NOVARA
CONSULTING GROUP

This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2025/05/27/the-policy-is-the-exhibit-generative-artificial-intelligence-in-human-resources-documentation-and-advice/.

政策即证据:生成式人工智能在人力资源文件与建议中的应用

Heather Grizzle · May 27, 2025

生成式人工智能可以快速起草员工手册、政策和职位描述。这些文件在诉讼中都是证据,而来源可信度正是决定其分量的关键。

目录

1. 曝光实际所在之处
2. 政策就是证物
3. 虚构的法律依据,以及人力资源部门为何比法务部门处境更糟
4. 持续对齐并非该技术的实际运作方式
5. 审计准备状态被颠倒了
6. 职位描述是一份法律文书
7. 聊天机器人给出的建议,雇主必须为之承担责任
8. 需要建立的事项
9. 结语
10. 参考文献

曝光实际所在之处

人工智能在人力资源领域受到的关注,大多集中在筛选环节:谁被筛查、排名、晋升,依据是什么。这种关注是有道理的,本机构对此已有大量论述。但这也使得增长最快的应用几乎完全未被审视。

Generative systems are now being used to draft, update, and localize the documents that constitute an employer's legal position. Handbooks. Accommodation procedures.

Leave policies. Job descriptions. Disciplinary standards. Investigation templates. Multi-state compliance addenda. The pitch is that this work, historically slow and expensive, can be produced instantly and kept continuously current as regulations change.

Novara Consulting Group's position is that this application carries greater and more immediate risk than automated screening, for a reason specific to the artifact: a policy is not an operational document. It is evidence. It is produced in litigation, cited in agency investigations, relied upon by employees, and read against the employer by people whose job is to find the gap between what it says and what happened. An organization that has generated its legal position without provenance has created a large volume of admissible material about itself that nobody at the organization can account for.

政策就是证物

In an employment dispute, the employer's written policy performs several functions at once, and each of them cuts.

它确立了雇主向员工传达的内容,这关系到信赖利益,并且在许多州,视措辞和免责声明而定,还关系到合同的成立。它确立了雇主为自己设定的标准,雇主在具体案件中的行为将据此标准被衡量;一个未能遵循自身既定程序的雇主,处境要比一个程序本就含糊的雇主更为不利。它确立了雇主已知悉的内容,因为一项针对某法律要求所制定的政策,证明了雇主对该要求的知悉。它还确立了做法是否一致,因为对不同员工适用不同标准的政策,正是通向歧视认定的常见证据路径。

这些都不是新问题。新的是,这份文件如今可能是由一个以“看似合理”为优化目标的系统所生成,起草者并非律师,整个过程也没有留下任何记录,说明曾参考过何种权威依据、行使过何种判断。

当对方律师询问是谁起草了这项便利安排程序、某一条款依据什么被纳入、有何法律依据支持时,“是生成式系统写的,有人审校了语气”这样的回答并不是一个好答案。它不仅无济于事;它还证明了雇主在未确定其正确与否的情况下,便采纳了某一法律立场。

它证明了雇主在未确定其正确与否的情况下,便采纳了某一法律立场。

虚构的法律依据,以及人力资源部门为何比法务部门处境更糟

生成式系统倾向于给出自信满满、却完全虚构的法律引证,这一现象已有据可查且后果严重。2023年,纽约南区联邦法院对提交了一份包含并不存在的判例、由聊天机器人生成且未经核实的诉状的律师予以制裁。类似事件此后在众多司法辖区相继出现,这一模式如今对任何执业者而言都已不陌生。

这场讨论中值得注意的一点是,那些当事人本身就是律师。他们受过识别“引证必须核实”这一原则的训练,拥有核实所需的数据库访问权限,并承担着一旦失职便可能受到制裁的职业责任。即便如此,错误仍反复发生。

而一位负责起草多州休假政策的人力资源通才,并不具备这些优势。她往往无法判断某条被引用的条款是否真实存在,没有相应的研究工具去核实,其工作也不附带任何专业核实义务。生成的文本读起来流畅、结构清晰、措辞得当,与她读过的每一份合格政策别无二致。而流畅,恰恰是这些系统最擅长生产的东西,也恰恰是非专业读者用来判断准确性的替代标准。

