节讨论的促进因素一起，这些安全措施将是确保人工智能的合乎道德、值得信赖和以用户为中心使用的关键。

除了立法之外，表2.6表明——尽管各国开始实施保障公共部门人工智能安全可靠使用的护栏——各国仍需要显著提升在人工智能使用方面的监管、透明度和问责力度。

表2.6. 管理人工智能工具使用的护栏措施

基于参与欧盟成员国与非欧盟国家（经合组织成员和伙伴）的国家级数据

	%欧盟	%非欧盟
生成式人工智能工具使用伦理指南	77% (17/22)	56% (10/18)
定期审计以确保符合人工智能政策和法规	23% (5/22)	17% (3/18)
外部机构或专家的独立监督	41% (9/22)	17% (3/18)
透明度措施，例如公开人工智能生成的内容	41% (9/22)	28% (5/18)
用于应对潜在滥用或错误的问责框架	23% (5/22)	17% (3/18)

	%欧盟	%非欧盟
用于确保数据安全性和隐私的数据标准或协议 用于生成式AI工具	32% (7/22)	33% (6/18)
其他	14% (3/22)	11% (2/18)

注意：此数据基于24个欧盟成员国和17个非欧盟经合组织成员及伙伴国的初步数据。资料来源：为2025年经合组织数字政府指数收集的数据。

立法框架

围绕政府中的人工智能立法对于确保人工智能系统得到负责任、合乎道德且符合公共利益的使用至关重要。随着政府越来越多地采用人工智能以改进服务、简化运营和提升决策能力，建立明确的法律框架以解决问责制、透明度、公平性和数据保护等问题就显得至关重要。立法还有助于设定采购、部署和监督的标准，确保人工智能系统支持民主价值观和人权。它使政府能够以身作则，在促进创新的同时保护民众免受新兴技术潜在风险和滥用的侵害。

The 欧盟的《人工智能法案》是首批人工智能综合立法框架之一。它根据风险等级将人工智能系统分为不可接受、高、有限和最低四类，并对高风险系统提出严格要求，特别是在执法、关键基础设施和招聘等领域。该法案强调透明度、人类监督和问责制，包括对高风险人工智能进行强制性文件记录、合规性评估和注册。通过执行这些标准，欧盟旨在平衡创新与基本权利保护，为全球负责任的人工智能治理提供典范。

虽然人工智能系统可以提升社会保障的发放，但如果缺乏合适的治理、控制与监督，也可能带来风险。社会保障因此是《法案》中强调的高风险具体领域之一，特别是在人工智能系统是：

“... intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services.”

(欧盟，2024年人工智能法案，附件III.5.a [13])

这仅适用于人工智能系统取代或显著影响有关福利的决策结果的情况，这种情况可能损害受益人的福祉和权利。第6.3条阐明，这不包括人工智能系统旨在执行狭义程序任务；改进先前已完成的人类活动结果；检测决策模式或与先前决策模式的偏差；且并非旨在在不进行适当人工审查的情况下取代或影响先前已完成的人类评估；或执行与附件III（欧盟，2024）中先前列出的使用案例相关的评估的准备工作。^[13]）。

在欧洲各地，各国正在着手实施必要的治理，以确保其人工智能系统的开发和使用符合欧盟人工智能法案的义务。例如，Kela——芬兰的社会保障机构——既是人工智能系统的开发者，也是提供者，这意味着该法案的所有义务都与它们相关。为此，该机构已实施了一个评估其人工智能系统的流程，该流程要求其团队在使用案例存在时，考虑分类、Kela的角色以及任何额外的监管考虑因素：

- 一个正在被设计的新人工智能系统；
- 一个正在采购的AI系统（可以是独立的，也可以集成到另一个解决方案中）；
- 添加或更改集成到另一个解决方案中的AI；

- 受系统使用数据变化的影响；
- 受法规变更影响；或
- 受系统所基于的模型变更影响。

为了帮助团队了解法案的义务以及这对他们的人工智能用例意味着什么，Kela 还建议团队使用合规检查工具。

伦理框架和标准

伦理框架对于指导政府中人工智能的负责任使用至关重要，确保系统公平、透明、问责且尊重人权。如果没有明确的标准，人工智能可能因偏见、歧视或个人数据的滥用而造成伤害——尤其是在社会保障领域。因此，伦理框架充当护栏，帮助公共机构做出明智、价值驱动的决策，同时建立公众信任。它们还确保符合法律法规，并支持各领域的一致实践。然而，要真正指导政府中人工智能的负责任使用，这些指南和原则不仅需要嵌入更广泛的治理生态系统，还需要融入公共部门员工的职业实践和日常工作中 (Fjeld 等人，2020)。^[14]，这在第三章中也将作为构建人工智能预备劳动力的一部分得到进一步探讨。

例如，在 **意大利**，agid 已制定了一系列关于在公共行政中采用人工智能的指南 (AGID, 2025^[15])。指南指出，公共行政部门可以自愿采纳道德和行为准则，以协助其实施标准。**奥地利** 提供有关人工智能伦理和治理的结构化指南，将欧盟人工智能法案纳入其监管框架。

立陶宛 已为公共部门中的人工智能确立了明确的原则，重点关注透明度、开放性和公众参与。**瑞典** 已针对生成式人工智能制定了专门指南，涵盖数据保护、伦理和劳动法等领域。**法国的ALLiance伦理中心** 提供基于透明度、公平性和人工监督的结构化指导。**斯洛文尼亚** 确保人工智能承诺对所有公共部门员工的可及性。**葡萄牙** 将人工智能伦理融入公共管理现代化努力中。**爱尔兰** 与欧盟和经合组织原则一致，提供负责任人工智能使用的结构化指南。尽管一些国家侧重于监管监督，但其他国家通过基于原则的方法鼓励更好的透明度、可及性和道德使用。这种情况适用于

加拿大政府，已制定了一系列关于在政府中使用人工智能的原则，包括一份关于使用生成式人工智能的指南（加拿大财政委员会秘书处，2025年）^[16]。

如第一章所述，社会保障领域的一个道德标准例子是 **德国**

联邦就业机构 (BA) 的人工智能愿景，其包含七项指导原则（参见框1.3），作为其人工智能伦理框架和标准的组成部分，设定了人工智能系统应关注人、自主和控制、安全、可靠性、数据保护、透明度和公平性的期望。BA的人工智能指导原则与内部数据伦理指南相连，该指南超越了人工智能实施。

更关注数据负责任使用，芬兰的 **科拉** 已采用乌德勒支大学 (2025^[17])—数据伦理决策辅助工具——为指导其员工基于以下标准对人工智能系统进行伦理评估：

- **协作团队评估** 通过在解决方案开发团队中进行结构化讨论。
- **多学科方法**：包括技术专家、法律顾问和商业代表。理想情况下，客户或最终用户也会参与。
- **引导式自我评估**：通过这种方式，团队必须通过有根据的对话而非预定答案来得出结论。
- **已记录的结果**：在一份最终报告中，包含推荐的下一步行动和分配的责任。

这两个来自社会保障领域的例子因此突出了制定员工遵循的指导原则的必要性，同时也需要考虑如何将这些原则嵌入到流程中，以便它们始终是首要考虑的事情，并成为他们日常工作中实践的内容。

定期审计以确保符合人工智能政策和法规

合规审查对于确保政府中的人工智能计划符合法律、道德标准和最佳实践至关重要。它们有助于识别偏见、数据滥用或缺乏透明度等风险，并确保在整个人工智能生命周期中都具有问责制。定期审查通过验证系统按预期运行、维护公共价值观且不造成伤害，起到护栏的作用。它们还建立信任，提高系统性能，并支持持续改进。通过追究机构的责任，合规审查有助于确保人工智能被负责任、合法和道德地使用——使其成为公共部门可信赖人工智能的关键保障。然而，也许是社会保障部门人工智能系统部署的早期阶段，定期审查的实践非常低。

在那些做的人当中，**匈牙利**已建立结构化审计框架，匈牙利竞争局（GVH）和人工智能法案强制要求对人工智能系统、透明度措施和责任框架进行定期评估。高风险人工智能系统接受独立评估，如Eötvös Lóránd大学（ELTE）等机构为公务员提供关于如何开展这些审计的培训。**爱尔兰**将其人工智能审计整合到其标准的政府监督机制中，内部审计由总审计长进行审查补充。最后，**法国的**联盟投资委员会**c**对AI项目进行定期审查，但侧重于影响指标和伦理一致性，而非报告起草时的正式审计。

这些定期审计非常重要，特别是在社会保障领域，因为没有严格的人工监督，存在出现错误和拒绝合格受益人享受福利的风险。例如，在**塞尔维亚**，实施了一种半自动化流程，用于评估社会卡资格，并减少申请人提供大量文件的管理负担。然而，算法决策往往使用不准确或分类错误的的数据，这影响了社会保护的使用——尤其是弱势群体。社会工作者没有足够的培训或识字能力来理解如何使用该系统，这意味着决策通常完全委托给自动化流程，而不是将其用作支持他们自身评估的辅助工具（大赦国际，2023）^[18]在这些示例中，定期审计有助于识别流程中的故障或人为疏忽的差距。

外部机构或专家的独立监督

外部机构或专家的独立监督对于确保政府人工智能计划的客观性、透明度和问责制至关重要——包括作为组织监督和监管一部分的独立审计和自我审计。它为内部团队可能忽视的潜在风险（如偏见、滥用或缺乏公平性）提供了中立核查。外部审查者可以提供多元化的视角，验证符合伦理和法律标准，并增强公众信任。这种监督通过防止利益冲突、鼓励负责任的创新以及确保人工智能服务于公共利益，发挥着关键的防线作用。最终，它有助于维护公共部门在人工智能应用中的诚信和问责制。

在整个欧盟，各国已采用不同方法对政府中的人工智能进行独立监督，通过结合外部机构和专家来确保合规性和伦理治理。**法国**

将指定多个当局——例如法国国家信息与自由委员会（CNIL）、跨部门数字管理局（DINUM）和视听与数字通信监管机构（ARCOM）——以保障欧盟人工智能法案下的基本权利，确保人工智能系统符合伦理和监管标准。**匈牙利** A I联盟，组成

由科技部牵头，包括政府机构、学术界和行业专家，共同监督人工智能的开发和实施。**爱尔兰**已建立分布式模型，将行业监管机构分配用于执行欧盟人工智能法案，并指定额外授权机构进行全面监督。这些方法已根据其国家背景进行调整，但所有目标都是在使用人工智能的公共部门时平衡创新、问责制和安全。

透明度措施

透明度措施对于确保政府对人工智能的使用负责任和获得公众信任至关重要，特别是披露人工智能的使用或人工智能生成的内容、数据的使用以及人工智能工具的开发方式。它们帮助用户了解人工智能何时以及如何被使用，从而实现有根据的参与和监督。缺乏透明度，由人工智能做出或支持的决策可能会显得神秘或不负责任，有滥用或失去公众信任的风险。明确的披露可以促进负责任的使用，突出局限性，并实现监督。因此，这也应包括上一节中强调的审计的性能和发现的记录证据。

作为关键的护栏，透明度确保人工智能支持开放、公平和民主治理，强化政府技术应始终服务于公众并对其负责的原则。尽管采用率仍然较低，但数据显示各国对此同样关注，41%的参与欧盟成员和28%的参与非欧盟成员已实施透明度措施。

例如，**匈牙利**通过实施《人工智能法案》的要求，该条例要求对高风险人工智能系统进行定期审查、责任框架和独立评估，以实现透明度。

爱尔兰在GDPR框架内整合AI透明度，确保遵守数据保护法律，并要求明确披露AI生成的内容。**瑞典**的信任模型是一种自我评估工具，用于维持公共管理中的开放性，确保人工智能应用遵循伦理和监管标准。**法国**鼓励公共机构披露人工智能的使用情况，特别是在向利益相关者提供的通信和官方文件中。最后，**芬兰**正在开发一项人工智能标签概念，以增强透明度，同时辅已建立的人工智能指南。

缺乏对人工智能运作方式和数据使用情况的透明度，存在破坏社会保障部门利用人工智能向善努力的风险。例如，在 **United Kingdom**，一个人工智能系统被实施用来筛选和优先处理劳动和社会保障部收到的海量信件。人们对于缺乏公众咨询和系统使用的透明度表示严重关切，但也对劳动和社会保障部为保护信件中包含的敏感个人信息所采取的措施表示关切(大赦国际,2025)^[19]；摊位, 2025^[20])

用于应对潜在滥用或错误的问责框架

问责框架对确保政府机构为人工智能系统的结果负责至关重要，包括错误、滥用或损害。它们明确了谁对人工智能做出的或受其影响的决策负责，确保在出现问题时能够追责。如果没有问责制，就很难纠正错误、从失败中学习或维持公众信任——尤其是在问题公开后。这些框架通过设定明确的角色、责任和后果，充当护栏，鼓励谨慎设计、监督和人工智能的合乎道德使用。最终，它们有助于确保公共部门的人工智能保持透明、公平，并与公共利益保持一致。尽管如此，数据显示这些框架的实施水平非常低，只有23%的参与欧盟成员和17%的参与非欧盟成员已经建立了这些框架。

