

إعادة تصوّر الوكالة التنظيمية في عصر وكلاء الذكاء الاصطناعي:

نحو إطار مفاهيمي جديد للفعل والقرار والمساءلة في المنظمات

الباحث: اشرف إبراهيم الشكري

باحث في مرحلة الدكتوراه

ashrfalshkri86@gmail.com

تاريخ الارسال: 2026/5/27- تاريخ القبول 2026/7/25- تاريخ النشر 2026/8/30

الملخص: تستطيع المنظمات أن تفوض إلى الوكلاء الاصطناعيين تحليل الأهداف، واختيار مسارات العمل، واستخدام الأدوات، وتنفيذ معاملات ذات أثر خارجي؛ لكنها لا تستطيع في الوقت الراهن نقل المساءلة التنظيمية أو الأخلاقية أو القانونية إليهم على نحو مماثل. تبني هذه الورقة نظرية لهذا اللاتماثل بالاستناد إلى مراجعة أدبية تكاملية نقدية تصل بين نظرية الوكالة، والتصميم التنظيمي والتفويض، والمساءلة، والتنظيم البشري-الخوارزمي، والسيطرة البشرية ذات المعنى. يبين فحص التداخل أن المفارقة العامة، ومفاهيم فجوة المسؤولية وتداولها وانجراف المساءلة، ليست جديدة. لذلك تتمثل المساهمة في إطار «قابلية التفويض-المساءلة» الذي يشرح كيف يقود عمق التفويض واستقلالية الوكيل وسلطة التنفيذ، ولا سيما مع ترابط المهام وضغط الزمن، إلى انكماش السلطة البشرية المتبقية. وتبقى المساءلة مقترنة جوهريًا بالفعل فقط عندما تتوافر ثلاثة مسارات متكاملة: إسناد معرفي قابل للتحقق، وقدرة واقعية سابقة أو متزامنة على التدخل، وقدرة لاحقة على الإصلاح وتحمل النتائج. وتعمل قابلية التدقيق والعكس بوصفهما شروط تمكن لا بدائل عن السلطة البشرية؛ كما ترفع شدة العواقب مستوى الاقتراح المطلوب. تُصاغ سبع قضايا نظرية، وتحدد حدود التطبيق والتفسيرات المنافسة. توسع الورقة نظرية الوكالة من مشكلة توافق الحوافز إلى مشكلة صلاحية المساءلة، وتضيف إلى التصميم التنظيمي قيدًا معياريًا-تشغيليًا على كفاءة التفويض، وتفرق في أدبيات المساءلة بين الإسناد الاسمي والقدرة الفعلية على تقديم الحساب والتدخل والإصلاح.

الكلمات المفتاحية

نظرية الوكالة؛ الوكلاء الاصطناعيون؛ التفويض؛ السلطة التنظيمية؛ المساءلة؛ السيطرة البشرية ذات المعنى؛ قابلية التدقيق؛ قابلية العكس.

Abstract: Organizations can delegate goal decomposition, action selection, tool use, and externally consequential execution to artificial agents, yet they cannot symmetrically transfer organizational, moral, or legal accountability to those systems. This conceptual paper theorizes that asymmetry through a critical integrative review of agency theory, organizational design and delegation, accountability, human–algorithm organizing, and meaningful human control. The overlap scan shows that the general paradox, responsibility gaps, circulating responsibility, and accountability drift are not novel. The contribution is instead a Delegation–Accountability Viability Framework explaining how delegation depth, agent autonomy, and execution authority—especially under task interdependence and time compression—produce residual-authority compression: the erosion of

accountable humans' real capacity to constrain, interrupt, verify, or remedy delegated action despite retaining formal authority. Substantive accountability remains coupled to action only when three channels coexist: verifiable epistemic attribution, real ex ante or concurrent intervention capacity, and ex post remedial and consequence-bearing capacity. Auditability and reversibility are enabling conditions rather than substitutes for human authority, while consequence severity raises the required level of coupling. Seven propositions, boundary conditions, and rival explanations are developed. The framework extends agency theory from incentive alignment to accountability viability, adds a normative-operational constraint to delegation design, and distinguishes nominal responsibility from the practical capacity to answer, intervene, and repair.

Keywords: agency theory; artificial agents; delegation; organizational authority; accountability; meaningful human control; auditability; reversibility.

1. المقدمة:

أدخلت الخوارزميات منذ سنوات القدرة على التنبؤ والتصنيف والتوصية إلى القرار التنظيمي. لكن الوكيل الاصطناعي يغير موضع المشكلة: فهو لا يكتفي بإنتاج معلومة يقرر الإنسان على أساسها، بل قد يحلل هدفًا مفتوحًا، ويفككه إلى مهام، ويختار أدوات وأطرًا، وينفذ سلسلة أفعال، ويراقب النتائج، ويعدل المسار، بل ويفوض إلى وكلاء فرعيين. هنا لا يعود السؤال الرئيس «هل الخوارزمية دقيقة؟» بل «من يملك سلطة الفعل، ومن يستطيع تقديم حساب عنه؟» (Shavit et al., 2023; Sapkota et al., 2026; Staufer et al., 2026).

تبدو الإجابة الإدارية المباشرة سهلة: المنظمة أو المدير يظل مسؤولًا لأن النظام مجرد أداة. غير أن هذه الإجابة تخلط بين بقاء المسؤولية الرسمية وبقاء القدرة الواقعية التي تجعل المساءلة عادلة وفعالة. فإذا كان حامل المسؤولية لا يعرف المسار الذي أنتج الفعل، ولا يستطيع اعتراضه في الزمن المناسب، ولا يملك مهارة التحقق أو سلطة إبطال النتيجة، فإن وضع اسمه في خانة «الموافق» لا يعيد السيطرة؛ بل قد يحوله إلى منطقة امتصاص أخلاقية تحمل اللوم عن نظام لم يملك عليه سيطرة ذات معنى (Bainbridge, 1983; Elish, 2019).

هذا التوتر ليس بلا سوابق. تفصل نظرية التنظيم بين السلطة الرسمية والسلطة الفعلية (Aghion & Tirole, 1997)، وتبين أدبيات اللامركزية أن السلطة قد تتوزع بينما تبقى المساءلة الخارجية لدى قمة المنظمة (Lee, 2024). كما أن أدبيات الذكاء الاصطناعي وصفت فجوات المسؤولية (Matthias, 2004)، ومحدودية الشفافية (Ananny & Crawford, 2018)، وتداول المسؤولية عند إسناد الوكالة للذكاء الاصطناعي (Leonardi, 2025). الجديد القابل للدفاع ليس المفارقة ذاتها، بل تحديد شروط تتحول فيها من توتر قابل للإدارة إلى خلل بنيوي في قابلية المساءلة.

تنطلق الورقة من التساؤل الآتي: متى يظل تفويض الفعل إلى وكيل اصطناعي متوافقًا مع بقاء المساءلة التنظيمية قابلة للممارسة، ومتى يؤدي إلى تدهور اقتراحها بالفعل؟ وتجب بناء إطار قابلية التفويض—المساءلة. حجته المركزية أن

التفويض لا يكون تنظيميًا صالحًا بمجرد تحسين الأداء أو حتى قابلية التدقيق؛ بل يجب أن يترك لدى الفاعل البشري أو التنظيمي الذي سيُسأل لاحقًا سلطة متبقية متناسبة مع عمق التفويض وشدة العواقب. تتجسد تلك السلطة في القدرة، لا في اللقب: القدرة على التحديد والتقييد والمقاطعة والتحقق والإصلاح.