最要命的错误往往也最不显眼。虚构的引证至少在有人想到要核实时是可以查证的。但一个被微妙曲解的标准却并非如此:一项将“互动流程”描述为始于员工提交文件之时、而非始于雇主意识到需求之时的政策,读起来完美无瑕,却会在两年后导致一项“未能提供便利安排”的认定。

持续对齐并非该技术的实际运作方式

“该生成式系统能随各州及各国法规变化而自动适应,确保持续对齐”这一说法,除非供应商能说明具体机制,否则应被视为不实。

语言模型存在训练截止日期。若没有针对权威且持续更新的法律来源进行检索,它输出的就是训练数据中所呈现的法律状态,且不会向用户发出答案已过时的任何信号。就业法恰恰是这一局限最致命的领域之一:薪酬透明度要求、带薪休假权利、竞业限制、薪资历史禁令以及排班规则,已在各司法辖区以远超任何静态语料库更新速度的节奏激增和演变。今天生成的一份政策,可能描述的是十八个月前已被修订的某项要求,而它会以十分自信的口吻加以描述。

与营销宣传的说法相反,多司法辖区场景恰恰是最糟糕的情形,而非最理想的情形。本地化要求判断五十多个法域中哪一个适用、它们如何相互作用、以及本地要求在何处超出

联邦最低标准。这是最依赖当下专业知识的任务,也是提出该请求的人最不可能察觉错误的场景,因为她正是由于自己不了解才提出本地化需求的。

An organization should also recognize the claim's shape. In January 2025 the Federal Trade Commission required an accessibility software provider to pay one million dollars over representations that its automated product could make any website conform to accessibility standards and maintain that conformance as content changed. Continuous automated compliance with a technical standard is the same claim, made about a different standard, and it did not survive contact with a regulator.

审计准备状态被颠倒了

该主张认为,生成的文档能带来审计准备状态。实际情况恰恰相反,而理解其原因也有助于厘清这类审查实际考察的是什么。

审计人员、监管机构调查员或对方专家,首要关心的并非某项政策是否存在——存在与否是显而易见的。他们真正关心的问题是:机构如何确定该政策满足了适用要求、是谁作出了这一判断、其依据了什么权威来源、上一次对照现行法律进行审查是何时、上一次修订时改动了什么以及为何改动。

生成的文档对这些问题一概无法回答。它制造出了文件本身,却摧毁了赋予该文件分量的记录。一家拥有内容单薄、日期陈旧、措辞朴素的政策,但配有由具名律师完成并记录在案的年度审查的机构,其处境要远远优于另一家拥有内容全面、看似最新、涵盖五十个州、却无人能说清其来源的政策体系的机构。

> **补救之道**
>
> 补救之道不是回避这些工具,而是保留来源记录:生成了什么、核实了什么、由谁核实、依据何种来源、于何时核实。这份记录在起草时几乎不费成本便可保留,但事后却无法重建。

职位描述是一份法律文书

标准清单中有一项值得单独讨论,因为它常被当作宣传文案对待,而实际上它是该职能所产出的最具后果性的文件之一。

Under the ADA (Americans with Disabilities Act), a qualified individual is one who can perform the essential functions of the position, and the implementing regulations identify written job descriptions prepared before advertising or interviewing as evidence of what those essential functions are. A generated job description that lists duties nobody actually performs, or that includes physical or sensory requirements copied from a generic template, does two things simultaneously: it screens out candidates who could perform the job, and it supplies the plaintiff with the employer's own document showing that the stated requirement was never essential.

The same document carries weight in exempt classification analysis under the Fair Labor Standards Act, where the inquiry turns on actual duties but the written description is evidence of what the employer represented those duties to be. It is also the source against which any selection procedure's job-relatedness is assessed under the Uniform Guidelines.