例如，**法国**的“公共利益人工智能巴黎宪章”³

强调针对人工智能相关危害的防范措施，确保遵守国际人权法，并促进透明度。该国针对人工智能提案的“评估网格”也强调明确的角色和责任，以及

停用 AI 工具，并始终有人监督。**匈牙利** 人工智能战略通过人工智能联盟整合问责制，该联盟包括政府机构、学者和行业专家，以监督人工智能发展并防止其被滥用。**爱尔兰** has launched an “**人工智能标准与保障路线图**”，与欧盟人工智能法案保持一致，以确保人工智能的负责任部署。此外，该国已指定九个国家级机构来保护基本权利，通过独立监督加强问责制。

特别是在社会保障领域，如果缺乏问责制，存在着削弱公众信任并造成损害的风险。例如，荷兰实施了一个算法风险分类模型来检测儿童保育津贴的欺诈性申领，该模型根据国籍和简单错误（例如文件上缺少签名）等因素，为申请人创建了更高的欺诈风险档案。该系统还计划从追回的欺诈性申领中筹集资金，据称这激励了当局尽可能多地处理虚构欺诈案例。结果，数万名父母和护理人员失去了他们的福利，这引发了一场丑闻，导致荷兰内阁在2021年解散（大赦国际，2021）。^[21] 一个强有力的问责框架将有助于确保在系统实施之前，适当的检查和制衡机制得到建立，从而避免此类问题在未来发生，并且当问题确实出现时，也要确保对此类决策的问责制是清晰的。

数据安全原则

数据安全原则是公共部门负责任部署人工智能的基本防线，确保敏感信息的完整性、保密性和可用性得到保障。政府机构通常管理包含个人、财务和国家安全信息的大量数据集，其滥用或未经授权的访问可能产生深远影响。

数据安全除了技术实施之外，还包括风险管理协议、风险评估以及遵守相关法律法规等多个方面。然而，这显然是欧盟需要集中精力的领域，因为在参与的国家中，仅有32%的欧盟成员国报告已制定具体标准或协议，而与之相比，33%的非成员国则报告已制定这些标准或协议。这也可能表明成员国完全依赖欧盟的《通用数据保护条例》（GDPR）来设定其国家环境中的数据安全与隐私方法。这一点在**法国**，该方案以符合GDPR、数据最小化和使用SecNumCloud认证基础设施为优先事项，以确保安全的人工智能部署。**爱尔兰** 同样地，它将其人工智能政策与欧盟和经合组织的原则保持一致，嵌入隐私设计框架和强大的数据治理机制。尽管这与人工智能更为具体，**瑞典** 的生成式人工智能指南强调在公共行政中负责任地使用人工智能，涵盖数据保护、道德和合规等领域。

参与度——用户如何参与使用人工智能系统？

与多元利益相关者进行有意义的参与对于在社会安全领域负责任地使用人工智能至关重要。这可以包括用户研究（包括调查、访谈、焦点小组和用户旅程地图）、用户测试（包括可用性、模拟测试、A/B测试和反馈机制），以及其他用户参与机制，如参与式设计研讨会和关于人工智能使用的公共咨询。

随着人工智能技术越来越多地影响关系到人们生活的决策，让用户、受影响的社群、专家、公民社会及其他公共机构参与其中，有助于确保系统保持透明、公平，并与社会价值观相符。早期的参与有助于建立信任，增强合法性，并使人们能够更深入地理解潜在风险，特别是对弱势群体而言。它还有助于通过纳入广泛的观点来支持更好的政策设计，帮助识别挑战、防止偏见，并确保社会保障领域的AI系统具有包容性、问责制和以人为本。

在法国和意大利的社会保障领域，人工智能参与工作主要集中于内部用户，采取了试点测试、用户调查和培训等富有成效的做法。两国的社会保障机构都采取了措施，让工作人员参与评估和塑造人工智能工具。然而，在整个人工智能生命周期内的参与——特别是与外部用户和受影响社区——仍然有限。随着人工智能试点的扩大，有机会通过清晰的沟通策略来规范更具包容性的参与做法。与民间社会、专家和其他政府机构进行更广泛的咨询，也有助于解决非采用等问题，并制定以用户为中心的可扩展解决方案。

然而，这并非这些国家社会保障领域的独特现象。表2.7显示，在公共部门制定AI政策和应用AI用例方面，与不同参与者的参与度普遍较低。因此，需要采取更一致的方法，在AI系统的设计和开发各个阶段参与用户，以建立信任并确保系统有效。

表 2.7. 参与公共部门 AI 政策和用例开发的参与者

基于参与欧盟成员国与非欧盟国家（经合组织成员和伙伴）的国家级数据

	%欧盟	%非欧盟
公共部门组织	82% (18/22)	89% (16/18)
公务员	55% (12/22)	78% (14/18)
服务用户	36% (8/22)	56% (10/18)
更广泛的生态系统（私营部门，CSO）	65% (14/22)	56% (10/18)
跨境行为体（例如其他政府、多边组织）	32% (7/22)	39% (7/18)

注意：此数据基于22个欧盟成员国和18个非欧盟经合组织成员国及伙伴国的初步数据。来源：为2025年经合组织数字政府指数收集的数据。

与公共部门组织

与其他公共部门组织合作对于确保在发展人工智能赋能工具和监管框架方面采取协调一致、有效的方法至关重要。这有助于促进知识、资源和最佳实践的分发，避免重复和分散的努力。合作确保各机构之间的标准、政策和道德框架保持一致，使人工智能系统更加互操作和值得信赖，这对于共享和整合数据至关重要，这些数据可以支持人工智能系统以更有针对性的方式被用于解决应用不足的问题。它还增强了集体能力，并培养了一种负责任的创新文化。通过共同努力，政府可以构建更强大、更有效和更具凝聚力的AI。因此，公共部门组织参与度最高，82%的参与欧盟成员和89%的所有参与非成员报告称与这些参与者合作。

欧盟成员国尤其采取了一系列方法，让公共部门组织参与人工智能的发展和部署。**立陶宛**的 GovTech Lab是一个结构化、创新驱动型模式，利用沙盒环境和共创项目，用于让多个公共机构参与测试人工智能解决方案。**芬兰**和**瑞典**更倾向于代理驱动的模式，通过网络引导对人工智能的部门探索。**法国**通过DINUM协调的ALLiance孵化器，将各部委直接纳入AI开发，而**斯洛文尼亚**在数字化转型部的领导下，采用了一个广泛的跨部门治理框架，以促进系统性协调。这些做法说明了对于AI治理中积极公共部门参与的共同认识，但在结构上有所不同——从集中式协调到自下而上的创新生态系统——这取决于制度环境和战略成熟度。

带有公务员

在开发和实施人工智能赋能工具和监管框架时，让公务员参与至关重要，因为他们通常会以高频率使用、管理并受到人工智能系统的影响。他们的见解有助于确保人工智能工具有实用性、用户友好性，并符合真实的公共服务需求——特别是在社会保障的一线服务中。尽早让他们参与进来能够建立理解、信任和所有权，减少阻力并促进负责任的使用。这也有助于识别风险、空白和培训需求。通过将公务员纳入流程，政府可以设计出更具效率和用户友好性的人工智能系统，真正支持公共部门目标，以及确保其使用符合伦理所必需的检查点。因此，数据显示，与这些行为体的参与率很高，55%的参与欧盟成员国和78%的参与非欧盟成员国在人工智能政策和使用案例的开发中让公务员参与进来。

例如，**法国的** 联盟计划采用以用户为中心的方法，将公务员嵌入为“内部企业家”，他们共同设计和测试人工智能解决方案，并得到工具（例如）的支持 **数字续集**。

奥地利的 人工智能政策论坛促进跨部门合作，创建一个用于对话和共同实施人工智能计划的平台，包括与欧盟人工智能法案的协同。**芬兰** 为公务员提供结构化的多语言指导，关于生成式人工智能，确保公共管理中的清晰度和一致性。最后，**爱尔兰** 和 **葡萄牙** 通过试点项目和参与式治理流程应用实践性参与。这些做法共同强调为公务员配备战略洞察力和运营工具对于人工智能的重要性。这些工具的一个例子是葡萄牙社会保障机构提供的对“Copilot-IA”培训和软件许可的强有力投资，以便工人能够利用并适应这些工具到他们的日常工作中，以便更快、更有效地完成他们的任务。

与服务用户

在开发人工智能赋能工具和规范时，吸引最终用户至关重要，以确保系统满足实际需求、易于使用，并公平有效地运行。最终用户可以就潜在风险、可用性问题和非预期影响提供宝贵反馈。他们的参与有助于建立信任、改进系统设计，并确保人工智能工具支持而非阻碍公共服务提供。纳入他们的观点也有助于促进透明度和问责制。最终，吸引最终用户将带来更实用、更具包容性和更负责任的人工智能系统，这些系统能更好地服务公众并符合共同价值观。尽管如此，与用户互动的比率是最低的，36%的参与欧盟成员和56%的参与非成员在人工智能系统的设计和开发中与他们互动——但这也可以是一个指标，表明目前公共部门的大多数人工智能用例仍处于试点阶段，目前可能仍侧重于内部测试流程。然而，应该注意的是，在解决方案的识别、设计、开发、测试和监控的每个阶段与用户互动都是关键，以确保人工智能系统满足他们的需求和期望。

在那些制定了与服务使用者互动的实践措施的人当中 **德国的** BA涉及用户参与ADEST开发的初步阶段。AI模型已被内部专家训练，系统训练方式的内部文档也已公开。试点解决方案与全国多个地方就业机构进行了一系列迭代测试，并在项目初期收集了持续的反馈。为了增强组织内部的透明度，AI团队收集了关于用户对ADEST认知和理解的反馈，以识别如何更好地支持用户。

法国 也通过在项目早期阶段进行访谈和迭代设计过程来整合用户反馈，以确保解决方案满足实际需求。**葡萄牙** 通过参与式平台如 *participa.gov.pt*，使公众能够对国家人工智能计划提出意见。这些模型突显了公众信任和可用性对于人工智能成功应用日益增长的认识。

随着更广泛的生态系统

与更广泛生态系统互动——包括私营部门和民间社会——对于政府在政府中开发有效的AI赋能者和护栏至关重要。这些利益相关者提供了多样化的专业知识、创新和观点，有助于识别风险、解决道德问题并确保负责任的使用。私营部门贡献了技术知识和工具，而民间社会则确保公共价值、权利和包容性得到优先考虑。合作促进了透明度、问责制和公众信任。通过让更广泛的生态系统参与进来，政府可以制定更平衡、更具前瞻性的AI政策，反映社会需求，并在公共部门促进AI的安全、合乎道德和有效使用。然而，数据显示，与这些行为者的参与程度较低，64%的参与欧盟成员和56%的非成员报告称在过程中涉及这些部门。

然而，欧盟中有国家正采用更广泛的人工智能生态系统——包括私营部门、民间社会和科研机构——来促进公共部门人工智能的创新、问责制和信任。**比利时的 AI4Belgium**作为一个全面的多利益相关方生态系统，促进政府部门、行业、民间社会和国际合作伙伴之间的持续合作。在**瑞典**，国家人工智能委员会提供一个结构化的咨询平台，汇集不同利益相关者以影响国家战略。**法国**整合初创企业、研究机构以及开源贡献者进入人工智能开发，支持技术进步和伦理监督。**捷克**和**匈牙利**通过战略咨询机构和跨部门协调机制促进生态系统参与。最后，**爱尔兰**通过应用实验和多参与者参与，鼓励合作，在其人工智能概念验证项目中。总体而言，这些方法强调了生态系统参与的重要性，以确保公共部门的人工智能具有创新性、包容性，并与社会价值观保持一致。

跨境行为体（例如其他政府、多边组织）

与跨境行为者——如其他政府、多边组织和国际专家机构——的互动对于促进公共部门中人工智能的负责任、可互操作和安全的运用也至关重要。合作参与使监管方法的协调、技术工具和政策最佳实践的交流以及共同治理框架的开发成为可能。它还促进了与国际研究网络、能力建设计划以及联合基础设施投资的途径。在一个快速变化的技术环境中，跨境合作增强了韧性，减少碎片化，并支持与民主价值观和国际法律标准一致的值得信赖的人工智能系统的开发。然而，这些行为者的整体参与率最低，其中32%的参与欧盟成员国和39%的参与非欧盟成员国在人工智能政策和用例的开发上超越其边境进行参与。

然而，在整个欧盟范围内都存在着这类合作的机会。例如，欧洲高性能计算联合企业（EuroHPC JU）是一个联合倡议。**欧盟**，欧洲国家以及私营伙伴，以开发超级计算基础设施并支持研究。EuroHPC JU将为大学和中型企业提供低成本的基础设施，用于算法开发。此外，**欧盟**资助机制——例如“数字欧洲计划”——旨在将欧盟成员国组成的联盟聚集起来，共同开发共享的人工智能解决方案（欧洲委员会，2025）^[22]。最后，像**欧洲社会保险平台**帮助将社会保障机构聚集在一起分享经验、制定共同方法，并确定合作的机会。