تسهم الورقة بثلاث طرق. توسع نظرية الوكالة من افتراض وكيل قابل للحوافز والجزاء إلى حالة منفذ غير قابل للمساءلة بذاته، فتستبدل تضارب التفضيلات وحده بمشكلة «عدم تماثل القابلية للمساءلة». وتضيف إلى التصميم التنظيمي قيدًا على التفويض هو قابلية المساءلة، لا الكفاءة المعلوماتية فقط. وأخيرًا، تفكك المساءلة إلى ثلاث مسارات مترابطة—الإسناد، والتدخل، والإصلاح—وتبين لماذا لا تكفي السجلات أو «الإنسان في الحلقة» كل على حدة.

2. المشكلة

2.1 لماذا لا تكفي الصياغة التقليدية لنظرية الوكالة؟

تعرف نظرية الوكالة العلاقة عندما يفوض أصيل عملًا إلى وكيل، فتظهر تكاليف مراقبة وربط وضياح متبقي بسبب اختلاف الأهداف وعدم تماثل المعلومات (Jensen & Meckling, 1976; Eisenhardt, 1989). يمكن إدراج الذكاء الاصطناعي ظاهريًا كوكيل: ينفذ نيابة عن المنظمة وقد ينحرف عن مقاصدها. لكن التشبيه ينهار في نقطة حاسمة. الوكيل البشري يملك مصالح، ويمكن التفاوض معه ومكافأته أو معاقبته واستدعاؤه لتفسير سلوكه. أما النظام الاصطناعي فيمكن التحكم في صلاحياته أو تحديثه أو إيقافه، لكنه ليس في الإطار القانوني والتنظيمي الراهن حاملًا مستقلًا لالتزام الجواب أمام منتدي أو لنتائج أخلاقية وعقابية.

لذلك لا تكون المشكلة مجرد «انحراف الوكيل» بل انفصال بين من ينتج التسلسل السببي ومن يتحمل علاقة المساءلة. يمكن للمنظمة بالطبع تحميل المطور أو المستخدم أو المورد التزامات تعاقدية، غير أن توزيع المسؤولية عبر سلسلة القيمة لا يجعل النظام ذاته قابلاً للجواب ولا يضمن أن أي طرف منفرد كان يملك المعرفة والسيطرة والقدرة على الإصلاح (Cooper et al., 2022; Cobbe et al., 2021). ومن هنا يلزم توسيع متغيرات النظرية من الحافز والمعلومات إلى سلطة التنفيذ وقابلية المقاطعة والعكس.

2.2 لماذا لا تكفي أدبيات الشفافية والتدقيق؟

تقدم المراجعات والخوارزميات القابلة للتدقيق سجلات مهمة تربط الأفعال بالمدخلات والإصدارات والأذونات. وقد طورت أدبيات المساءلة الخوارزمية أطرًا للتدقيق الداخلي عبر دورة الحياة والمراجعة والاعتراض (Kroll et al., 2017; Raji et al., 2020; Wieringa, 2020). لكن رؤية السجل لا تعني فهم سبب الفعل أو القدرة على اختبار مبرراته، كما أن الفهم اللاحق لا يعيد معاملة غير قابلة للعكس ولا يمنح المشرف سلطة سابقة لم يملكها. الشفافية قد تنتج «رؤية بلا معرفة» إذا كان حجم الآثار أو تعقيد السلسلة يتجاوز قدرة المنتدي أو المنظمة على الفحص (Ananny & Crawford, 2018).

لهذا تميز الورقة بين ثلاثة أبعاد لقابلية التدقيق: الإسناد، أي ربط الفعل بنظام وإصدار ومفوض وسياق؛ الفهم، أي تقديم منطوق وبدائل وقيود بطريقة تسمح بالحكم؛ والتحقق، أي قدرة طرف مستقل على إعادة اختبار الادعاءات ومقارنتها بمعيار. لا تخفف قابلية التدقيق تدهور المساءلة إلا إذا اتصلت هذه الأبعاد بمنتهى ذي سلطة وبمسار تدخل أو علاج.

2.3 موضع الفجوة الدقيقة

يقدم Hillebrand وزملاؤه (2025) منظورًا تنظيميًا لإدارة الذكاء الاصطناعي، ويشرح Murray وزملاؤه (2021) الوكالة المقترنة، وبين (Leonardi, 2025) كيف تحرك إسنادات الوكالة شعور القوة والمسؤولية. ويقترح South وزملاؤه (2025) بنية تفويض موثق ومصرح ومدقق، فيما يضع Shavit وزملاؤه (2023) الإسناد والمقاطعة ووضوح نشاط الوكيل ضمن ممارسات الحوكمة. ومع ذلك، تظل الحلقة النظرية الناقصة هي آلية تربط مقدار التنفيذ المفوض بالحد الأدنى من القدرة الإنسانية اللازمة لجعل المساءلة غير صورية، وتبين كيف تتفاعل قابلية التدقيق والعكس وشدة العواقب مع هذا المقدار.

الفجوة النظرية المحددة

لا تفتقر الأدبيات إلى مفاهيم المسؤولية أو الوكالة أو الإشراف: بل تفتقر إلى قاعدة تنظيمية تركيبية تقول إن صلاحية التفويض تتوقف على تكافؤ نسي بين «القدرة المفوضة على إحداث الأثر» و«القدرة المحتفظ بها على معرفة الأثر وتقبيده وإصلاحه». عندما ينكسر هذا التكافؤ تصبح المساءلة اسمًا بلا قدرة.

3. المنهج: مراجعة أدبية تكاملية نقدية

اعتمدت الورقة مراجعة تكاملية نقدية ملائمة لبناء المفاهيم والنظرية عبر أدبيات متباعدة، لا مراجعة منهجية بغرض حساب أثر متوسط أو حصر كل منشور (Torraco, 2005; Snyder, 2019; Post et al., 2020). اتبعت العملية أربع حركات تحليلية. الأولى تحديد الظاهرة: انتقال الذكاء الاصطناعي من التوصية إلى التنفيذ المتسلسل. والثانية تفكيك المشكلة إلى أسئلة: من يملك حق القرار، من يملك السيطرة الفعلية، من يعرف، من يستطيع التدخل، ومن يجيب أمام من. والثالثة مقارنة إجابات التيارات النظرية واكتشاف افتراضاتها غير المشتركة. والرابعة تركيب مفاهيم وقضايا قابلة للتخطئة بدل تلخيص وصفي للأدبيات (Jaakkola, 2020).

4. الأساس الأول: من نظرية الوكالة إلى عدم تماثل القابلية للمساءلة

4.1 العلاقة الثلاثية بدل الثنائية

في العلاقة الكلاسيكية يوجد أصيل ووكيل وعقد أو ترتيب يحدد الأداء والحوافز. أما في الحالة الوكيلية تنظيميًا فتوجد على الأقل ثلاث طبقات: صاحب غاية أو حق قرار، وفاعل اصطناعي ينفذ، وفاعل بشري/تنظيمي يظل قابلاً للمساءلة.