生成式职位描述之所以有吸引力,是因为撰写它们繁琐乏味,而生成出来的内容读起来又很不错。但正是在这一环节,一个微小的起草便利变成了证据层面的问题,而“体能要求”部分正是这种情况最常发生之处。

聊天机器人给出的建议,雇主必须为之承担责任

最后一种应用——由人工智能助手按需回答员工的政策问题——转移了一种特定且被低估的风险。

An employee who asks about leave eligibility, accommodation procedure, or complaint routing and receives an answer has received it from her employer. If the system tells her she is not eligible for leave she is entitled to, and she does not take it, the employer has interfered with a statutory right, and interference under the Family and Medical Leave Act does not require intent. If it tells her the accommodation process requires medical documentation she is not obliged to provide, and she abandons the request, the employer has a failure-to-accommodate problem with a written record of the deterrent. If it directs a harassment complaint into a channel that never reaches a decision-maker, the employer's notice defense is compromised by its own system.

由此可直接得出三项管控措施。对话应当被记录并保存,因为它们是雇主的陈述,日后会被要求提供。任何涉及休假、便利安排、投诉或调查的问题,都应转交给具名的真人处

理,而不是由系统直接作答,因为这些正是错误答案可能引发诉讼责任、且即时回复价值最低的类别。界面还应明确声明,它提供的是一般性信息,不对资格或权益作出认定。

需要建立的事项

在生成式系统产出任何机构将依赖或员工将阅读的内容之前,有六件事需要落实。

1. 为每一份生成的文件保留来源记录:产出了什么、核实了何种来源、由谁核实、于何日核实,以及审查时改动了什么。无日期、无归属的政策,是失败的表现形式。
2. Require legal review for anything stating a legal standard, entitlement, or procedure, and treat the reviewer's name and date as part of the document rather than as an internal note.
3. 对任何声称能保持监管现行性的供应商,要求其说明具体机制:检索的是何种权威来源、检索频率如何、以及当某条款发生变化时用户将如何被告知。如果答案是“该模型是基于法律材料训练的”,那么该产品并不具备其所宣称的功能。
4. 将职位描述视为法律文书。对照任职者实际从事的工作,审查其基本职能以及任何体能、感官或沟通方面的要求,并为此次审查注明日期。
5. 记录并保留面向员工的助手对话内容,并将休假、便利安排、投诉和调查方面的问题转交给专人处理。
6. 不要将生成的文档在内部或外部宣称为审计准备就绪。审计准备就绪指的是审查记录,而非文档本身。

结语

在人力资源文档工作中运用生成式辅助工具,理由是切实存在的。起草工作确实耗时,该职能部门确实资源不足,在有效监督下形成初稿是这项技术的合理用途。

这类应用与其他应用不同之处在于,其输出结果就是证据。筛选系统产生的决定,雇主事后可能难以复原其推理过程。而文档系统产生的是雇主必须为之负责的陈述——以雇主自己的口吻、以书面形式、大批量、高速度产生。生成员工手册的组织并非只是自动化了一项琐事,而是在未曾读过自己说了什么的情况下,书写了未来纠纷中属于自己的那一方立场,而且日后还会在自己无法掌控的情形下再次面对这份文件。

所需的规范其实很简单:核实、标明来源、注明日期,并保留决策人的记录。这项规范正是使用该工具与被该工具所暴露之间的全部区别。

参考文献

《1990年美国残疾人法》,42 U.S.C. § 12111(8);29 C.F.R. § 1630.2(n)(3),关于基本职能的证据,包括在招聘广告发布或面试之前准备的书面职位描述。

《家庭与医疗休假法》,29 U.S.C. § 2615(a)(1),禁止干涉受保护权利。

《公平劳动标准法》,29 U.S.C. § 201及其后条款,基于实际职责的豁免分析。

美国联邦贸易委员会,《关于*accessiBe, Inc.*案》,案号2223156,投诉与拟议裁决于2025年1月3日公布。

*Mata*诉*Avianca*航空公司案,案号22-cv-1461(南纽约地区联邦地方法院,2023年),因提交由人工智能工具生成的虚构司法引证而受到制裁。

《雇员甄选程序统一指南》,29 C.F.R. Part 1607。

> **TIPS**
>
> Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

引用方式

Grizzle, H. M. (2025, May 27). 政策即证据:生成式人工智能在人力资源文件与建议中的应用. Novara Consulting Group. https://www.novaracg.com/zh/2025/05/27/the-policy-is-the-exhibit-generative-artificial-intelligence-in-human-resources-documentation-and-advice/