关键发现

要充分发挥可信人工智能在社会安全领域的潜力，强有力的治理至关重要，以确保其发展和应用具有战略性、有效性和可信性。为了支持社会安全领域的可信人工智能，各国必须从零散的实验转向更具战略性、包容性和问责性的方法。加强支持因素、嵌入安全机制，并扩大参与——尤其是与公众的参与——对于确保人工智能系统不仅有效，而且公平并值得公众信任将是至关重要的。

基于对欧盟各国的近期分析，包括来自法国和意大利的详细案例研究，强有力的治理正在显现，但在使能者、护栏和参与支柱方面却是不均衡的：

国家在实施战略人工智能应用的赋能方面更强，中央政府和社会安全机构之间有一种协同方法：

- 各国正在制定战略和治理框架以支持人工智能的采用，尽管方法各异。无论是由中央政府还是社会保障机构领导，清晰的理念、支持性环境以及明确定义的职责对于成功至关重要。
- 战略投资仍然分散。然而，能够支持社会保障领域更专注的人工智能发展的专项资金和采购模式正在涌现。
- 基础设施和数据进展参差不齐。虽然可扩展计算正在发展，但需要更多地围绕数据治理和其他关键的支持性基础设施进行努力，这些基础设施对于在社会保护中有效使用人工智能至关重要。

为安全及合乎道德地使用人工智能而设立的护栏正在涌现，但可以通过在社会安全领域实施欧盟人工智能法案的要求得到加强：

- 该法案为人工智能监管提供了全面框架，各国开始应用。这在社会安全领域尤为重要，因为人工智能系统可能被归类为高风险，因此在法案下有更多要求进行主动监管。
- 伦理框架正在建立，但挑战在于将伦理考量贯穿人工智能的整个生命周期，包括设计、部署和持续监督。
- 在各国准备履行《法案》的同时，已在一些实施透明度、审计和合规机制方面取得进展，以满足欧盟要求。这对于确保人工智能不会损害依赖社会保障体系的人来说至关重要。因此，各国应在《法案》的全面要求生效之前，继续集中精力广泛实施这些机制。
- 在国家层面，关于数据隐私和安全的相关实践报告相对较少，这可能是由于GDPR已经提供了强大的保护。

超越政府层面的参与可以加以改进，以建立信任并交付以用户为中心的人工智能系统：

- 与公务员和公共部门机构的内部合作相对紧密，这有助于确保人工智能工具在政府运营中具有实用性和适用性。
- 在设计、开发和监控阶段，与服务用户的互动可以进一步加强。这至关重要，尤其是在社会保障领域，人工智能直接影响公民获得福利和服务的途径。
- 跨部门、跨境合作可以改进，以最大限度地扩大推广良好实践和通过共享学习和合作加速创新的机会。

参考文献

- AGID (2025), 人工智能, <https://www.agid.gov.it/it/ambiti-intervento/intelligenza-artificiale> [15]
- agid (2023), 关于数据开放和再利用的技术规则指南
公共部门信息, https://www.agid.gov.it/sites/default/files/repository_files/lg-open-data_v.1.0_1.pdf 访问
在2025年6月24日)。
- 大赦国际 (2025), 技术太多, 同理心不足, <https://www.amnesty.org/en/documents/eur45/9478/2025/en/> (于
2025年9月17日)。
- 大赦国际 (2023) 塞尔维亚: 陷入自动化: 贫困与歧视
塞尔维亚的福利国家, <https://www.amnesty.org/en/documents/eur70/7443/2023/en/>。
- 大赦国际 (2021), 排外机器: 通过不受监管的使用进行歧视 [21]
荷兰儿童保育津贴丑闻中的算法, <https://www.amnesty.org/en/documents/eur35/4686/2021/en/>。
- 布思, R. (2025) 对DWP使用人工智能读取福利相关信函表示“严重关切” [20]
主张人, <https://www.theguardian.com/society/2025/jan/27/dwp-ai-whitemail-benefit-claimants-applicants>。
- 欧洲委员会 (2025) 数字欧洲计划 - 项目提案征集, [22]
https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/digital/wp-call/2025/call-fiche_digital-2025-ai-08_en.pdf。
- 欧盟 (2024) 欧洲议会和欧盟理事会法规 (EU) 2024/1689 [13]
2024年6月13日理事会决定制定有关人工智能的统一规则并修订法规 (EC) 第300/2008号, (EU) 第167/2013号,
(EU) 第168/2013号, (EU) 2018/858号, (EU) 2018/1139号, <http://data.europa.eu/eli/reg/2024/1689/oj>。
- Fjeld, J.等(2020), “原则性人工智能: 绘制伦理和
基于权利的方法: AI 原则” SSRN 电子期刊, <https://doi.org/10.2139/ssrn.3518482>。
- 李, D. (2021), data.gov.ie 开放数据发布指南, <https://data.gov.ie/guidelines/index.html> (2025年6月24日访问)。
- 穆罕默德, S. 等. (2025), “数据质量对机器学习性能的影响
表格数据” 信息系统, 第 132 卷, 第 102549 页,
<https://doi.org/10.1016/j.is.2025.102549>。
- OECD (2025), 人工智能治国: 核心领域的现状与未来方向
政府职能, 经合组织出版商, 巴黎, <https://doi.org/10.1787/795de142-en>。
- 经合组织 (2023年), “2023年经合组织开放、有用和可重用数据(OURdata)指数: 结果和
关键发现”, 经合组织公共治理政策论文, 第43号, 经济合作与发展组织出版局, 巴黎,
<https://doi.org/10.1787/a37f51c3-en>。

经济合作与发展组织 (2022年), “经济合作与发展组织人工智能系统分类框架”, *经合组织数字经济* [11] 论文, 第323号, 经济合作与发展组织出版社, 巴黎, <https://doi.org/10.1787/cb6d9eca-en> .

经合组织 (2021年), “经合组织人工智能原则实施情况: 各国人工智能的见解” [12] 政策”, *经济合作与发展组织数字经济论文*, 第311号, 经济合作与发展组织出版, 巴黎, <https://doi.org/10.1787/1cd40c44-en> .

OECD (2019), *成为数据驱动型公共部门之路*, 经合组织数字政府 [4] 研究, 经合组织出版物, 巴黎, <https://doi.org/10.1787/059814a7-en> .

OECD (2017), *理事会对增强数据获取和共享的推荐*, [9]

经合组织 (即将出版), “2025年经合组织数字政府指数”, 经合组织出版社, 巴黎。 [2]

OECD (未发表), *差距分析报告 - 利用人工智能解决社会福利未领取问题* [3] 在法国和意大利 .

芬兰统计局 (2022), *国家数据质量标准与指标*, [8] [https://stat.fi/media/uploads/org/tilastokeskus/tiedonlaatu/data_quality_criteria_and_indicators .pdf](https://stat.fi/media/uploads/org/tilastokeskus/tiedonlaatu/data_quality_criteria_and_indicators.pdf) (2025年6月24日访问)。

加拿大财政部秘书处 (2025), *微软员工Copilot (浏览器聊天)* [16] 基于应用程序)政策实施通知, <https://www.canada.ca/en/government/system/digital-government/policies-standards/microsoft-copilot-for-work-policy-implementation-notice.html> .

乌得勒支大学 (2025) *道德数据助手 (DEDA)*, <https://deda.dataschool.nl/> . [17]

注释

¹ 欧盟成员国包括奥地利、比利时、克罗地亚、捷克、丹麦、爱沙尼亚、芬兰、法国、希腊、匈牙利、爱尔兰、意大利、拉脱维亚、立陶宛、卢森堡、荷兰、葡萄牙、罗马尼亚、斯洛伐克、斯洛文尼亚、西班牙、瑞典。非欧盟成员国包括阿根廷、澳大利亚、巴西、加拿大、智利、哥伦比亚、哥斯达黎加、冰岛、以色列、日本、韩国、墨西哥、新西兰、挪威、秘鲁、瑞士、土耳其和英国。

² 看 <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses> .

³ 看 <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/the-paris-charter-on-artificial-intelligence-in-the-public-interest> .

3 构建人工智能预备劳动力

人工智能在社会保障组织对劳动力意味着什么？本章讨论了人工智能对领域内工作及技能的影响。它还探讨了通过招聘和培训来为劳动力做准备和提高内部人工智能能力的良好方法。借鉴经合组织社会保障机构和行政部门的经验，本章强调了支持有效和可信赖地使用人工智能以符合组织目标的劳动力战略。

公共部门组织对人工智能的采用不仅是技术问题，更重要的是，是员工发展、实施和使用人工智能工具的问题。成功的人工智能采用需要具备正确技能、了解人工智能的益处和风险，并协同合作改进其机构向公众提供价值的方式的人员。

人工智能工具在信息处理等认知任务上的性能不断提升。这引发了关于人工智能应用对公共部门劳动力带来的变化的问题。人工智能将对劳动力需求和技能产生影响，因为它可以增强人类工作、自动化特定的行政任务，并转变角色和组织流程。为了有效地、负责任地并以符合机构目标的方式利用人工智能的益处，公共机构需要将人工智能实施与战略劳动力发展相协调。

经合组织人工智能原则指出政府的关键作用，即“使人们能够有效地使用和与人工智能原则互动[.....]，包括为他们提供必要的技能”。劳动力发展战略——涵盖招聘、外包和培训——应是构建人工智能能力的制度方法的重要组成部分。社会保障机构可以通过招聘、外包和培训这三个主要杠杆来实现这一目标。稀缺技术人才的招聘具有挑战性，尽管外包是政府机构的一种常见做法，但它存在需要管理的风险。

最终因此，通过有针对性的技能发展和有利于创新和持续学习的组织文化来建立内部能力，可能会带来人工智能解决方案效率更高、问责制更强以及与公共价值的更好一致性等好处。培训是政府机构建立能够利用人工智能的益处和管理人工智能风险的人工智能人才队伍的重要杠杆。在这种背景下，对技能的理解遵循Mergel等人（2023^[1]）它超越了技术和横向技能，涵盖了技术趋势、政策、法律和伦理问题以及管理实践的意识与知识。

本章首先从公共部门人工智能对技能和劳动力影响的文献综述开始。其次，它讨论了支持公共行政机构吸纳所需技能、建设能力的招聘和外包国际良好实践。第三，它基于利益相关者访谈，讨论了公共机构内部建设人工智能能力的培训计划案例。最后，本章就支持人工智能应用以改善公共服务的人力资源发展战略提出了更广泛的结论。

理解人工智能对劳动力及技能的影响

随着经合组织各国的社会安全机构和公共行政机构采用人工智能工具，这对他们的员工及其所需要的能力产生了影响。事实上，越来越多的证据表明，人工智能可以帮助个人在工作中更快地完成某些任务并产出更好的结果（Czarnitzki, Fernández和Rammer, 2023^[2]；布林约夫森、李和雷蒙德, 2023^[3]）。目前，人工智能似乎更可能自动化繁琐、重复的任务。私营部门的员工调查表明，因此，工作的速度以及员工的表现和享受都增加了（Lane, Williams和Broecke, 2023^[4]）。与此同时，人工智能工具的能力持续提升，包括在处理非常规认知任务（如信息处理）（经合组织, 2023^[5]）。

现有关于人工智能对公共部门劳动力潜在影响的研究有限（Mergel等人, 2023年^[1]），但关于人工智能和劳动力市场的一般文献提供了重要的见解。在整体经济中，虽然许多工作暴露于人工智能的高风险中，但目前由于人工智能导致大规模失业的证据仍然是有限的（Acemoglu等人, 2022^[6]；经济合作与发展组织, 2023^[5]；Gmyrek等, 2025^[7]）。重要的是，自动化某些任务的技术可行性可能不会立即转化为实际应用，因为组织会评估引入人工智能的成本、收益、风险和监管障碍。社会保障机构对新技术的审查非常严格，而且它们在重新分配人力资源方面也比

然而，从现有文献中关于人工智能对社会安全组织劳动力影响的研究，可以得出三个主要结论。

可以通过采用人工智能来提高公共部门的生产力。

社会保障机构是在所有经合组织国家中的主要雇主，其中许多人面临人员短缺、工作量大以及财政压力的问题。其核心职能包括行政和支持任务，例如处理福利索赔或征收标准化贡献。因此，人工智能的采用对减轻行政负担、释放时间和精力以从事更有趣和更高价值任务所产生的潜在影响是实质性的（Mergel等人，2023）^[1]；阿特金森和布朗，2024^[8]）。在公共部门内重新分配工作角色和人力资源可能具有挑战性，并且可能需要更长的时间。

人工智能确实已经应用于公共部门组织内部的特定任务和流程自动化。本报告第一章中的例子富有启发性。德国联邦通讯社最近委托进行的一项研究（*联邦印刷局*）表明，即使在政府内部，相当不同的行政程序也可以分解为相似的流程组件，这些组件可以通过技术来增强。特别是，基于规则的行政程序在应用数据和人工智能解决方案方面具有加速的潜力（Bundesdruckerei和Possible，2025）^[9]）。

一个有趣的例子是芬兰国家社会保险机构Kela利用人工智能自动化处理与福利申请相关的附件接收和处理。他们的AI平台通过文字识别、图像校正、格式转换等，在单一工作流程中标准化多种不同类型的文件，为行政案件工作者节省时间（另见第一章）。