وقد تتضاعف الطبقات بوجود مزود نموذج، وباني منصة، ومنفذ، ومستخدم، وأطراف متأثرة. هذا التعدد لا يلغي المساءلة، لكنه يرفع كلفة تحديد أي طرف كان قادرًا على منع الضرر أو تفسيره أو علاجه (Wieringa, 2020; Cooper et al., 2022).

تسمي الورقة الفجوة الناتجة «عدم تماثل القابلية للمساءلة»: القدرة على العمل تنتقل نحو مكون لا يستطيع أن يكون طرفًا كاملاً في علاقة المساءلة، بينما يبقى واجب الجواب على فاعلين قد تتناقص قدرتهم على السيطرة. يختلف هذا عن عدم تماثل المعلومات التقليدي؛ فقد يعرف المطور والناشر والمستخدم أجزاء مختلفة، ولا يمتلك أي منهم صورة كاملة أو حق تدخل شامل. كما يختلف عن تضارب الأهداف؛ فقد يحدث الخلل حتى لو اتبع الوكيل الهدف المبرمج، لأن الهدف نفسه ناقص أو يتفاعل مع سياق غير متوقع.

4.2 السلطة الرسمية والسلطة الفعلية

تفصل Aghion و(1997 Tirole) بين الحق الرسمي في اتخاذ القرار والسيطرة الفعلية التي تتولد من المعلومات والمبادرة. ويوضح هذا الفصل لماذا قد يبقى المدير «موافقاً» رسميًا بينما يملك الوكيل السلطة الفعلية: يقترح بسرعة، ويجمع البيانات، ويحدد البدائل، وينفذ قبل أن يستطيع الإنسان تكوين حكم مستقل. وتبين دراسة Dobrjaska وزملائها (2015) أن ترابط القرارات وكثافة المعلومات يؤثران في عمق التفويض داخل التسلسل الهرمي. في الوكلاء الاصطناعيين تتضاعف هذه الآلية لأن تفويض طبقة مبكرة—كتفكيك الهدف—يعيد تشكيل البدائل التي ستصل أصلاً إلى الإنسان.

تؤدي الأتمتة أيضًا إلى تآكل مهارة المشرف ونموذجه الذهني، وهي مفارقة قديمة في أتمتة الأنظمة المعقدة (Bainbridge, 1983). وكلما نجح الوكيل مدة أطول، قل تدريب البشر على حالات الاستثناء، وزادت تبعية المنظمة له، وارتفعت كلفة إيقافه. لذلك قد تقلص السلطة الفعلية بمرور الزمن من دون أي تعديل رسمي للصلاحيات.

5. الأساس الثاني: المساءلة كعلاقة مؤسسية

5.1 من المسؤولية إلى تقديم الحساب

يعرّف (Bovens 2007) المساءلة كعلاقة يكون فيها فاعل ملزمًا بشرح سلوكه وتبريره أمام منتدى يستطيع الاستجواب والحكم وفرض نتائج. وتفرق هذه الصياغة بين كون المنظمة «مسؤولة» بصورة عامة ووجود عملية قابلة للتشغيل. كما يميز (Bovens 2010) بين المساءلة كفضيلة وبينها كآلية. يفيد هذا الفرق في الذكاء الاصطناعي: يمكن للمنظمة إعلان فضائل الشفافية والنزاهة، لكن آلية المساءلة تفشل إذا لم يوجد سجل صالح، أو منتدى مخول، أو شخص يملك جوابًا، أو نتيجة مترتبة.

يحلل Novelli وزملاؤه (2024) المساءلة في الذكاء الاصطناعي عبر قابلية الجواب وشروط الاعتراف بالسلطة والاستجواب وتقييد القوة. وتبين (Busuioc 2021) أن تعقيد الخوارزميات وعتامتها يقيدان قدرة المؤسسات على أداء واجب تقديم الحساب. أما (Cech 2021) فيعيد الانتباه إلى وكالة المنتدى ذاته: لا تكفي مطالبة المنظمة بالحساب إذا لم يملك المنتدى معرفة أو وصولًا أو سلطة. لذلك يعتمد الاقتران الجوهرى على جانبي العلاقة، لا على النظام وحده.

5.2 ثلاث مسارات لاقتران المساءلة

تقترح الورقة أن المساءلة تقتزن بالفعل عبر ثلاث مسارات. مسار الإسناد المعرفي يجيب: ما الذي حدث، وبأي نموذج وإصدار وأذونات، وبأي تفويض، وما البدائل التي استُبعدت؟ مسار التدخل يجيب: من كان يستطيع تقييد مجال العمل، أو طلب موافقة، أو إيقاف السلسلة، وبأي زمن وموارد؟ ومسار الإصلاح وتحمل النتائج يجيب: من يستطيع إبطال النتيجة، وتعويض المتضرر، وتصحيح النظام، ومن يواجه حكم المنتدى؟

هذه المسارات متكاملة وليست بدائل. قد توجد قدرة توقف بلا معرفة كافية، فتنتج قرارات عشوائية أو اعتماداً أعمى. وقد توجد سجلات ممتازة بلا فرصة تدخل، فتتحول المساءلة إلى تشريح متأخر. وقد يوجد تدخل ومعرفة بلا قدرة علاج أو جزاء، فتفقد علاقة المنتدى أثرها. الاقتران الجوهرى هو الحد الأدنى المشترك الذي يجعل واجب الجواب قابلاً للإنصاف والفعالية.

6. الأساس الثالث: الوكالة البشرية-الاصطناعية والتنظيم

تعامل Faraj وزملاؤه (2018) مع الخوارزميات المتعلمة بوصفها مشاركا في التنظيم، وتفحص Shrestha وزملاؤه (2019) هياكل قرار بشرية وخوارزمية هجينة، وبين Kellogg وزملاؤها (2020) أن السيطرة الخوارزمية تخلق مجالات جديدة للصراع والمقاومة. ويفصل Murray وزملاؤه (2021) أشكال الوكالة المقترنة بحسب موضع التقنية في تطوير بروتوكول الروتين واختيار الفعل. هذه الأدبيات مهمة لأنها تمنع تصوير الوكيل ككيان مستقل تماماً: الأداء ينتج من ترتيب اجتماعي-تقني، لكن هذا لا يعني أن الإسناد المؤسسي يجب أن يذوب في الشبكة.

يكشف منظور Hillebrand وزملائه (2025) أن البشر يستخدمون الذكاء الاصطناعي للإدارة ويُدارون به في آن، وأن التحليل يجب أن ينتقل من مهمة إلى سياق تنظيمي ومن وكيل فردي إلى وكالة جماعية. ويضيف (Krakowski 2025) أن الذكاء التوليدي يعيد تشكيل الوكالة أكثر من مجرد التنبؤ، بينما يبين (Leonardi 2025) أن البشر الذين ينسبون الوكالة إلى النظام قد يشعرون بانخفاض السيطرة والمسؤولية حتى عندما يحتفظون رسمياً بهما. ينتج عن ذلك حلقة: الأداء الناجح يشجع الإسناد، والإسناد يشجع تفويضاً أعمق، والتفويض يخفف الاحتكاك البشري، فيتراجع النموذج الذهني وقدرة الاعتراض، فتزداد السلطة الفعلية للنظام.

