

### **Leandry Junior Jieutsa**

*Investigador en governança de la IA a les ciutats,  
Càtedra UNESCO de Paisatge Urbà*

## **1. Introducció**

L'equitat i la no-discriminació són valors fonamentals de la IA urbana a les ciutats intel·ligents centrades en les persones. El debat cada vegada major entre investigadors i legisladors palesa la importància creixent d'abordar els biaixos i la discriminació en els sistemes d'IA. L'equitat deriva del judici moral, és a dir, el procés pel qual els individus determinen què és moralment correcte o reprovable (Weinkauff, 2023). Tot i que la IA presenta molts avantatges per a les ciutats, el seu desplegament posa a prova l'objectiu d'aconseguir una ciutat justa, en generar o reforçar la discriminació i les desigualtats. Així doncs, integrar els principis d'equitat i no-discriminació en els cicles de vida de la IA urbana és fonamental per garantir el benestar de les persones i les comunitats a les ciutats intel·ligents. No obstant això, la posada en pràctica d'aquest principi segueix sent complexa i poc concreta. Per aconseguir-ho, les ciutats han d'articular diferents papers en la governança de la IA, com a desenvolupadores de solucions internes, responsables del desplegament de sistemes externs o reguladores. Per això, cal adoptar diferents mecanismes, incloent-hi innovacions sociotecnològiques, creació de normes locals d'equitat en la IA i estàndards d'adquisició pública. A més, cal introduir la legislació urbana per protegir els més vulnerables i garantir que els ciutadans poden exercir els seus drets digitals. Tanmateix, aquestes mesures requereixen recursos, que les ciutats poden mobilitzar fomentant la cooperació i el treball en xarxa.

## **2. Un sistema d'IA equitatiu ha de ser imparcial i utilitzar-se de manera responsable**

L'equitat i la no-discriminació són conceptes complexos i fonamentals en la societat contemporània (Barocas et al., 2023a). L'equitat, que el diccionari Cambridge defineix com la qualitat de tractar les persones per igual de manera justa o raonable, respecta les persones com a individus i com a membres de la societat. Tres elements primaris, articulats en dimensions distributiva i sociorelacional, constitueixen aquest concepte que es refereix als individus o als grups (Barocas et al., 2023b): la igualtat

Les ciutats han d'articular diferents papers en la governança de la IA, com a desenvolupadores de solucions internes, com a responsables del desplegament de sistemes externs o com a reguladores.

d'oportunitats, el dret a la justificació i la igualtat en les relacions (Giovanola i Tiribelli, 2022). Una societat equitativa ha de considerar cada individu o grup d'individus tenint en compte les seves circumstàncies i característiques particulars per garantir un tracte i uns resultats equitatis (Giovanola i Tiribelli, 2022; Lyu et al., 2023). Així doncs, l'equitat abraça el concepte de no-discriminació, que implica que ningú no ha de ser exclòs. Els grups o les persones vulnerables són més susceptibles a patir discriminació.

El sorgiment de tecnologies disruptives com la IA suposa un repte per als aspectes relacionats amb l'equitat. En la discriminació relacionada amb la IA hi intervenen dos factors principals: els biaixos algorítmics i la utilització de sistemes basats en IA (Ferrara, 2023; O'Neil, 2016; Wachter et al., 2021).

El primer factor, el biaix algorítmic, distorsiona les dades d'entrenament originals o l'algoritme d'IA, cosa que genera resultats esbiaixats i potencialment perjudicials (Holdsworth, 2023). Aquests biaixos redueixen la precisió i el potencial de la IA en diferent grau segons l'aplicació. Hi ha dos tipus principals de biaixos en la IA: biaixos d'automatització i biaixos per correlació (Barocas et al., 2023a; González-Sendino et al., 2023). Els biaixos d'automatització consisteixen en la propagació a gran escala a través dels processos del sistema d'IA de biaixos socioculturals que estan profundament enquistats en les dades d'entrenament històriques utilitzades per alimentar el sistema d'IA. S'hi inclouen els biaixos humans, els biaixos de les dades, els biaixos d'aprenentatge i els biaixos de desplegament. El biaix per correlació, per la seva banda, té lloc quan la correlació involuntària de variables protegides (per exemple, el sexe o la raça) permeten inferir biaixos, tot i els esforços per excloure'ls de les dades d'entrenament.