大多数利益相关者指出，某些任务或流程的完全自动化目前尚不可能，因为任何结果或决策都需要人为批准。在大多数机构中，明确决策的责任归属至关重要，以避免技术故障、滥用、偏见或人工智能使用缺乏透明性带来的风险。

因此，机构面临的挑战不仅在于确定人工智能的应用在哪里以及如何应用有意义，还在于理解工作流程可能需要如何改变才能更充分地利用人工智能技术的优势。换句话说，狭隘地将人工智能应用于现有任务将无法充分发挥该技术为更有效和更好的服务所带来的潜力。

本报告采访的大多数机构似乎都意识到人工智能可能为公共服务和福利提供带来的巨大生产力提升。随着人工智能系统能力的不断提高以及政府在建立必要的AI和数据治理以及基础设施方面所做的努力，以及为员工采用人工智能做准备，这种潜力可能会进一步增长。

人工智能改变了公共行政中的技能需求。

技术采纳，包括人工智能，增加了对专业化数字技能的需求。与私营部门一样，公共机构也面临着对专业化数字和数据专业人才日益增长的需求，这些人才在数据科学和机器学习等领域拥有先进的技术技能。为了应对这些需求，劳动力培训、技能再培养以及招聘是应对措施的重要组成部分（Chinn等人，2020）^[10]）。这是一个挑战，因为经合组织中的许多政府机构在与私营部门竞争以吸引和留住合格的数字和数据专业人员（法国政府、DINUM和INSEE，2021年）^[11]；梅达吉，米卡莱夫和唐吉，2024^[12]）。同时，经合组织的研究发现，尽管对专业AI技能的需求在增长，但它仍然只影响极少部分的工种（格林，2024）^[13]；borgonovi等人，2023^[14]）。

除了少数专门从事数字和数据领域的专业人员外，社会保险机构的绝大多数员工将不需要与人工智能相关的高级技术技能。然而，人工智能的采用将改变这些员工的任务、他们工作的组织方式以及执行工作所需的技能（Green，2024）。^[13]；经济合作与发展组织，2023^[5] 大多数员工都需要具备能够有效且负责任地使用和与工作中的人工智能工具的能力，这在欧盟的《人工智能法案》中被称为人工智能素养。人工智能技能的关键基础是扎实的基本数字技能，这在一些公共部门机构中是一项挑战。此外，所有员工群体都需要持续学习，以获得关于人工智能如何运作及其局限性的一些基础知识。

各级员工将日益需要广泛而互补的技能，以配合人工智能系统，特别是在高级认知技能方面，如解决问题能力或批判性思维能力，以及横向技能，如社交能力、管理能力或团队合作能力（Theben、Plamenova和Freire，2023）^[15]；经济合作与发展组织，2023^[5]）。这些互补能力不仅将允许人工智能的采用和有效使用，还将使员工能够参与持续学习，适应和塑造工作组织的变化，并找到创新的方法使用人工智能工具来支持他们实现其机构的目标。

aitv作为奥地利社会保障机构的信息技术服务商所开发的ai解决方案的实例，突出了任务内容和技能需求可能发生的变化。aitv开发了一个ai工具来支持流程索赔管理，该工具自动化了简单常规索赔的管理。ai工具的引入产生了次要效应，即主要将复杂案件留给行政案件工作人员，并且员工认为他们的工作难度因此增加了。在这种情况下，该ai解决方案的价值可能部分被不断变化的技能要求和日益增加的任务难度所掩盖。

最后，公共机构中的领导者与决策者在推动战略人工智能采用，同时保持问责制、管理风险、确保合规和保护数据方面发挥着关键作用。优秀的领导者对于确保人工智能技术的有效和负责任的应用至关重要（Valle-Cruz, García-Contreras和Patricia Muñoz-Chávez，2024）^[16]）。他们的技能、态度和参与度对于建立学习和创新文化、重组组织工作流程以及采用符合其机构目标的AI至关重要。摒弃传统的自上而下的管理角色，领导者将越来越需要鼓励员工的发展、卓越和自主性（经合组织，2023）。^[17]）。

公共机构需要积极管理人工智能的使用及其风险。

第二章论证了在公共服务机构中发展强大的AI治理的重要性，既包括促进AI应用的实施者，也包括防范风险的护栏，以及确保与用户参与的方式。考虑到AI应用的日益广泛，政府需要加强保护其免受风险机制，例如伦理宪章、审计机制和AI监管机构，或问责框架。特定的一组风险与直接员工的不当使用AI有关。鉴于开放获取的genAI工具的广泛性，例如，公共服务机构需要积极和战略性地管理其对AI的机构方法，包括员工个人层面的AI使用。

证据表明许多公务员已经使用AI工具 其中最常见的是通用型和开放获取的genAI工具 如ChatGPT (Gillespie等人，2025年)^[18]）。一项针对英国教育、医疗、社会工作及应急服务领域公务员的最新调查发现，22%的受访者在工作中使用GenAI，而且大多数人是每天或每周使用（Bright等人，2025年）。^[19]）。随着人工智能工具的普及，这类数字可能会随着时间的推移而增加。一项针对加拿大公务员的早期研究发现，有11%的受访者报告称使用过GenAI工作（全球政府论坛，2023^[20]）而私有部门员工中GenAI的使用可能略为先进，但公共部门领导者可能低估了AI已经在他们的组织中扮演的角色。

根据一项大规模的具有代表性的全球调查，报告在工作中使用AI的受访者中有半数承认违反组织政策使用AI，例如将敏感信息上传到公共AI工具（Gillespie等人，2025）^[18]。在社会保障组织中，考虑到他们处理的敏感个人数据，这种对GenAI的滥用对它们服务的个人构成了重大风险。一些公共机构对此做出了反应，例如，在美国社会保障管理局（FedScoop，2024年），暂时封锁了政府设备上的开放访问AI系统。^[21]，而其他人则为其员工介绍了关于使用该技术的指南或培训，明确警告员工无权将任何机密或私人信息与 ChatGPT 等开放获取的 AI 工具分享。先进的公共部门组织通过引入基于内部内容训练的自有大型语言模型（LLM）采取了积极行动，丹麦格拉斯克斯市就运行了一个基于 GPT 技术的自托管 LLM（Medaglia, Mikalef and Tangi, 2024^[12]）。

关于人工智能对劳动力技能影响以及公共行政部门当前人工智能使用实践的证据，突显了双重需求。一方面，组织越来越需要在劳动力中具备与人工智能相关的核心技术和辅助技能与能力。另一方面，它们需要采取有意识的方法来规范员工的人工智能使用，以符合人工智能和数据方面的法规。对人工智能使用的有意识方法和积极管理将有助于避免数据隐私、偏见或缺乏透明度的风险，但它也将避免员工群体之间的差距。一些员工可能会独立有效地和安全地开始采用人工智能，而其他人则需要量身定制的技能发展机会、明确的行为规范和清晰的激励措施，以开始在工作中使用人工智能工具。

社会保障领域与AI相关的技能需求

随着他们构建人工智能能力，社会保险机构需要一支了解使用人工智能的益处和风险的劳动力队伍。依靠 *经合组织公共部门数字人才与技能框架* (2021^[22])，以及艾伦·图灵研究所的企业AI技能能力框架 (2024^[23])，可以区分三个关键劳动力群体来定义采用人工智能系统所需的技能。

- **普通员工** 需要能够以负责任和有效的方式与使用人工智能工具进行交互。
- **领袖** 需要战略性理解人工智能技术，包括其对于组织目标的潜力和风险。
- **数字和数据专业人员** 需要支持人工智能系统开发、实施和维护的技能。

其中的每个目标群体都有不同的学习需求，针对他们的培训内容可被描述为基础性、战略性或技术性。学习需求基于他们在机构中的角色和任务，以及他们与人工智能采用相关以完成这些任务所需的技能（见表3.1）。

表3.1. 目标群体相关的AI技能需求

目标群体	类型	任务	技能与知识
普通员工	基础性	与并使用人工智能负责地使用工具有效方法	• 基础数字技能 • 人工智能新兴知识与意识 • 技术、其潜力、风险和伦理使用 • 数据保护的必要意识 • 批判性思维与独立判断
领袖	战略	采取战略决策 针对人工智能的采用行为 technology	• 关于新兴人工智能技术及其作用的相关知识 • 实现组织目标 • 战略性地考虑风险、伦理的能力 • 与AI相关的监管方面

目标群体	类型	任务	技能与知识
			- 实施人工智能的战略思维和能力建筑（人工智能治理、数据基础设施和劳动力准备状态） - 强强合作，沟通，利益相关者管理和变革管理技能
数字和数据专业人士	技术	支持开发实现，和维护AI系统	- 其专业领域的技术技能（例如：数据科学） - 关于伦理和监管的详细知识围绕人工智能的考量 - 技术技能以减轻风险和管理人工智能系统的复杂性 - 协作与沟通技能以在跨学科团队，并支持领导和整个组织

来源：经合组织参考阿兰·图灵研究所的《商业人工智能技能能力框架》（2024[23]）。

本报告中受访的机构均未以结构化的方式分析其人工智能采用和使用所需技能，并且均未建立人工智能技能框架。然而，一些机构已经开始使用涵盖人工智能元素的数字技能框架。例如，为近63 000名公务员提供培训的公共管理学院的拉脱维亚数字学院，与外部供应商合作开发了一个数字技能框架，以加强公务员的数字技能。该框架基于欧洲数字技能参考框架DigComp。¹

拉脱维亚框架包含六个能力领域：信息与数据素养、数字沟通与协作、数字内容开发、网络安全、数字化战略与进步，以及横贯能力。针对每个能力，它提出了从入门到专家四个技能水平。其中一项能力特别涉及人工智能的使用：在入门水平，员工知道人工智能是什么，理解其基本概念及其在工作中的潜在应用。在基础水平，员工能够创建高质量的AI提示，批判性地评估AI输出，了解人工智能对其工作质量和效率的影响以及数据安全要求。在高级水平，他们能够为特定需求开发或调整人工智能，在专家水平，他们推广人工智能的合乎道德和负责任的使用，关注人工智能工具的发展趋势，规划其在机构中的使用，并组建能够使用人工智能的项目团队。该框架是一个基础工具，旨在确保公务员具备使用数字技术（包括人工智能）所需的数字和横贯能力。目前，它在拉脱维亚的几个地区进行试点。一旦验证通过，它将用于评估员工的技能并识别个人的培训需求。

通过招聘和战略合作伙伴关系建立内部人工智能能力

社会保障机构与其他组织一样，有三项主要手段来整合新技能并为人工智能的采用建立能力：招聘、外包和培训。与私营部门组织相比，政府机构的招聘流程对人工智能的采用和使用相关的短期技能和能力差距的反应不够灵活。同时，招聘可以战略性地遵循劳动力队伍和组织发展的长期目标。公共部门的工作保障通常很高，这既为机构投资和激励员工技能发展提供了机会，也为其带来了义务。对许多政府机构来说，吸引和留住优秀候选人是一项挑战，特别是在ICT等技能短缺领域（欧洲基金，2023）。^[24]。为了解决这些问题，各国政府采取了不同的方法来创建有吸引力、高影响力的

政府和数字数据专业人员的机遇，以及改善该职业群体的条件和职业机遇。

外包是政府机构（包括社会保障机构）的标准替代方案，用以应对技能缺乏或内部能力不足，并实施新的AI解决方案。然而，将某些任务外包的决定需要权衡作为公共服务提供者的问责责任。如第二章所述，在内部保留关键技能和职能，并战略性地管理第三方供应商具有一系列优势。在AI应用背景下，需要更加重视保证数据隐私和安全，确保合规和问责，以迭代方式整合AI工具以支持组织目标，并在采购中避免信息不对称和供应商锁定（Kupi, Jankin和Hammerschmid, 2022）^[25]；梅达吉，米卡莱夫和唐吉，2024^[12]）。这些考虑也体现在芬兰国家社会保障机构的战略选择中，如第一章所述，其倾向于建设高度自主的技术基础设施。公共部门管理采购的一种方式是通过GovTech项目，这些项目是与初创企业、公共或私人创新单位或学术界的战略合作，以支持创新的数字和人工智能解决方案（OECD, 2024）^[26]）。

提供具有高影响力的项目以吸引技术人才

一些政府设计了针对数字和数据专业人员（通常是高技能毕业生或年轻专业人士）的项目，并提供有吸引力、高影响力的机会。例如，德国联邦政府的数字服务推出的Tech4Germany和Work4Germany奖学金，将私营部门专家和联邦部门及机构的专家聚集在一起参与转型项目。该竞争性奖学金将个人安排到跨学科团队中，在几个月内采用敏捷和以用户为中心的方法处理特定项目。Tech4Germany专注于产品、设计和工程专业人员，要求他们在三个月内创建原型软件产品；而Work4Germany专注于数字变革管理、工作流程的重新设计和迭代工作方法（数字服务，2025）^[27]）。

自2012年起运营，总统创新研究员（PIF）是美国一个类似的、竞争激烈的研究员项目，该项目在12个月内将私营部门成功的技术领导者与联邦政府的机构配对，在高规格的转型项目上合作。研究员是敏捷的、跨学科团队的一部分，该团队在产品设计、软件工程或其他数字和数据领域拥有专业知识。过去的项目包括，例如，改善疾病控制预防中心的数据使用和技术能力，或加强医疗保健生态系统内（美国通用服务管理局）医疗保险与医疗补助服务中心的数据共享（美国通用服务管理局，2025）^[28]）。