في المنظمات المنسقة بالذكاء الاصطناعي لا يظل التفويض ثنائياً بين مدير وأداة؛ بل ينتشر عبر وكلاء وخدمات وواجهات وأطراف خارجية، وتتحرك حقوق القرار على مستوى البنية التنظيمية (van Esch, 2026). لذلك يجب أن تكون وحدة التحليل «سلسلة فعل مفوضة» تربط الغاية والإذن والتخطيط والتنفيذ والنتيجة والمنتدى، لا نموذجاً منفرداً أو مستخدماً منفرداً.

7. إعادة بناء مفهوم الوكالة التنظيمية

تُعرّف الوكالة التنظيمية هنا بأنها قدرة منظمة على تكوين غايات معترف بها، وتخصيص سلطة العمل لتحقيقها، والحفاظ على علاقة قابلة للتتبع والتدخل والإصلاح بين الأفعال والفاعلين الذين يقدمون الحساب عنها. يميز التعريف بين

القدرة السببية على إحداث أثر وبين المكانة المؤسسية التي تسمح بإسناد الحساب. يمكن لوكيل اصطناعي أن يمتلك قدرة سببية وتشغيلية عالية من دون أن يمتلك المكانة الثانية.

بذلك لا تُلغى فكرة الوكالة الموزعة؛ فالنتيجة قد تكون ثمرة بشر ونماذج وبيانات وواجهات وموردين. لكن المنظمة تظل «مركز اقتران» لأنها تصمم أو تقبل ترتيبات التفويض، وتحدد مجال الصلاحيات، وتختار المنتديات ومسارات التصعيد والتعويض. توزيع السببية لا يبرر تبخر الجواب. وفي الوقت نفسه، لا يجوز تحويل أقرب إنسان إلى حامل صوري للذنب إذا لم يكن يملك سلطة متناسبة مع ما طُلب منه مراقبته؛ وإلا نشأت منطقة امتصاص أخلاقية (Elish, 2019).

8. المتغيرات والمفاهيم التشغيلية

الجدول رقم (1): المفاهيم والمتغيرات التشغيلية لإطار قابلية التفويض—المساءلة ومؤشرات قياسها المقترحة

المفهوم	التعريف	مؤشرات قياس مستقبلية
استقلالية الاصطناعي	الوكيل	مدى قدرة النظام على اختيار التسلسل والأداة والتوقيت والتكيف دون موافقة مباشرة.
عمق التفويض	عدد طبقات سلسلة الفعل وأهميتها التي تنتقل من الإنسان إلى النظام.	عدد الخطوات غير المعتمدة؛ تنوع الأدوات؛ الاستمرارية؛ التكيف مع هدف مفتوح.
سلطة التنفيذ	اتساع الموارد والأفعال الخارجية التي يستطيع الوكيل التأثير فيها فعلياً.	حدود مالية/بيانية؛ صلاحيات الكتابة؛ التواصل مع أطراف؛ أثر مادي أو قانوني.
السلطة البشرية المتبقية	القدرة الفعلية—لا الرسمية فقط—على تقييد الوكيل ومقاطعته والتحقق منه وإصلاح أثره.	زمن التدخل؛ مهارة المشرف؛ وصوله للأدلة؛ حق النقض؛ دعم مؤسسي؛ بديل تشغيلي.
قابلية التدقيق	قابلية ربط الفعل بسياقه وفهم مبرره والتحقق منه مستقلاً.	اكتمال؛ قابلية تفسير المسار؛ إعادة الاختبار؛ حفظ الأدلة.
قابلية العكس	إمكان وقف الأثر أو إعادته أو تعويضه قبل انتشاره.	زمن الاسترجاع؛ نسبة الأفعال القابلة للإلغاء؛ كفاية التعويض؛ نطاق الانتشار.
شدة العواقب	حجم الضرر المحتمل واتساعه واحتمال عدم إصلاحه.	سلامة وحقوق؛ أثر مالي/سمعة؛ عدد المتأثرين؛ حساسية الوقت.

المفهوم	التعريف	مؤشرات قياس مستقبلية
اقتران المساءلة	مدى اتصال سلطة الفعل بمسارات الإسناد والتدخل والإصلاح لدى فاعل ومنتدى قابلين للعمل.	توافر المسارات الثلاثة وجودتها؛ وضوح صاحب الحساب؛ فعالية الحكم والجزاء.
انكماش السلطة البشرية المتبقية	عملية تدهور قدرة حامل المساءلة على السيطرة الواقعية مع بقاء سلطته الرسمية.	فجوة بين الحق الرسمي والزمن/المعرفة/المهارة/البديل الفعلي.
تدهور اقتران المساءلة	تحول ديناميكي يصبح فيه مستوى الاقتران المتاح أدنى من المستوى الذي تتطلبه شدة الأثر.	تزايد الحوادث غير القابلة للإسناد أو الاعتراض أو الإصلاح؛ لجان شكلية؛ تأخر الاستجابة.

9. إطار قابلية التفويض-المساءلة

يفترض الإطار أن عمق التفويض وترايط المهام يرفعان استقلالية الوكيل وسلطة تنفيذه. لا ينتج الخلل من الاستقلالية في حد ذاتها، بل من نسبتها إلى السلطة البشرية المتبقية. عندما يصبح مسار الفعل أسرع أو أعقد أو أكثر اعتماداً على بنى خارجية مما يستطيع حامل المساءلة فهمه أو مقاطعته، يحدث انكماش للسلطة الفعلية. وتحدد قابلية التدقيق والعكس مقدار ما يمكن استعادته من القدرة، بينما تحدد شدة العواقب مقدار الاقتران المطلوب أصلاً.

يمكن التعبير عن منطق الإطار بصورة معيارية مبسطة: تكون المساءلة قابلة للممارسة عندما تكون قدرة الاقتران المتاحة—الإسناد × التدخل × الإصلاح—متناسبة على الأقل مع الطلب الذي تخلقه شدة العواقب، وبعد تعديله بعمق التفويض وسلطة التنفيذ وقابلية العكس. ليست هذه معادلة قياس جاهزة، بل منطق تكويني: ضعف شديد في أي مسار قد لا يعوضه ارتفاع المسارين الآخرين، خصوصاً في الأفعال غير القابلة للعكس.

10. القضايا النظرية

القضية 1 — الاستقلالية النسبية

كلما ارتفعت استقلالية الوكيل وسلطة تنفيذه نسبياً إلى السلطة البشرية المتبقية، زاد احتمال انكماش السيطرة الفعلية وتدهور اقتران المساءلة، حتى لو ظلت حقوق القرار الرسمية لدى الإنسان.

توضح القضية أن الاستقلالية ليست خاصية تقنية مطلقة؛ إذ قد يكون وكيل عالي الاستقلالية داخل حيز محدود وعكوس أقل إشكالاً من نظام أقل استقلالية يملك وصولاً إلى أفعال شديدة الأثر. المتغير الحاسم هو النسبة بين ما يستطيع النظام فعله وما يستطيع حامل الحساب معرفته ومنعه أو إصلاحه.

القضية 2 — عمق التفويض وترابط المهام

يقوى الأثر الموجب لاستقلالية الوكيل وسلطة تنفيذه في انكماش السلطة البشرية المتبقية كلما زاد عمق التفويض وترابط القرارات؛ لأن تفويض الطبقات المبكرة يعيد تشكيل البدائل والمعلومات المتاحة للرقابة اللاحقة.

تفويض تنفيذ خطوة بعد قرار بشري أقل عمقاً من تفويض صياغة المشكلة وتحديد المعيار وتوليد الخيارات والاختيار والتنفيذ. وعندما تعتمد الخطوات على بعضها، يصبح فحص النهاية دون إعادة بناء السلسلة غير كافٍ. لذلك لا ينبغي قياس التفويض بعدد الموافقات البشرية فقط، بل بموضعها داخل سلسلة تكوين الفعل.

القضية 3 — قابلية التدقيق المشروطة

تخفف قابلية التدقيق تدهور اقتران المساءلة فقط عندما تجمع بين الإسناد والفهم والتحقق، وتتصل بمنتدى ذي سلطة وبمسار فعلي للتدخل أو الإصلاح؛ أما السجلات غير المفهومة أو غير القابلة للاختبار فلا تكفي.

يمكن لسجل كامل أن يوثق مساراً لا يستطيع أحد الحكم عليه، كما يمكن لتفسير مقنع لغوياً أن يكون غير وفيّ للسبب الفعلي. ولهذا تُفاس قابلية التدقيق بالقدرة التي تمنحها للفاعل والمنتدى، لا بحجم البيانات المخزنة.

القضية 4 — قابلية العكس كمساحة تعلم

ترفع قابلية العكس مستوى التفويض الذي يمكن للمنظمة تحمله دون تدهور جوهري في اقتران المساءلة، لأنها تحول بعض الأخطاء من نتائج نهائية إلى تجارب قابلة للتصحيح؛ ويضعف هذا الأثر عندما تكون الأضرار الاجتماعية أو القانونية غير قابلة للإرجاع تقنياً.

إلغاء معاملة مالية قد يكون ممكناً، لكن كشف بيانات أو إهانة متلقٍ أو حرمانه من حق لا يعود كاملاً بحذف السجل. لذلك يجب أن تشمل قابلية العكس الإصلاح الاجتماعي والقانوني، لا مجرد rollback تقني.

القضية 5 — شدة العواقب والسلطة المتبقية

كلما ارتفعت شدة العواقب وانخفضت قابلية العكس، ارتفع الحد الأدنى المطلوب من السلطة البشرية المتبقية قبل التنفيذ، بما في ذلك الموافقة ذات المعلومات الكافية، وحق النقض، والوقت والموارد لممارسة الحكم المستقل.

تتوافق القضية مع منطق قانون الذكاء الاصطناعي الأوروبي الذي يربط الإشراف البشري بالكفاءة والتدريب والسلطة والدعم، لكنها تحوله إلى حجة نظرية: ليست هذه سمات امتثال شكلية، بل مكونات للقدرة التي تجعل الإسناد الأخلاقي والتنظيمي معقولاً.

القضية 6 — قصور الإشراف الاسمي

لا يؤدي وجود إنسان في الحلقة إلى تحسين اقتران المساءلة إذا افتقر إلى الوقت أو الكفاءة أو المعلومات أو السلطة أو البديل التشغيلي؛ وقد يزيد تركّز اللوم على المشرف مع تقليل سيطرته، فتظهر منطقة امتصاص أخلاقية.

الموافقة تحت ضغط عشرات الحالات، أو الاعتماد على ملخص الوكيل نفسه، أو استحالة تعطيل الخدمة دون شل المنظمة، كلها تحول الإشراف إلى طقس. وبهذا يمكن لتصميم يُعلن أنه بشري-المركز أن يفاقم عدم عدالة الإسناد.

القضية 7 — التكوينات ومسار التدهور

يبلغ تدهور اقتران المساءلة أقصاه في تكوين يجمع استقلالية وسلطة تنفيذ وعمق تفويض مرتفعة مع قابلية تدقيق وعكس وسلطة بشرية متبقية منخفضة؛ وكلما زادت شدة العواقب ولّد هذا التكوين ضغطاً لإعادة المركزية أو تقوية القيود والتدقيق ومسارات الإصلاح.

هذه قضية تكوينية لا خطية. قد تعوض قابلية العكس المؤقتة عن ضعف التدخل في بيئة اختبار، وقد تعوض حدود صلاحيات ضيقة عن عتامة نسبية في مهمة منخفضة المخاطر. لكن لا يوجد متغير سحري: عندما تجتمع الشروط السلبية تصبح إعادة ربط المساءلة ضرورة تنظيمية، إما بتقليل التفويض أو بزيادة السلطة المتبقية أو بكليهما.

11. الشروط الحدودية

الجدول رقم (2): الشروط والقيود الحدودية الحاكمة لتطبيق الإطار النظري

الحد	كيف يقيد النظرية
نوع النظام	ينطبق بقوة على الأنظمة التي تخطط وتستخدم أدوات وتنفذ أفعالاً خارجية؛ يضعف في أدوات الصياغة أو التوصية التي لا تملك صلاحية تنفيذ.
نوع المهمة	المهام المهيكلية منخفضة المخاطر والعكوسة تسمح بتفويض أعمق؛ المهام الغامضة القيمة أو المرتبطة بالحقوق تتطلب اقتراناً أعلى.
الزمن	في البيئات البطيئة يمكن للإنسان التحقق والتدخل؛ ضغط الزمن يجعل حق النقض الرسمي غير قابل للممارسة.
الترايط والمنظومة	التبعيات بين وكلاء وموردين وخدمات خارجية تزيد تشتت المعلومات والقدرة على العلاج.
المؤسسة القانونية	الإطار يفترض عدم تمتع الوكيل الحالي بشخصية قانونية مستقلة كاملة؛ تغير هذا الشرط سيغير توزيع الحساب لا الحاجة إلى الإسناد والتدخل والإصلاح.
المعيارية التنظيمية	في القطاعات المنظمة تُرفع عتبة الاقتران المطلوبة بفعل حقوق المتأثرين وواجبات الإشراف؛ في البيئات التجريبية قد تكون العتبة أدنى.
الاعتمادية المتراكمة	النجاح الطويل قد يضعف مهارات البديل البشري ويزيد كلفة الإيقاف، فتتغير صلاحية التفويض مع الزمن حتى دون تغير تقني.

12. التفسيرات المنافسة

12.1 التفسير التقني: السجلات والمحاذاة تكفي

قد يقال إن تحسين المحاذاة والاختبار والسجلات يحل المشكلة. يرد الإطار بأن هذه الآليات تخفض احتمال الخطأ وتحسن الإسناد، لكنها لا تخلق تلقائياً منتدى قادراً أو حق تدخل أو موارد علاج. نظام موثوق ليس نظاماً قابلاً للمساءلة بالضرورة؛ فقد ينجح وفق هدف غير مشروع أو يفشل نادراً في حالة كارثية لا يستطيع أحد إيقافها.