在法国，beta.gouv.fr被设立为一个孵化器项目，旨在将雄心勃勃的数字和数据专业人士与法国公共行政领域内具有高影响力且有时间限制的数字化转型项目相连接。被指定为“国家初创企业”，它的目标是将公共和私营部门的人员吸引到各个拥有特定使命的团队中。Beta.gouv.fr还帮助机构建立它们自己的数字公共服务孵化器，并在政府内部建立专家网络。该项目遵循以用户为中心的方法，注重成果，其管理旨在基于团队（信任和自主权）（法兰西共和国，2025）^[29]）。

这些举措都依赖于对具备高数字和数据技能的竞争性选拔的专业人士。它们吸引具有目标驱动使命的个人，并且通常在选拔的学者之间促进合作和团队合作。通过将个人置于具有潜在重大影响的时间限制型项目中，它们推动公共部门内的创新和技术采用。

提升数字和数据专业人员的职业发展路径

社会保障机构通常通过国家级行政层面的倡议，受益于各种举措，使公共部门对数字和数据专业人才更具吸引力，并优先考虑数字和数据技能。这包括专门职位的定义、技能框架和职业发展机会。虽然通常没有明确关注人工智能，但这些方法是构建人工智能应用工作能力的一部分。

上述拉脱维亚数字学院开发的数字技能框架，例如，包括有关数据分析师、ICT安全经理、服务开发者、AI专家、网络安全专业人士和数字领导者等各种关键角色以及项目经理、政策制定者、沟通专业人士、客户服务专业人士、律师和人力资源管理专家等支持角色的技能要求信息。有反思，可能会使用该框架来指导未来数字和数据专业人士绩效评估系统开发及学习路径构建。

英国政府数字化和数据专业能力框架是一个于2017年引入的技能框架，旨在提高公务员队伍中数字化专业人员的招聘、保留和技能发展。该框架定义了政府机构执行的核心数字化功能相关的160多项技能。这些技能可以配置为六个职业系列中的43个职业角色：架构、数据、IT运营、产品与交付、质量保证测试、软件开发和以用户为中心的设计。内阁办公室内的中央数字化和数据办公室（CDDO）根据专家反馈管理和定期更新该框架（Gov.uk, 2025）。^[30]；布尔特施尔，钢琴和韦尔比，2024^[31]）。

澳大利亚政府已采取不同措施来吸引数字和数据专业人才，并增强其数字能力。自2019年以来，澳大利亚公共服务（APS）职业已开发出具体职业系列，包括“数字职业”和“数据职业”，以更好地定义这些领域的职业发展路径，并促进专业人才的流动、培训和招聘。澳大利亚公共服务委员会（APSC）负责进一步制定技能战略、培训计划以及数据驱动的劳动力规划（澳大利亚政府，2023）^[32]。它依赖于整个国家的SFIA（信息时代技能框架）许可证。APS职业路径图——一个面向公众的应用程序——利用SFIA帮助机构和工作寻求者了解数字、数据和网络安全角色的职责和技能要求（澳大利亚政府，2025）^[33]。公共服务委员会与部门协商确定角色和技能要求，这些数据现在也用于采购数字和数据任务。澳大利亚政府还采用了《2025-30年政府数据、数字和网络安全人才规划》，该规划符合政府的更广泛战略，并定义了一种协调的方法来吸引、培养和留住这些领域的人才，加强公共部门数字化能力，并创建实践社区。

建设 GovTech 项目

公技指的是公共部门与初创企业、创新型中小企业、学术界和研究中心合作开发创新技术，以改善公共服务（经合组织，2024年）^[26]）。在这个模型中，与外包相反，政府不依赖于大型技术提供商，并且可以访问沙盒来测试新解决方案。对欧洲GovTech的研究表明，项目可能位于政府中心，政府机构内，或作为独立、非营利组织存在。它们的作用从促进者到公共部门创新的中心节点不一。常见活动包括公开竞赛、黑客马拉松、加速器计划、试点实验、研发拨款以及同行学习网络（Kuziemski等人，2022）^[34]）。

法国农业互助社 (MSA)，作为负责法国农业领域社会安全和医疗保险的公共机构，战略性地与大学和研究机构合作，其中包括技术研究机构IRT圣埃克苏佩里和图卢兹人工智能与自然智能研究所ANITI。这些合作支持与研究人员在特定人工智能用例方面的合作，并可能有助于吸引人才。

在奥地利，促进创新公共采购 (PPPI) 服务中心连接了公共机构和私营公司。2013年，它作为联邦采购局内一个小组成立，提供信息资源和同行学习网络、为政府合作伙伴提供咨询服务，以及一个创新采购平台。它还在其网站上发布公开竞争，邀请私营公司提交提案以解决公共部门中的 (AI) 挑战，例如支持拨款申请处理的AI解决方案 (IOB, 2025)。^[35]；库齐姆斯基等, 2022^[34])。

2021年成立，德国公技园区是一个注册的非营利组织，专注于连接政府机构与科技公司，以开发和扩展数字解决方案。该组织的董事会包括国家、联邦和地区各级政府的代表。其成员包括IT或网络安全公司、学术机构、初创企业和社会组织参与者。活动以创建交流和创新平台以及学习生态系统为中心 (另见下一节) (GovTech Campus Deutschland, 2025)^[36])。

gove技术合作也在地区层面进行。通过2020年的“催化剂项目”，埃塞克斯大学分析学与数据科学研究所与埃塞克斯郡和萨福克郡的郡议会合作，提供预测性大数据分析，以改善对弱势群体的服务。自2019年以来，该大学还任命了埃塞克斯郡议会的首席科学顾问，旨在提高数据驱动政府服务 (埃塞克斯大学, 2025)^[37])。

通过培训发展人工智能相关技能

在社保机构中培训人工智能至关重要。如前所述，招聘新员工存在局限性和挑战，外包核心职能也存在风险。公共行政领域 (包括社保领域) 的高工作稳定性意味着机构必须随着时间的推移投资于现有员工的技能发展，并创造一个有利于持续学习的环境。在此背景下，技能发展机会是建立内部人工智能能力的战略杠杆。

通用人工智能工具的广泛可用性和使用意味着公共部门组织有强烈的动力来确保员工具备使用人工智能工具有效且可信赖的方式所需的技能和知识。如果并非所有员工都需要对使用人工智能工具的潜力、局限性和风险有基本的了解，并基于良好的数字技能基础。

自2025年2月起，根据欧盟人工智能法案第4条，在欧盟提供或部署人工智能系统的组织，包括公共管理机构，必须依法确保其员工具备“足够的人工智能素养水平”。

人工智能系统的提供者和部署者应当采取措施，尽其所能确保其员工及其他代表其操作和使用人工智能系统的人员具备足够的AI素养，并考虑其技术知识、经验、教育和培训，以及人工智能系统将要使用的环境，并考虑到人工智能系统将要使用的人员或群体。

(European Union, 2024 AI Act, Article 4 [38])

为确保员工具备使用人工智能工具进行有效和可信工作的技能和知识，并遵守人工智能法案，公共机构正日益投资于人工智能相关培训。通常，政府以课程或研讨会的形式为其员工提供结构化培训。极少情况下，员工能够获得人工智能的长期培训，从而获得认可资格或学位（见附件表3.A.2）。通过实践学习、阅读内部指南、参加教育活动、同行交流和实践社区等非正式学习方式，也在机构内部发挥着重要作用，并可以通过领导力来推动（见战略性人力发展部分）。以下小节对选定经合组织国家的结构化培训计划进行了比较分析，并深入探讨了五个选定的案例。

为不同目标群体提供人工智能培训

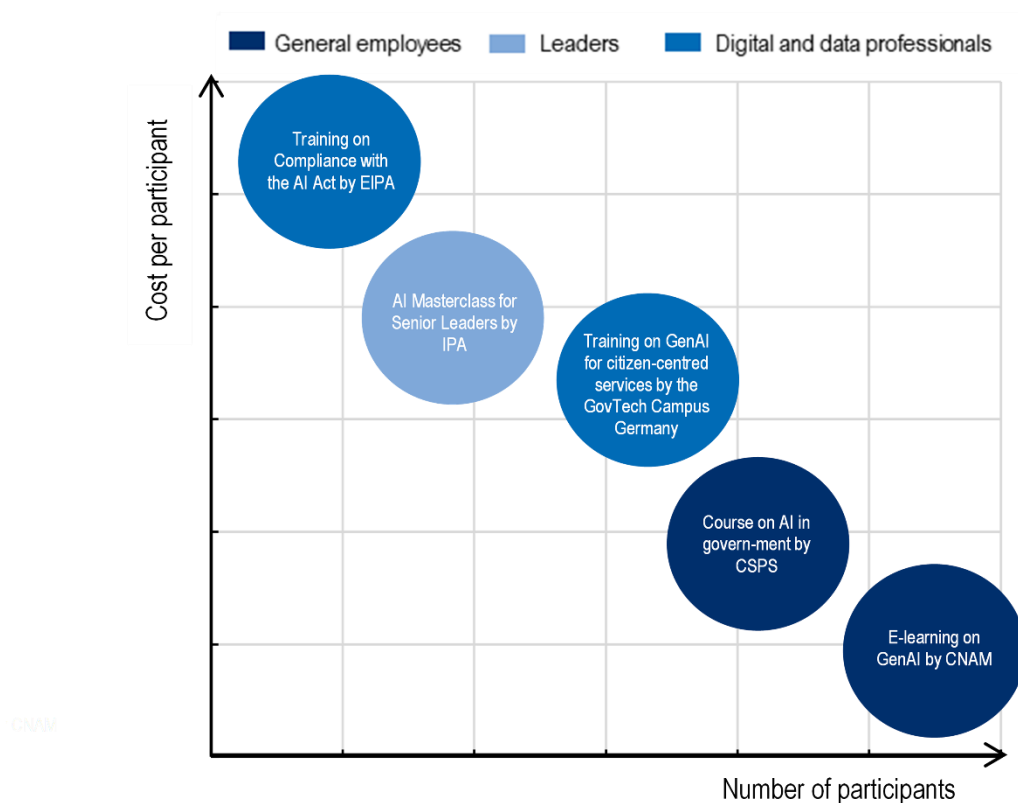
许多社会保障组织和更广泛的公共部门机构已经开始为员工开发和提供人工智能培训。人工智能相关培训的内容很大程度上取决于目标群体的学习需求，这些群体包括普通员工、领导者或数字和数据专业人士。针对公共部门普通员工的培训通常介绍人工智能概念和工具，并强调该技术的益处和风险。例如，爱沙尼亚数字国家学院提供的简短在线学习课程“人工智能入门”，公务员可以在其中了解人工智能是什么，如何使用它来创建更好的数字服务，以及在开发或实施人工智能项目时需要考虑哪些事项。这种基础的人工智能培训可以很容易地进行扩展，挑战在于实现广泛的员工参与。

对领导者的培训通常重点关注其组织及业务领域使用人工智能的战略潜力。例如，英国政府数字服务部门下属的公务员学习机构提供的在线培训课程就专注于人工智能的商业价值。对数字化和数据专家的培训可能旨在加强特定技术技能，提供关于人工智能监管的深度知识，并涵盖人工智能实施案例研究。赫尔辛基市例如，在2021年和2023年根据技术人员的需求提供了数据科学和数据工程培训机会。在目标群体中，有些课程更偏理论性，而另一些则非常实用，让参与者测试具体的AI工具并从实际应用案例中学习。有些培训具有产品导向，并且提供给组织以支持特定AI工具的实施。

在资源受限的现实情况下，人工智能相关培训的规模与强度之间存在权衡（图3.1）。换句话说，参与者人数需要与培训持续时间相平衡。例如，短期的自定进度在线课程开发和交付所需的资源相对较少，并且可以扩展到大量参与者。另一方面，面向公共部门高管的为期多天的面对面课程则更昂贵，并且只能提供给一小部分选定的参与者。图3.1指的是OECD政府实施的五个人工智能相关培训的具体案例。本节末尾对这些案例进行了进一步描述。

图3.1. 按目标群体划分的AI训练示例说明

按目标群体，训练强度（每人成本）与范围（参与者数量）之间的权衡的示意图



来源：经合组织。

培训形式涵盖员工可自行安排时间观看的短期在线课程，以及可在虚拟或线下进行的结构化培训项目，还有需要定期参与并最终获得认证的模块化项目。例如，在奥地利，联邦行政学院的数据公共服务学院为在三年内完成八天人工智能、数字化或数据相关主题培训的参与者提供证书（联邦行政学院，2025）^[39]。交付方式可以是自己节奏的在线学习内容、基于课堂的课程，可以虚拟或面对面进行，或者互动式面对面课程和研讨会。