12.2 التفسير القانوني: منح الوكيل شخصية أو تأمينا

قد يُقترح نقل المسؤولية إلى كيان اصطناعي أو صندوق تأمين. يمكن لذلك إعادة توزيع التعويض، لكنه لا يعالج بالضرورة من صمم مجال الفعل أو من كان يستطيع منع الضرر. كما قد يسمح بغسل المسؤولية عبر كيان محدود الأصول. لذلك تظل الشخصية القانونية المحتملة شرطاً حدودياً، لا بديلاً عن تحليل القدرة الفعلية.

12.3 التفسير الاجتماعي-التقني: الوكالة موزعة فلا معنى لمساءلة طرف

صحيح أن السببية موزعة، لكن المساءلة ليست نسخة مطابقة من السببية؛ إنها علاقة مؤسسية مصممة لتوليد التعلم والشرعية والجزاء. يمكن توزيع واجبات متعددة مع تعيين مركز حساب نهائي، شريطة ألا يتحول هذا المركز إلى حامل لوم بلا قدرة. الإطار لا يعيد الفردانية، بل يطلب وصل التوزيع السببي بتوزيع صريح للقدرة والجواب.

12.4 التفسير التقليدي: الذكاء الاصطناعي مجرد أداة

يصح الوصف عندما لا يتجاوز النظام توصية قابلة للمراجعة. لكنه يصبح مضللاً عندما يختار تسلسلاً، ويستدعي أدوات، ويجري معاملات، ويتكيف ويستمر دون مراجعة لكل خطوة. ليست القضية أن النظام أصبح شخصاً، بل أن الأداة حصلت على سلطة تنفيذ تغير بنية التفويض. تجاهل هذا التغير يحافظ على لغة المسؤولية ويخسر السيطرة الفعلية.

13. المناقشة

تغير الوكلاء الاصطناعيون السؤال التنظيمي من «أين نضع القرار؟» إلى «أي مزيج من القرار والتنفيذ والمعرفة والإصلاح يمكن نقله مع بقاء المنظمة قادرة على الجواب؟». هذا التحول يفسر لماذا تبدو بعض ترتيبات الحوكمة سليمة على الورق وفارغة في التشغيل. فالمدير يملك حق النقض، لكن التنبيه يصل بعد التنفيذ؛ والمراجع يملك السجل، لكنه لا يفهم اعتماداته؛ والمنظمة تملك عقد المورد، لكنها لا تستطيع إعادة بناء النموذج أو إصلاح أثره؛ والمستخدم يضغط موافقة لأن البديل غير متاح.

يقدم مفهوم انكماش السلطة البشرية المتبقية آلية تصل هذه الأعراض. يحدث الانكماش عبر أربع عمليات: سبق زمني، عندما ينفذ الوكيل أسرع من دورة القرار؛ ضغط معرفي، عندما يتجاوز مسار الفعل قدرة المراجع؛ اعتماد تشغيلي، عندما يصبح الإيقاف مكلفاً أو مستحيلاً؛ وتشتت مؤسسي، عندما تتوزع المعرفة والحقوق بين موردين ووحدات. قد تجتمع العمليات أو تعمل منفردة، وهي قابلة للقياس مستقبلاً.

كما يميز الإطار بين مستويين للضبط. «حدود الإذن» تحدد ما يستطيع الوكيل فعله تقنياً أو تعاقدياً، و«حدود التفويض المشروع» تحدد ما يجوز نقله في ضوء مستوى الاقتران المتاح وشدة العواقب. قد يكون الفعل مصرحاً تقنياً لكنه غير صالح تنظيمياً لأن المنظمة لا تستطيع تفسيره أو إصلاحه. هذه التفرقة تحول الحوكمة من قائمة ضوابط عامة إلى قرار تصميمي حول صلاحية كل سلسلة فعل.

14. الإسهامات النظرية

14.1 إسهام في نظرية الوكالة

تضيف الورقة بُعد قابلية للمساءلة إلى مشكلة الوكالة. فإضافة إلى اختلاف الأهداف وعدم تماثل المعلومات والمخاطر، توجد لا تماثلية في قابلية أطراف العلاقة لتقديم الحساب وتحمل الجزاء. ويصبح ضبط الوكيل ليس فقط محاذاة حوافزه—وهو تعبير غير دقيق للأنظمة—بل تصميم حيز عمل وأذونات ومقاطعة وتحقيق يجعل الفاعلين القانونيين قادرين على الوفاء بالحساب الذي يبقى عليهم.

وتحول وحدة التحليل من عقد ثنائي إلى سلسلة تفويض متعددة الأطراف، وتدخل الزمن والعكس ضمن تكاليف الوكالة. التكلفة ليست مراقبة الأداء فحسب؛ بل الحفاظ على مهارة التدخل والبدل التشغيلي وحقوق الوصول والأدلة ومسار التعويض. ومن ثم قد يكون تفويض يزيد الإنتاجية المباشرة غير كفؤ إذا خلق التزامًا لا يمكن إدارته أو ضررًا غير قابل للعكس.

14.2 إسهام في التصميم التنظيمي والتفويض

تضيف الورقة «قيد قابلية المساءلة» إلى منطق التفويض. تقليديًا يُفاضل التصميم بين مزايا المعلومات المحلية والمبادرة وبين فقدان السيطرة (Aghion & Tirole, 1997). في حالة الوكيل الاصطناعي قد يمتلك النظام معالجة مركزية وسرعة وتوسعًا، لكن المنظمة تحتاج أيضًا إلى الاحتفاظ بسلطة متبقية عملية. بذلك يصبح السؤال ليس مدى عمق التفويض الأمثل فقط، بل كيفية تقسيم السلسلة بحيث تقع نقاط الموافقة قبل التحولات غير القابلة للعكس.

كما تفسر الورقة لماذا لا تتطابق السلطة الرسمية والفعلية في أنظمة الوكلاء. تتكون السلطة الفعلية من موضع النظام في التدفق، وسرعته، وسيطرته على عرض البدائل، وصلاحيات الكتابة، وتبعية المنظمة له. ولذلك ينبغي أن تقيس بحوث التصميم التنظيمي «سلطة التنفيذ» بوصفها متغيرًا منفصلًا عن حق القرار الرسمي.

14.3 إسهام في أدبيات المساءلة

تفكك الورقة الاقتران الجوهرى إلى الإسناد والتدخل والإصلاح. يوضح هذا لماذا تختلف الشفافية عن المساءلة، ولماذا قد يكون الإشراف اسميًا، ولماذا لا يكفي تعيين مالك للنظام. كما تجعل شدة العواقب محددًا لكمية الاقتران المطلوبة: ليست كل سلسلة فعل بحاجة إلى المستوى نفسه من الموافقة البشرية، لكن كل زيادة في الأثر أو اللاعكوسية ترفع عبء الاحتفاظ بالقدرة.

وتعيد صياغة «انجراف المساءلة» كعملية قابلة للتفسير لا شعار: هو تدهور نسبي بين الاقتران المتاح والمطلوب، ينشأ غالبًا مع زيادة الاعتماد والسرعة والتوسع من دون تحديث للسلطة المتبقية. هذه الصياغة لا تدعي ملكية المصطلح، بل تعطيه شروطًا وآلية ومخرجات قابلة للاختبار.

15. الآثار العملية

لا ينبغي للمنظمات اعتماد حوكمة موحدة لكل وكيل. يترجم الإطار إلى «مراجعة صلاحية التفويض» قبل الإطلاق ولكل تغيير جوهري. تبدأ المراجعة برسم سلسلة الفعل: من يحدد الهدف، ويولد البدائل، ويختار، وينفذ، ويتكيف، ومن يستطيع إيقاف كل انتقال. ثم تصنف شدة العواقب واللاعكوسية، وتختبر المسارات الثلاثة للمساءلة بأدلة تشغيلية لا سياسات فقط.

الجدول رقم (3): مجالات التصميم التنظيمي والممارسات الإجرائية لمراجعة صلاحية التفويض

مجال التصميم	اختبار أو ممارسة
حدود الإذن	صلاحيات بأقل امتياز، حدود مالية وبيانية، مدة قصيرة للأذونات، منع التفويض الفرعي غير المصرح.
نقاط عدم الرجوع	وضع الموافقة البشرية قبل النشر أو الدفع أو الإفصاح أو تغيير الحقوق، لا بعده.
جودة الإشراف	قياس زمن المراجعة، الوصول للأدلة الأولية، كفاءة المشرف، ومعدل الاعتراض المبرر؛ لا مجرد تسجيل وجود موافق.
قابلية التدقيق	حفظ هوية النظام والإصدار والأذونات والبيانات والأدوات ومسار القرار ونتائج التحقق المستقل.
قابلية العكس والإصلاح	rollback تقني، وقف الانتشار، قنوات تظلم وتعويض، وتمويل مسبق للإصلاح.
احتياطي الكفاءة	تدريب دوري وتمارين فشل وبديل يدوي أو مستقل يمنع اعتمادًا يجعل الإيقاف غير ممكن.
مالك حساب نهائي	تعيين وحدة/قيادة تجيب للمنتدى مع منحها فعلاً ميزانية وحق إيقاف ووصولاً للموردين.
مراقبة الانكماش	مؤشرات فجوة زمن التدخل، نسبة الأفعال غير القابلة للعكس، تبعية المورد، وتراجع خبرة البشر.

تتسق هذه الممارسات مع AI RMF في وظائف الحوكمة والرسم والقياس والإدارة (NIST, 2023)، ومع اشتراط قانون الذكاء الاصطناعي الأوروبي أن يُسند الإشراف إلى أشخاص ذوي كفاءة وتدريب وسلطة ودعم (European Union, 2024). لكنها لا تقدم امتثالاً قانونياً؛ فهي منطق تصميم عام يجب مواءمته مع القطاع والولاية القانونية.

16. أجندة بحث مستقبلية

✓ تطوير مقياس لعمق التفويض يميز بين تفويض صياغة المشكلة وتوليد البدائل والاختيار والتنفيذ والتعديل والتفويض الفرعي.

- ✓ قياس فجوة السلطة الرسمية-الفعلية عبر زمن المقاطعة، وتوافر البديل، وإمكانية وصول المشرف إلى أدلة مستقلة عن الوكيل.
- ✓ اختبار القضايا بتصاميم سيناريو وتجارب ميدانية تقارن بين تكوينات الاستقلالية والتدقيق والعكس وشدة العواقب.
- ✓ دراسة طولية لكيفية تأكل خبرة البشر وارتفاع تبعية المنظمة، وما إذا كانت تمارين الفشل والاحتياطي البشري تمنعان الانكماش.
- ✓ تحليل حوادث متعددة الأطراف لرسم انتقال الحساب عبر المورد والمطور والمنفذ والمستخدم والمنتدى والمتضرر.
- ✓ مقارنة قطاعات عالية التنظيم—الصحة والمال والموارد البشرية—بمهام منخفضة المخاطر لاختبار الشروط الحدودية.
- ✓ فحص قدرة المنتديات نفسها: هل يملك مجلس الإدارة أو المدقق أو المنظم معرفة وحقوق وصول وجزاء تكفي لمساءلة منظمة تستخدم وكلاء؟
- ✓ دراسة كيفية تغير النظرية إذا مُنحت بعض الأنظمة وضعًا قانونيًا أو ترتيبات تأمين إلزامي، وهل ينتج ذلك نقلًا حقيقيًا للحساب أم غسلًا له.

17. القيود

أولاً، المجال سريع التغير، وبعض أهم مصادر 2025-2026 preprints أو نسخ مقبولة لم تستقر حجيتها بعد. عوملت هذه المصادر كإشارات حديثة وفحص تداخل، لا كأساس وحيد. ثانيًا، لم يُبحث مباشرة في Scopus أو Web of Science، ومن ثم لا تدعي المراجعة حصرًا ببليومتريًا. ثالثًا، الورقة نظرية ولا تقدم اختبارًا تجريبيًا لمقاييسها أو قضاياها. رابعًا، يركز الإطار على السياق التنظيمي المعاصر حيث لا يكون الوكيل شخصًا قانونيًا كاملاً. خامسًا، قد تتعارض متطلبات قابلية التدقيق مع الخصوصية والأمن والملكية الفكرية، وقد تكون قابلية العكس مكلفة أو جزئية.

سادسًا، مفهوم «السلطة البشرية المتبقية» له جذور واضحة في نظريات السلطة والملكية والسيطرة البشرية، ومصطلحات قريبة مستخدمة في مجالات أخرى؛ لذا لا يُدعى أنه اختراع من فراغ. القيمة النظرية المقترحة هي تعريفه في سياق التفويض الوكيل وربه تكوينيًا بالمسارات الثلاثة للمساءلة. وأخيرًا، المخطوط عربي؛ يلزم تحويله إلى إنجليزية أكاديمية وإعادة ضبط الطول والبنية قبل الإرسال إلى المجلة المختارة.

18. الخاتمة

تستطيع المنظمة أن تمنح وكيلًا اصطناعيًا قدرة متزايدة على الفعل، لكنها لا تستطيع بمجرد تسمية النظام «وكيلًا» أن تنقل إليه واجب الجواب الأخلاقي أو التنظيمي أو القانوني. ومع ذلك، لا يكفي القول إن الإنسان أو المنظمة «يظل مسؤولًا». فالمساءلة بلا معرفة أو سلطة تدخل أو قدرة إصلاح قد تكون توزيعًا للوم لا حوكمة.

يقترح إطار قابلية التفويض-المساءلة أن التصميم الصالح يحافظ على تكافؤ نسبي بين الأثر المفوض والقدرة المحتفظ بها على إسناده وتقييده وإصلاحه. وتفسر آلية انكماش السلطة البشرية المتبقية كيف ينهار هذا التكافؤ مع العمق والترابط والسرعة والتبعية. لذلك ليس السؤال الحدي هو «إنسان أم آلة؟»، بل: هل يحتفظ الفاعل الذي سيُسأل لاحقًا بقدرة حقيقية ومتناسبة على أن يعرف ويتدخل ويصلح؟ عندما تكون الإجابة لا، يكون التفويض قد سبق البنية الأخلاقية والتنظيمية التي تجعله مشروعًا.