提供人工智能相关培训的机构通常有多个，在不同经合组织国家中有所不同，包括公共部门的国家级培训机构、政府特定领域的培训机构、政府机构、非营利培训机构或技术提供商。在许多国家，公共部门的国家级培训机构，如加拿大的公共服务学院或爱沙尼亚的数字政府学院，已经开始提供人工智能相关培训。在法国，有EN3S（国立社会保障高等学院）和4.10学院，为法国社会保障部门提供一些人工智能相关培训。数字挪威、瑞典AI或葡萄牙AI4PA等创新机构为公共和私营部门提供学习机会。在德国GovTech Campus的案例中，作为一个注册的非营利机构，人工智能培训是与私营第三方提供商合作提供的。

以下示例突出了针对不同目标群体、培训模式和交付方式的培训具体案例研究。在不声称全面的情况下，它们被选中是为了突出政府正在发展和实施的与人工智能相关的培训的多样性。附件A提供了经选择的经济合作与发展组织管理机构中更长的AI培训项目清单。

示例 1：法国国家健康保险局 (CNAM) 的 GenAI 在线学习

- **描述：**这是一个介绍性在线学习模块，旨在帮助员工熟悉GenAI的主要概念，了解机遇和风险，采用负责任的实践，并学习如何有效编写提示。该在线学习是更大计划的一部分，旨在为GenAI的采用做准备，并且可以免费在属于法国健康保险系统 (CNAM) 的机构网络中获取。该在线学习于2025年2月启动，自那时起已获得93%的参与者认可。
- **提供者：** CNAM人力资源部门与AI部门之间的合作，以及第三方支持。
- **目标群体：** 所有员工
- **配送方式：** 自我学习在线课程
- **持续时间：** 1小时
- **优势与挑战：** 为机构的大规模人工智能战略制定GenAI使用的参考标准。该格式和交付方式允许大规模分发培训 (CNAM约有100000名员工)。培训内容简短且参与自愿，这引发了如何实现足够高的参与率和积极影响的问题。

示例2：加拿大政府公务员数字学院提供的关于人工智能使用的课程

- **描述：** 加拿大政府在GenAI应用方面的入门课程。探索GenAI工具及其潜在应用、高效提示、包容性AI实践以及不准确性的检测与预防。培训交付是迭代的，培训材料定期更新。目标是广泛推广，覆盖加拿大联邦公共服务中的10000名参与者。²
- **提供者：** 加拿大公共服务学院数字学院，联邦学习提供者，隶属于财政部。
- **目标人群：** 加拿大联邦政府的所有公务员
- **配送方式：** 基于虚拟课堂的课程
- **持续时间：** 2小时
- **优势与挑战：** 该课程允许大型教室，参与人数上限为700人，补充了一份关于使用GenAI的政府指南，免费提供给公务员，并有助于创建参考标准。培训内容相对一般，仅在固定日期进行。

示例3：欧洲行政学院关于欧盟人工智能法案的合规性培训

- **描述：** 符合欧盟人工智能法案要求的深度课程，旨在提供对人工智能法案风险分级方法的理解，以及其与更广泛监管环境 (包括欧盟的通用数据保护条例) 的相互作用。通过互动研讨会，参与者将获得知识和方法，例如在机构内部负责地采用和整合人工智能工具的风险缓解措施。³

- **提供者：** 欧洲公共行政学院（EIPA），是一个独立的教学与研究中心，总部设在荷兰马斯特里赫特，并由欧盟成员国及欧洲委员会通过其伊拉斯谟+计划共同资助。

- **目标人群：** 人工智能项目经理、数据保护官员、法律专业人士和其他数字化专业人士
- **配送方式：** 互动线下课程
- **持续时间：** 两天
- **优势与挑战：** 该强化课程允许深入学习合规性欧盟人工智能法案和为欧盟不同行政部门的专业人士构建网络国家。该课程费用为每位参与者1230欧元，并要求前往马斯特里赫特。

示例 4：爱尔兰公共行政学院面向高级领导层的 AI 大师班

- **描述：** 为爱尔兰公共服务领域的高级管理人员提供培训，以进一步发展他们对人工智能及其在政策制定和服务提供中的战略潜力的了解。学习目标包括理解人工智能相关的术语、伦理和监管考量、数据治理和质量的相关性，以及人工智能工具在不同业务领域的潜力。课程协调员采用不同的方法，包括讲座、讨论、实践挑战和同行交流。课程费用为每人460欧元，由相应的机构或个人支付。⁴

- **提供者：** 爱尔兰公共支出、基础设施、公共服务改革与数字化部下属的公共行政学院（IPA）。它提供正式培训作为本科和研究生课程的一部分，以及为爱尔兰不同机构和政府层级的公务员提供专业发展。

- **目标人群：** 领导者（高级公务员）
- **配送方式：** 互动线下课程
- **持续时间：** 有一天
- **优势与挑战：** 该课程专注于AI的战略应用，与领导者的学习需求相对应。课程的线下性质便于参与者建立联系。同时，这也限制了参与人数，并使课程资源密集。

示例 5：由德国 GovTech 校园提供以人为本服务的 GenAI 培训

- **描述：** 一门关于不同AI应用案例的课程，具有提高公共行政效率、生产力和服务质量的潜力。参与者将学习从规划到实施AI工具时需要考虑的重要方面，包括合规性、数据保护和IT安全。他们建立了一个同行网络，并学习了克服公共行政中实施AI工具相关挑战的策略。⁵

- **提供者：** 德国GovTech园区，一个政府资助的非营利组织，与第三方私营供应商合作。

- **目标人群：** 数字和数据专业人员，其他公务员。
- **配送方式：** 互动线下课程
- **持续时间：** 4小时
- **优势与挑战：** 该课程对参与者免费，并促进网络建设建筑。同时，每场活动仅限25名参与者，且仅在柏林进行。

德国的非营利组织地位使得其能够与私营部门进行紧密合作，但也带来了资源限制。

政府也可以依赖现有的AI培训和免费的学习资源，要么鼓励其员工直接参与其中，要么作为开发自有课程材料的灵感。例如，赫尔辛基市鼓励其39000名员工报名参加公开在线课程“AI基础”，除了提供AI基础知识在线学习课程、每月学习活动以及关于Microsoft CoPilot和其他生成式AI工具的产品相关培训外。框3.1总结了与公共管理相关的免费培训和学习资源。

3.1. 人工智能免费培训和学习资源箱

以下示例是关于人工智能的培训材料和教材，对政府机构可能相关，并且可以在线免费获取。⁶ 它们由研究中心或私营公司提供：

- 政府在线学习AI **AI4GOV** (<https://ai4gov-project.eu/home/resources/training-learning/>), 欧盟资助的平台提供视频资料、活动邀请函，以及三门围绕“可信人工智能促进透明公共治理、培育民主价值观”主题的开放在线课程。

The **人工智能的要素** 交互式在线课程 (<https://www.elementsofai.com/>) 由MinnaLearn和赫尔辛基大学提供。它是一门需要花费4-8小时完成的巨型开放式在线课程 (MOOC)，完成后为学习者提供证书。它为学习者提供了人工智能的介绍，包括其在现实世界的应用、工作原理及其对社会的影响。

- The **图灵在线学习平台** (<https://www.turing.ac.uk/courses>) 提供了一系列关于负责任人工智能的在线课程，面向普通和技术受众，课程时长不同 (5-40小时)。包括人工智能伦理、人工智能中的偏见与歧视或透明机器学习等主题。

- 课程、活动和其他学习资源由 **非政治性** 关于面向公务员和高级政府领导人的政府人工智能 (<https://apolitical.co/learning-hub/government-ai-campus/>), 包括一个“人工智能准备情况检查” (<https://apolitical-arc.co/>)。课程时长为1-4小时，涵盖主题例如人工智能在医疗保健中的应用、人工智能领导力或公务员人工智能基础知识。

私营部门的公司和组织经常面临与公共机构相似的挑战。表3.2提供了关于私营部门培训的见解，旨在建立与《人工智能法案》关于人工智能素养的第4条的合规性。

3.2. 私营部门为遵守人工智能法案第4条所采取的措施

欧盟人工智能办公室已汇编了一份符合《人工智能法案》第4条的私营部门培训计划库，该条款要求人工智能系统的提供者和部署者确保其用户具备足够的人工智能素养。以下示例更具体地适用于公共行政和社会保障机构，因为它们为公共部门提供服务或处理类似于社会保障机构的非常敏感的数据。

- **Gjensidige Forsikring ASA** 是挪威的一家保险公司，使用多个AI系统。它采用多渠道方法，为员工提供不同类型的AI培训，包括内部

知识交流平台。公司为所有员工开发了必修的AI基础知识在线学习课程。此外，根据员工的角色和职责，还会提供特定的培训，例如，分析师的培训将侧重于模型风险管理、数据治理以及他们在AI价值链中的具体职责。参与索赔管理的人员会根据他们使用的AI系统接受产品特定的培训。此外，在适当的情况下，AI基础知识将在招聘过程中进行验证，针对外部候选人的ICT和AI系统培训是入职流程的一部分。

- **Dedalus Healthcare** 是一家总部位于意大利的全球医疗保健和诊断软件公司。该公司开发和使用的AI系统根据《AI法案》可被归类为高风险。面向整个组织员工，提供有关AI基础知识、其益处与风险、伦理考量及实际应用的培训。针对特定角色，特别是高管、数据保护官员（DPO）、开发人员和工程师，以及法律、合规、质量保证和监管事务部门，计划进行定制培训。

- **Ineco** 是一家专注于可持续交通和公共部门服务数字化转型的西班牙国有工程咨询公司。Ineco为全体员工提供公司范围内的AI培训，重点在于基础AI知识并强调最佳实践。针对直接参与开发或监督AI系统和模型的技术角色和项目经理，还提供专业培训。这确保了每个人都具备AI安全和伦理的基线理解，而参与AI项目的员工则获得更深入、更有针对性的指导。该公司面临的一个挑战是合格专家的短缺，他们能够设计和提供AI培训。为应对这一问题，公司为高技能员工提供“培训培训师”项目，为他们提供教学工具，使他们能够有效地向其他员工传授AI相关主题。

- **斯马尔斯** 是一家比利时的非政府组织，专门为比利时的公共和社会保障部门提供ICT服务，开发人工智能解决方案的试点、概念验证和原型。该组织建立了一套全面的培训策略，针对所有员工提供通用的人工智能素养课程，以及针对数字和数据专业人员的各种群体（例如人工智能专家；数据工程师；基础设施专家；数据保护官和法律团队；开发者、项目经理）提供专门培训。此外，在启动新的AI项目之前，Smals会为所有相关利益相关者开展针对产品的特定培训。Smals的另一个显著举措是任命AI大使，他们的职责是推广和促进人工智能技术的应用，并最大化其价值和有效性。

来源：欧洲人工智能办公室，2025，《人工智能素养实践存储库》v.16.04.2025 <https://digital-strategy.ec.europa.eu/en/library/living-repository-foster-learning-and-exchange-ai-literacy> .

战略人才队伍建设

公共部门组织可以战略性地调整劳动力发展，以构建AI能力，包括通过招聘和培训，如上所述。本节基于经合组织的良好实践，考虑了两个支持劳动力中AI技能和能力的要素，以及机构内任务和流程的潜在变化：将劳动力发展与更广泛的AI战略保持一致，并培养创新和持续学习的文化（经合组织，2023）。^[17]。

将劳动力发展与更广泛的AI战略相结合

将劳动力发展与更广泛的AI战略相结合，有助于采取更协调一致的方法，并更好地利用招聘和培训的资源。2022年，大约三分之二的经济合作与发展组织（OECD）国家的政府报告称，他们在中央层面以及部委和机构层面为其员工制定了学习和发展战略（OECD，2023）。^[17]。然而，人工智能与技术采纳和劳动力发展之间的联系并不明显，因此，组织将从采取协调方法中受益。招聘和培训是改善服务和福利的关键杠杆，采用数字和人工智能技术也是如此。如果这些杠杆能良好协同，它们就有潜力相互增强。

爱沙尼亚是一个在人工智能应用和工作force发展方面进行战略协调的良好范例。该国通过一个全面、涵盖整个政府的战略，以结构化和渐进的方式在公共部门采用人工智能，并强调了技能发展的重要性。自2019年起，该国发布了人工智能任务小组报告⁷ 人工智能战略 2019-2021⁸，标志着精心规划、分阶段推出的第一步。随后是更新的2022-2023年人工智能战略⁹ 以及一个2024-2026年的人工智能行动计划。¹⁰ 所有这些文件都包括一系列旨在使公共部门具备有效采用和利用AI所需技能的措施。

早在2019年，人工智能工作小组报告就确定了关键的人力资源发展措施，以促进人工智能的采纳，例如提高领导层的意识并提升公务员的技能。后续的策略和行动计划包括具体目标，例如任命人工智能领导者以及提供各种人工智能学习活动。更普遍而言，该方法依赖于加强人工智能和数据素养，并提升公共机构使用数据的能力。

在不同战略和行动计划中提到的培训和学习机会是通过两个主要平台提供的：数字国家学院¹¹ 以及Kratid平台。¹² 爱沙尼亚数字经济学院是爱沙尼亚所有公共部门官员使用的在线学习平台，由经济事务与通信部与塔林理工大学合作开发。课程通常由该部内部开发，但有些课程由外部供应商设计。目前，该学院提供66门免费课程，涵盖数字和人工智能主题以及其他对公共部门具有战略意义的领域，且课程提供范围不断扩展。大多数课程相对较短，通常约为一小时，旨在提供基础和入门知识。主题包括数据管理、网络安全、使用政府云服务、人工智能相关课程，例如项目管理和平等政策。该部监控课程参与度、完成情况和参与者的满意度。通常，这些课程用于培训新员工。

除了数字国家学院，公务员还可以通过Kratid平台获得培训机会。由司法与数字事务部管理，该平台提供有关人工智能和数据的各种资源，包括人工智能支持工具箱或数据科学主题的培训视频。它提供有关公共部门虚拟助手Bürokratt的详细信息以及数据治理和数据经济的详细信息。通过Kratid，司法与数字事务部还提供其他几项能力建设措施，以促进人工智能的采用，包括定期研讨会和专家网络。每季度都会组织关于公共部门人工智能用例的知识分享会，由爱沙尼亚或国际专家进行演示，并与观众进行讨论。

针对希望采用人工智能的公共机构，有一项特定的AI支持服务可供使用。该服务包括头脑风暴会议和会议，以评估需求、识别挑战，并提出潜在的AI解决方案。有时会根据特定机构的需求开发定制化的研讨会，例如，一个旨在提高市镇内部对人工智能认识的参与式研讨会。另一个例子是关于如何为AI工具编写更好提示的课堂。这些具体的研讨会通常由外部服务提供商开发。

爱沙尼亚在公共部门采用人工智能的方法强调了从一开始就将劳动力发展纳入全面的分阶段国家人工智能战略的重要性，不仅要提供普通培训（人工智能素养），还要提供针对不同专业群体（领导者、人工智能专家）需求的培训，并通过数字国家学院和Kratid等平台使其获得，并定期与公共部门官员互动，以确保他们的认同和广泛参与学习机会。

培养创新和持续学习文化

经合组织之前的工作强调了学习文化对支持更具灵活性的公共部门的重要性（经合组织, 2023）^[17]。这在采用人工智能工具的情况下尤其紧迫。采用人工智能系统依赖于组织内所有潜在人工智能用户的基本意识和开放性。同时，技术解决方案若与组织流程和任务的重新思考相结合，可能会产生最大的效果，而这反过来又需要员工批判性地、创新地反思如何最好地实现其组织的目标。虽然领导和数字专家当然可以推动人工智能创新，但全体员工都对变革管理的速度和便捷性以及人工智能的有效和负责任采用做出了贡献。虽然难以衡量，但一个鼓励和激发好奇心、开放性和实验的环境可以有意地创造出来，例如通过人工智能的学习活动或专家社区。创造和维护创新和学习文化可以被视为随着时间的推移支持改进而准备肥沃的土壤。

为了支持组织学习，许多公共部门机构都设立了中央创新和技术团队。将数字和数据专业人员与熟悉业务流程和需求的专家相结合的跨学科创新团队可以取得良好成果。一个例子是正在意大利国家社会保障机构（INPS）内部开发的“人工智能能力中心”，它是意大利的主要社会保障机构。该能力中心将设在罗马的INPS中央总部，成为一个专注于人工智能、并由具有跨学科背景的专业人士提供支持的内部创新、协调和治理单位。其目标是促进员工参与和协作，在人工智能系统实施时为整个机构提供技术支持，为员工提供人工智能技能发展机会，并改善人工智能的内部沟通。

另一个例子是由IT Systemhaus（如第一章所述）建立的卓越中心AI中心，它是德国联邦就业机构的IT服务提供商。卓越中心AI中心成立于2019年。目前该中心由约15人组成，其中许多人具有数据科学背景，但也有其他专业背景。每当卓越中心AI开发新产品时，他们都会组建一个跨学科项目团队，该团队既包括业务专业人士，也包括技术专家。随着时间的推移，他们意识到业务专业人士的参与不仅对构建机器学习模型的训练集很重要，而且对开发以用户为中心、满足业务需求的解决方案至关重要。如今，在部署任何AI系统之前，数据和AI伦理、数据保护和IT安全功能都会评估系统大规模部署相关的风险，确保符合GDPR和AI法案，并获得批准方可使用该AI系统。

需要刻意在企业内部创建创新和持续学习的文化。一个有趣的例子是由加拿大公共服务学院组织的关于人工智能的学习周。加拿大联邦政府鼓励公务员在这个星期参加各种浓缩的学习活动，包括现场活动、讲师指导的课程和自我学习（加拿大政府，2025）。^[40]

另一个例子是爱沙尼亚的年度会议Kratitreff，它汇集了爱沙尼亚公共和私营部门的AI专家和用户，以支持使用AI进行更好的决策和服务提供，与该国的更广泛的AI战略一致（见上文）。类似地，葡萄牙的公共行政人工智能和数据分析（AI4PA）计划提供了一系列培训活动

聚焦公共管理和中小企业中的人工智能和数字化转型。全天活动包括培训课程、互动研讨会、资金机会信息和IT解决方案。注册费约为每人30欧元，并在结束时提供结业证书。

爱沙尼亚司法和数字事务部运营不同的网络来建立一个专家社区，以便公共部门可以在此进行合作和交流良好实践。一个数据专家小组汇集了国家级的各个机构和机构的IT领导者，以分享人工智能开发和实施的经验。该部还建立了一个实践社区，由约700名专业人士组成，他们从事人工智能、数据或信息技术工作，服务于统计局（爱沙尼亚的国家统计机构）、公共就业服务、社会保险委员会或土地与空间发展委员会等公共组织。该部通过电子邮件、简报、会议和研讨会定期与该实践社区互动，有时还会就他们的需求提出反馈，尤其是在培训方面。

公务员，无论是一般职位还是数字和数据分析职位，都可以从允许进行人工智能实验的空间中获益。澳大利亚的人工智能合作实验室旨在提供此类空间，让公务员学习、测试不同的AI工具，并与各级政府部门、非政府组织、学术界和私营部门建立联系。人工智能合作实验室举办不同类型的人工智能活动、研讨会和培训，例如一个定期的线下研讨会“面向政策制定者的实用人工智能”，该研讨会会让公务员以及跨部门参与者能够熟悉具体的人工智能工具。

赫尔辛基市设有专门的“实验加速器”，旨在培育和奖励人工智能及数字技术的创新。该加速器于2019年作为城市数字化转型计划的一部分成立，旨在支持使用技术及新型工作方式进行的快速内部实验。在几次项目征集过程中，城市员工提交了120份实验提案，其中65个被选中。每个被选中的提案获得了约1万欧元的资助，这允许获得第三方支持。实验并不总是以新技术的实施而结束，而是有明确的开发“经验教训”的目标，这些经验教训会收集在专门网页上，并可作为后续项目的基石。¹³

关键发现

人工智能的采用将对社会保安劳动力产生影响

- 社保领域成功应用人工智能系统，有潜力提高服务效率并释放员工时间用于更高价值的工作。最重要的是，人工智能的应用可能会随着时间的推移改变组织内的现有工作角色和流程。
- 虽然人工智能的采用可能会增加对数据和数字专业人才的需求，但全体员工具备良好的数字技能和人工智能素养是关键。解决问题、批判性思维或社交能力等互补性和横向技能将促进组织变革。
- 社保机构中人工智能工具的使用日益增多，需要积极管理，包括通过培养劳动力队伍中与人工智能相关的核心技能和能力来
赋能公共部门员工的人工智能素养；使领导者能够推动组织变革
；以及吸引和留住数字和数据专业人才。

社会保障机构需要内部人工智能能力

- 内部战略性地保持关键技能和功能，有助于提高问责制、透明度和数据隐私，使人工智能系统更好地与组织目标保持一致，并降低对技术供应商的依赖。
- 政府有不同的工具来提高数字和数据专业人员招聘和留住率，例如，具有时限性项目、明确工作角色和这条职业群体的职业发展路径的高影响项目，以及战略性的政府科技项目。
- 鉴于招聘的灵活性有限，员工持续技能发展是关键。公共机构正越来越多地为员工提供人工智能相关的培训。他们可能针对普通员工以培养人工智能基础技能，针对领导者以获取人工智能战略知识，或针对数字和数据专业人员以提升人工智能技术相关技能。

劳动力发展可以战略性地支持人工智能的采纳

- 劳动力发展，包括通过招聘、外包和培训，应与组织更广泛的ai战略保持一致，并为其组成部分。
- 公共部门机构中强烈的创新和学习文化将支持变革过程，并利用人工智能等技术，例如通过跨学科团队、实验空间和持续学习的激励措施，不断提高服务和福利。

参考文献

阿克莫格鲁, D. 等人 (2022), “人工智能与就业: 来自网络招聘的证据” [6]
劳动经济学杂志, 第40卷/S1, 第S293-S340页, <https://doi.org/10.1086/718327> .

阿特金森, I. 和 J. 布朗 (2024), *人工智能对公共部门劳动力的潜在影响*, [8]
 托尼·布莱尔全球变革研究所 <https://institute.global/insights/economic-prosperity/the-potential-impact-of-ai-on-the-public-sector-workforce> (2025年5月12日访问)

澳大利亚政府 (2025), *APS职业发展路径工具*, [33]
<https://www.apsprofessions.gov.au/career-development/aps-career-pathfinder-tool> 访问
 2025年11月24日).

澳大利亚政府 (2023) *移动机会 / 数字职业*, [32]
<https://www.digitalprofession.gov.au/career-development/mobility-opportunities> (访问于 2023 年 6 月 27 日)。

博尔戈诺维, F. 等人 (2023), “14个经合组织国家人工智能技能需求的兴起趋势”, [14]
经合组织人工智能论文, 第2号, 经合组织出版局, 巴黎,
<https://doi.org/10.1787/7c691b9a-en> .

布里特, J 等 (2025年), “生成式人工智能已广泛应用于公共部门: 来自一项研究的证据” [19]
“英国公共部门专业人士调查” 数字政府: 研究与实务, 第6卷/1期, 第1-13页, <https://doi.org/10.1145/3700140> .

布赖恩·约尔弗森, E., 李D.和雷蒙德L. (2023年), “工作中的生成式人工智能” *Nber 工作论文* [3]
系列, 第31161卷, <http://www.nber.org/papers/w31161> (2025年5月12日访问)

联邦印刷厂和可能 (2025) *专业方法解码。数据和AI应用* [9]
高效行政,
https://scnem2.com/art_resource.php?sid=8vufi.j1b3di,force_download=1&u=5a59803f51917537a55471fcb5fd688f (于2025年7月28日访问)。

伯特施尔, M., S. 皮亚诺和B. 韦尔比 (2024), “培养数字政府技能: 一项综述” [31]
经合组织各成员政府在良好实践方面的经验, *经合组织社会、就业与移民工作组论文*, 第303号, 经济合作与发展组织出版局, 巴黎, <https://doi.org/10.1787/f4dab2e9-en> .

奇恩, D. 等人 (2020), *未来已来: 弥合欧洲公共部门的人才技能差距*, [10]
 麦肯锡公司,
<https://www.mckinsey.com.br/~media/McKinsey/Industries/Public%20and%20Social%20Sect%20欧洲的%20公共部门/The-future-is-now-Closing-the-skills-gap-in-Europes-公共部门-vF.pdf> (2025年5月23日访问)。

Czarnitzki, D., G. Fernández和C. Rammer(2023), “人工智能和企业层面的” [2]
生产力, *经济行为与组织杂志*, 第211卷, 第188-205页,
<https://doi.org/10.1016/j.jebo.2023.05.008> .