المراجع

- Adusei Brobbey, A. (2026). Agentic AI accountability in organizations [Working paper]. SSRN. <https://ssrn.com/abstract=6663958>
- Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy*, 105(1), 1–29. <https://doi.org/10.1086/262063>
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>
- Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779. [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*, 45(3), 1433–1450. <https://doi.org/10.25300/MISQ/2021/16274>
- Binns, R. (2018). Algorithmic accountability and public reason. *Philosophy & Technology*, 31, 543–556. <https://doi.org/10.1007/s13347-017-0263-5>
- Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal*, 13(4), 447–468. <https://doi.org/10.1111/j.1468-0386.2007.00378.x>
- Bovens, M. (2010). Two concepts of accountability: Accountability as a virtue and as a mechanism. *West European Politics*, 33(5), 946–967. <https://doi.org/10.1080/01402382.2010.486119>
- Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, 81(5), 825–836. <https://doi.org/10.1111/puar.13293>
- Cech, F. (2021). The agency of the forum: Mechanisms for algorithmic accountability through the lens of agency. *Journal of Responsible Technology*, 7–8, Article 100015. <https://doi.org/10.1016/j.jrt.2021.100015>

Chinnaraju, A. (2026). When AI agents act: Governance, accountability, and strategic risk in autonomous organizations. *International Journal of Research and Scientific Innovation*, 12(12), 547–612. <https://doi.org/10.51244/IJRSI.2025.12120050>

Cobbe, J., Lee, M. S. A., & Singh, J. (2021). Reviewable automated decision-making: A framework for accountable algorithmic systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 598–609). ACM. <https://doi.org/10.1145/3442188.3445921>

Cooper, A. F., Moss, E., Laufer, B., & Nissenbaum, H. (2022). Accountability in an algorithmic society: Relationality, responsibility, and robustness in machine learning. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 864–876). ACM. <https://doi.org/10.1145/3531146.3533150>

Dobrajska, M., Billinger, S., & Karim, S. (2015). Delegation within hierarchies: How information processing and knowledge characteristics influence the allocation of formal and real decision authority. *Organization Science*, 26(3), 687–704. <https://doi.org/10.1287/orsc.2014.0954>

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57–74. <https://doi.org/10.5465/amr.1989.4279003>

Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human–robot interaction. *Engaging Science, Technology, and Society*, 5, 40–60. <https://doi.org/10.17351/ests2019.260>

European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. <http://data.europa.eu/eli/reg/2024/1689/oj>

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, 28(1), 62–70. <https://doi.org/10.1016/j.infoandorg.2018.02.005>

Hillebrand, L., Raisch, S., & Schad, J. (2025). Managing with artificial intelligence: An integrative framework. *Academy of Management Annals*, 19(1), 343–375. <https://doi.org/10.5465/annals.2022.0072>

Horneber, D., & Laumer, S. (2023). Algorithmic accountability. *Business & Information Systems Engineering*, 65(6), 723–730. <https://doi.org/10.1007/s12599-023-00817-8>

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10, 18–26. <https://doi.org/10.1007/s13162-020-00161-0>

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)

Karimova, G.-S., & LeMay, S. A. (2026). When no one is responsible: How leaders can restore accountability in complex organizations. *Organizational Dynamics*, 55(3), Article 101245. <https://doi.org/10.1016/j.orgdyn.2026.101245>

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366–410. <https://doi.org/10.5465/annals.2018.0174>

Krakowski, S. (2025). Human–AI agency in the age of generative AI. *Information and Organization*, 35(1), Article 100560. <https://doi.org/10.1016/j.infoandorg.2025.100560>

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review*, 165(3), 633–705. https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/

Kuss, P. M., & Meske, C. (2025). From entity to relation? Agency in the era of artificial intelligence. *Communications of the Association for Information Systems*, 56, 633–674. <https://doi.org/10.17705/1CAIS.05626>

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126–148. <https://doi.org/10.1287/orsc.2021.1549>

Lee, M. Y. (2024). Enacting decentralized authority: The practices and limits of moving beyond hierarchy. *Administrative Science Quarterly*, 69(3), 791–833. <https://doi.org/10.1177/00018392241257372>

Leonardi, P. M. (2025). Homo agenticus in the age of agentic AI: Agency loops, power displacement, and the circulation of responsibility. *Information and Organization*, 35(3), Article 100582. <https://doi.org/10.1016/j.infoandorg.2025.100582>

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), 175–183. <https://doi.org/10.1007/s10676-004-3422-1>

Mukherjee, A., & Chang, H. H. (2025). Agentic AI: Autonomy, accountability, and the algorithmic society [Working paper]. SSRN. <https://doi.org/10.2139/ssrn.5123621>

Murray, A., Rhymer, J., & Sirmon, D. G. (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review*, 46(3), 552–571. <https://doi.org/10.5465/amr.2019.0186>

National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). <https://doi.org/10.6028/NIST.AI.100-1>

National Institute of Standards and Technology. (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1). <https://doi.org/10.6028/NIST.AI.600-1>

Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society*, 39, 1871–1882. <https://doi.org/10.1007/s00146-023-01635-y>

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(3), 286–297. <https://doi.org/10.1109/3468.844354>

Post, C., Sarala, R., Gatrell, C., & Prescott, J. E. (2020). Advancing theory with review articles. *Journal of Management Studies*, 57(2), 351–376. <https://doi.org/10.1111/joms.12549>

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210. <https://doi.org/10.5465/amr.2018.0072>

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44). ACM. <https://doi.org/10.1145/3351095.3372873>

Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5, Article 15. <https://doi.org/10.3389/frobt.2018.00015>

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2026). AI agents vs. agentic AI: A conceptual taxonomy, applications and challenges. *Information Fusion*, 126, Article 103599. <https://doi.org/10.1016/j.inffus.2025.103599>

Sengul, M., Gimeno, J., & Dial, J. (2012). Strategic delegation: A review, theoretical integration, and research agenda. *Journal of Management*, 38(1), 375–414. <https://doi.org/10.1177/0149206311424317>

Shavit, Y., Agarwal, S., Brundage, M., Adler, S., O’Keefe, C., Campbell, R., Lee, T., Mishkin, P., Eloundou, T., Hickey, A., Slama, K., Ahmad, L., McMillan, P., Beutel, A., Passos, A., & Robinson, D. G. (2023). Practices for governing agentic AI systems. OpenAI. <https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf>

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66–83. <https://doi.org/10.1177/0008125619862257>

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>

South, T., Marro, S., Hardjono, T., Mahari, R., Deslandes Whitney, C., Greenwood, D., Chan, A., & Pentland, A. (2025). Authenticated delegation and authorized AI agents [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2501.09674>

Staufer, L., Feng, K., Wei, K., Bailey, L., Duan, Y., Yang, M., Ozisik, A. P., Casper, S., & Kolt, N. (2026). The 2025 AI Agent Index: Documenting technical and safety features of deployed agentic AI systems [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2602.17753>

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>

van Esch, P. (2026). From agentic AI to AI-orchestrated organizations: Understanding the next surge in artificial intelligence. *Business Horizons*. Advance online publication. <https://doi.org/10.1016/j.bushor.2026.03.003>

Wieringa, M. A. (2020). What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 1–18). ACM. <https://doi.org/10.1145/3351095.3372833>