数字服务 (2025) *数字先锋和变革专家奖学金*, [27]
<https://digitalservice.bund.de/en/fellowships> (于2025年6月6日访问)。

- 欧洲基金 (2023) *解决劳动力短缺问题的措施：对未来政策的启示*，出版物 [24]
 欧盟委员会办公室，卢森堡。
<https://www.eurofound.europa.eu/en/publications/2023/measures-tackle-labour-shortages-future-policy-course> (于2025年1月28日访问)。
- 欧盟 (2024) *欧洲议会和欧盟理事会法规 (EU) 2024/1689* [38]
 2024年6月13日理事会决定制定统一的人工智能规则并修订
 法规 (EC) 第300/2008号，(EU) 第167/2013号，(EU) 第168/2013号，(EU) 2018/858，(EU) 2018/1139 an，<http://data.europa.eu/eli/reg/2024/1689/oj>。
- FedScoop (2024)，“更多联邦机构加入临时阻止或禁止ChatGPT |” [21]
<https://fedscoop.com/more-federal-agencies-join-in-temporarily-blocking-or-banning-chatgpt/>
 (2025年5月26日访问)。
- 吉尔斯皮，N.等(2025) *信任、态度与人工智能的使用：一项2025年全球研究*， [18]
 墨尔本大学；毕马威 <https://doi.org/10.26188/28822919>。
- 全球政府论坛 (2023)， *加拿大公共服务对人工智能的态度*， [20]
 Gmyrek，P等(2025)， *生成式人工智能与工作：2025年更新*，国际劳工 [7]
 组织， <https://www.ilo.org/publications/generative-ai-and-jobs-refined-> (2025年6月4日访问)。
- 法国政府，DINUM和INSEE (2021)， *评估国家需求* [11]
数据方面的能力和专业知识，
<https://www.numerique.gouv.fr/uploads/RAPPORT-besoins-competences-donnee.pdf>
 (访问于2025年2月11日)。
- gov.uk (2025) *政府数字和数据职业能力框架*， <https://ddat-capability-framework.service.gov.uk/> [30]
 (2025年5月27日访问)。
- 加拿大政府 (2025) *人工智能学习周*， <https://www.cspsexpfc.gc.ca/events/learning-week-ai/index-zh-cn.aspx> [40]
 (于2025年6月6日访问)。
- 德国GovTech校区 (2025年) *govtech校区德国*， [36]
<https://govtechcampus.de/> (2025年5月27日访问)。
- 格林，A. (2024)，“人工智能与劳动力市场技能需求的变化” [13]
经合组织人工智能论文，第14号，经合组织出版，巴黎，
<https://doi.org/10.1787/88684e36-en>。
- IÖB (2025)， *公共采购促进创新*， <https://www.ioeb.at/en/> [35]
 2025年5月27日)。
- 库皮，M.，S.扬金和G.哈默施密德 (2022)， *政府中的数据科学与人工智能。为什么* [25]
公共部门组织需要内部的数据科学和人工智能专业知识，
 赫特里学院 https://www.hertie-school.org/fileadmin/2_Research/2_Research_directory/Research_Centres/Centre_for_Digital_I_Governance/5_论文/其他论文/HS_政策简报_英文最终版本打印.pdf
 (于2025年5月14日访问)。
- 库齐姆斯基等(2022)， *欧盟的政府技术实践*，欧洲出版局 [34]
 卢森堡联合 <https://doi.org/10.2760/74735>。

莱恩, M., 威廉姆斯, M.和布鲁克, S. (2023年), “人工智能对工作场所的影响: 主要发现 [4]
源自经合组织对雇主和工人的AI调查”, 经合组织社会、就业与移民工作组论文 第288号, 经合组织出版, 巴黎,
<https://doi.org/10.1787/ea0a0fe1-en> .

梅达利亚, R., P.米卡拉夫和L.唐吉 (2024), 人工智能的能力和治理实践 [12]
公共部门, 欧盟出版局,
<https://data.europa.eu/doi/10.2760/7895569> (2025年5月26日访问)。

梅格尔, I 等人 (2023), “在公共部门实施人工智能” 公共管理评论 , [1]
第1-14页 <https://doi.org/10.1080/14719037.2023.2231950> .

OECD (2024), 在政府中实现数字创新: 经合组织GovTech政策 [26]
框架, 经合组织数字政府研究, 经合组织出版, 巴黎,
<https://doi.org/10.1787/a51eb9b2-en> .

OECD (2023), 经合组织就业展望2023: 人工智能与劳动力市场 , [5]
经合组织出版商, 巴黎, <https://doi.org/10.1787/08785bba-en> .

OECD (2023), 公共就业与管理 2023: 迈向更灵活的公共 [17]
服务, 经合组织出版商, 巴黎, <https://doi.org/10.1787/5b378e11-en> .

经合组织 (2021), “经合组织关于公共部门数字化人才和技能的框架”, 经合组织 [22]
公共治理工作文件 45号, 经合组织出版, 巴黎
<https://doi.org/10.1787/4e7c3f58-en> .

法兰西共和国 (2025) 公共服务数字孵化器 -beta.gouv.fr , [29]
<https://beta.gouv.fr/> (于2025年7月29日访问)。

艾伦·图灵研究所 (2024) 商业能力框架中的人工智能技能 , [23]
https://zenodo.org/records/1109267/preview/AISkillsForBusinessCompetencyFramework_V2.pdf?include_delete_d=0 (访问于2025年6月2日)。

Theben, A., N. Plamenova和A. Freire (2023), “未来的“新货币”: 回顾” [15]
关于加速数字化时代下劳动力技能需求方面的文献
管理评论季刊 , 第75卷/第1期, 第495-526页, <https://doi.org/10.1007/S11301-023-00387-9/METRICS> .

美国通用行政管理局 (2025) 总统创新研究员 , [28]
<https://presidentialinnovationfellows.gov/> (于2025年6月6日访问)。

埃塞克斯大学 (2025) 与分析数据科学学院合作 , [37]
<https://www.essex.ac.uk/centres-and-institutes/institute-for-analytics-and-data-science/working-with-business#policy> (2025年5月23日访问)。

瓦莱-克鲁斯, D., R. 加西亚-孔特拉雷斯和J. 赫拉蒂西亚·穆诺兹-查韦斯 (2024), “领导力与 [16]
公共服务领域的转型: 第四次工业革命期间人工智能采用与效率的实证探索 数字政府研究 , 第 794-806 页,
<https://doi.org/10.1145/3657054.3657146> .

联邦行政学院 (2025) , 数据公共服务学院 , [39]
<https://www.vab.gv.at/bildungsprogramm/fachbereiche/public-management-und-governance/digital-government.html> (2025年5月26日访问)。

注释

- ¹ 看 https://joint-research-centre.ec.europa.eu/projects-and-activities/education-and-training/digital- transformation-education/digital-competence-framework-citizens-digcomp/digcomp-framework_en (访问于2025年5月26日)
- ² 看 https://catalogue.csps-efpc.gc.ca/product?catalog=DDN321&cm_locale=en (访问于2025年6月6日)。
- ³ 看 <https://www.eipa.eu/courses/artificial-intelligence/> (在6日访问) th 2025年6月)
- ⁴ 看 <https://www.ipa.ie/information-systemstechnology/artificial-intelligence-masterclass-for-senior- public-ser vice-leaders.6741.html> (在6日访问) th 2025年6月)
- ⁵ 看 https://www.linkedin.com/search/results/all/?fetchDeterministicClustersOnly=true&heroEntityKey=urn%3Ali%3Aor ganization%3A72086001&keywords=govtech%20campus%20deutschland&origin=RICH_QUE_RY_SUGGESTION_&position=0&searchId=7e6cbf68-910f-4110-93a8-1641acbc3af1&sid=uXH&spellCorrectionEnabled=false (于2025年6月6日访问)。
- ⁶ 这些资源仅供参考，但并非经合组织官方认可。
- ⁷ <https://cdn-assets.inwink.com/b0269dea-af7b-4460-b29a-c40b9941c4c5/9042bd3d-7e92-45ad-bf58-28a412c3a13d> (访问于2025年5月21日)。
- ⁸ https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_27a618cb80a648c38be427194affa2f3.pdf (访问于2025年5月21日)。
- ⁹ https://www.kratid.ee/en/_files/ugd/980182_4434a890f1e64c66b1190b0bd2665dc2.pdf (访问于2025年5月21日)。
- ¹⁰ <https://www.mkm.ee/sites/default/files/documents/2024-02/Tehisintellekti%20tegevuskava%202024-2026.pdf> (在爱沙尼亚，访问于2025年5月21日)。
- ¹¹ <https://digiriigiakadeemia.ee/> .
- ¹² <https://www.kratid.ee> .
- ¹³ 看 <https://kokeilukiihdyttamo.hel.fi/results> (访问于2025年6月18日)

附件 3.A. 经合组织国家中与人工智能相关的培训示例

附件 表 3.A.1 公共行政中人工智能培训示例

机构	国家	目标组	类型	配送方法	持续时间	描述	链接
加拿大大学校公共服务	加拿大	所有民事仆人	入门	在线，教室 m-based 课程	2h	人工智能培训课程政府，补充政府的指南生成式人工智能的使用	https://catalogue.csp-s-efpc.gc.ca/product?catalog=DDN321&cm_locale=zh
公共学院管理	爱尔兰	公民仆人，领导们，数字和数据职业 als	入门；战略；技术	在线或面对面课程	8-24小时	人工智能课程范围公共行政中的数据包括实施国家人工智能指南，人工智能高级民事大师班仆人	https://www.ipa.ie/training-at ion-系统技术。1875.html
公务员学习政府数字化服务，部门科学，创新和科技	英国	公民仆人，领导者，数字和数据职业 als	入门；战略；技术性，	在线，互动性 e 课程	-	人工智能课程范围不同主题，包括AI伦理，人工智能工具和应用，业务人工智能或机器的价值学习。	https://www.gov.uk/政府/新闻/n ew-courses-on-人工智能-启动
人工智能和数据科学 for 公开管理 (AI4PA) 欧洲数字创新 Hub	葡萄牙	领导们，数字和数据职业 als	战略；技术性，	线下训练 events/ 工作坊 s	6h	面对面培训周期人工智能和数字事件公共转换行政管理与中小企业	https://ai4pa.pt/noticias/事件循环-formativos-em-ia-e-transformacao-digital-na-administracao-公开-并-支付
欧洲研究所公共管理	欧盟能源	领袖	战略	面对面	8-16小时	几个与公共行政中的AI在人工智能风险管理方面，法律考量，AI在采购，知识管理，或组织转换	https://www.eipa.eu/product-category/artificial-intelligence/
澳大利亚公共服务学院	澳大利亚	所有公共服务员工	入门	自我-定向网络学习课程	-	提供入门基础课程人工智能在政府中的由数字开发转换机构 (DTA), 事件和记录涵盖人工智能主题的网络研讨会资源	https://www.apsa.gov.au/artificial-intelligence-ai
部门金融	澳大利亚	所有公共服务员工	入门	自我-定向 e-学习		初学者和中级模块嵌入式 GovAI 平台允许动手培训 (例如)	https://www.govai.gov.au/了解人工智能

机构	国家	目标组	类型	配送方法	持续时间	描述	链接
				课程		入门到中级提示工程和检查/评估技术)。知识中心资源：事实手册和餐垫提供实用指南关于人工智能基础	
联合国发展项目 (联合国开发计划署), 数字容量实验室	全球	所有 government 官员	入门	在线和在-人课程	-	人工智能培训计划对于政府, 关于人工智能应用, 法规实现与伦理。	https://www.undp.org/digital/ai-政府
联合国教科文组织和全球国际远程通信 n 并	全球	公民仆人和政策制定ers	入门	在线讲师 LED	40h	人工智能计划公共部门	https://academy.itu.int/training-courses/full-目录/人工-intelligence-public-部门
伙伴关系公共服务 (非政府组织)	美国	领袖	战略	在线课程	18h	领导力培训人工智能项目与微软和谷歌致美国高管联邦政府。	https://ourpublicservice.org/course/ai-政府-leadership-program/
联邦学院行政管理, 数据学院公共服务	奥地利	领导们, 全部	入门战略	-	8-16 小时	系列关于人工智能的课程政府, 例如在使用人工智能工具提示, 面向高管的AI道德考量, 或管理AI项目	https://www.vab.gv.at/教育培训计划m/fachbereiche/public-management-und-governance/数字-government.html

注意：链接最后访问时间为2025年6月30日。

附件 表格 3.A.2. 公共行政中人工智能训练导致学位或资格的实例

名称	机构	国家	目标组	持续时间	描述	链接
AI4Gov 硕士程度 (人工智能智能公开服务)	理工大学 de 马德里和米兰理工大学曾受支持欧盟的全球网关计划	意大利和西班牙国际	领袖	12 月份	硕士项目要培养所需的技能以设计、实现和在公共领域管理人工智能行政管理。面对面和混合	https://www.polidesign.net/en/xuexi/shangye-设计/主--人工-公共智能-services/
证书：人工智能和数据科学公开管理 on	赫特里学院政府合作 government technology campus 德国和 others	德国国际	领袖	18天 of 指令 n (部分-超时 6个月)	选择性但免费项目支持基础和设计用于提供战略对人工智能的理解公共管理。	https://www.hertie-school.org/zh/study/executive/ki-und-data-science-为公众-管理
APS数字实习程序	澳大利亚公共服务 (APS) 工作	澳大利亚	数字和数据职业nals	12-24 月 (全职或部分-时间)	技能提升和转岗培训公务员项目向数字化转型通过在-职业在岗学习与培训课程, 一个正式的资格在最后。	https://content.apsjobs.gov.au/career-pathways/digital-实习项目

注意：链接最后访问时间为2025年6月30日。

利用人工智能于社会保障

用例、治理和劳动力准备

人工智能可用于提高社保福利的获取途径，并确保在恰当的时间将恰当的福利发放给恰当的人群。为探索如何利用人工智能达成此目标，本报告记录了国家公共服务中具体的政策杠杆、工具、实施策略以及能力建设努力，尤其关注数据质量、治理以及劳动力准备度等问题。报告为社会安全部门如何受益并同国家层面的努力保持一致，以实现更紧密、更有影响力及更值得信赖的人工智能应用提供了指导。进一步的努力将是必要的，以探索在社会保障部门中人工智能的有效应用，尤其是在将这些做法适应社会保护系统的具体运营、法律和伦理环境时。持续的实验、评估和跨部门合作对于确保人工智能的采用能够带来切实利益，同时保障公平、透明和公众信任至关重要。

该报告强调了利用人工智能作为更包容、更有效、更透明的社会保障体系战略推动者的关键机遇。实现这一潜力将取决于嵌入健全的治理框架、投资数据和基础设施，以及建立公共部门值得信赖的人工智能所需的制度和人力能力。