El segon factor és la utilització de sistemes basats en IA. De fet, quan s'utilitzen per elaborar perfils o per al control social, els sistemes vulneren els drets digitals (Calzada, 2021; Cugurullo et al., 2022). Per exemple, en recollir i utilitzar informació personal, les tecnologies de reconeixement facial vulneren la intimitat i les dades personals dels ciutadans (UN-Habitat, 2023). Els drets digitals es consideren drets humans existents que cal protegir en el context de les tecnologies digitals, atès que l'espai físic i el digital cada cop estan més entrellaçats (UN-Habitat, 2020).

L'equitat algorítmica es basa en variables interrelacionades (Weinkauf, 2023). Un sistema de decisions automatitzades es considera que és equitatiu quan no depèn de dades sensibles com el sexe o la religió, no desfavoreix les minories i es fa servir de manera responsable.

### **3. El dilema de la IA: l'equilibri entre les oportunitats i els impactes dels sistemes d'IA a les ciutats**

L'urbanisme històricament ha contribuït a crear i reforçar diferents formes de desigualtat i discriminació urbanes (Fainstein, 2009; Hall, 2014). Les poblacions més afectades en són principalment les minories i les persones més vulnerables, que poden variar en funció del context.

En conseqüència, han sorgit nombrosos conceptes, com «dret a la ciutat» o «ciutat justa» de Henry Lefebvre (Fainstein, 2009; Fincher i Iveson, 2012; Harvey i Potter, 2009; Lefebvre, 1968), que aspiren a fer que les ciutats siguin més equitatives, especialment proporcionant oportunitats i serveis urbans que millorin la qualitat de vida.

El sorgiment de la IA suposa un repte per a la ciutat justa, ja que proporciona oportunitats per aconseguir ciutats més inclusives i alhora crea i reforça diferents formes de desigualtats i discriminació. De fet, els sistemes d'IA tenen la capacitat per filtrar i processar grans volums de dades connectades amb xarxes extenses i a l'entorn urbà. Per això, faciliten que es prenguin decisions complexes de manera autònoma o semiautònoma (Marvin et al., 2022; Sherman, 2023; Yigitcanlar et al., 2021). Les metodologies d'IA explicable poden ajudar els ajuntaments a entendre el càlcul de l'equitat i a millorar-la (Lyu et al., 2023). La implementació d'infraestructures basades en IA ajuda a créixer la participació entre els ciutadans i la municipalitat i optimitza la prestació de serveis, especialment per a les poblacions més vulnerables.

Per exemple, les eines d'aprenentatge profund milloren la gestió de les dades espacials per optimitzar la prestació de serveis en barris desafavorits de [Durban \(Sud-àfrica\)](#). La [IA generativa](#) facilita els processos de planificació participatius generant escenaris urbans en temps real, cosa que permet un urbanisme més inclusiu que incorpora perspectives diverses. A més, els xats de conversa municipals, com els que s'han implementat a [Hèlsinki \(Finlàndia\)](#) o a [Saint-Lin-Laurentides \(Canadà\)](#), automatitzen la interacció dels ciutadans amb els ajuntaments, cosa que millora la gestió dels serveis públics, especialment per als individus que no estan familiaritzats amb els tràmits administratius complexos o per a les persones que tenen dificultats per accedir als serveis en persona.

Tanmateix, com hem assenyalat anteriorment, els sistemes d'IA i l'accent en la competitivitat econòmica de les ciutats suposen un repte per a la ciutat justa en produir resultats discriminatoris i poc equitatius. A més, a diferència de les formes tradicionals de discriminació, la discriminació automatitzada generada pels algoritmes resulta més abstracta o opaca i poc intuïtiva, subtil, intangible, difícil de detectar i a gran escala (Kleinberg et al., 2018; O'Neil, 2016; Sanchez et al., 2024; Wachter et al., 2021).

Per exemple, els algoritmes fiscals que singularitzaven «noms que sonaven estrangers» i persones amb «doble nacionalitat» van provocar que [milers de famílies racialitzades fossin acusades falsament de frau](#) als Països Baixos. Globalment, els sistemes de vigilància predictiva, com [Clearview AI](#), suposen problemes de privacitat i alhora reforcen els biaixos (Dauvergne, 2022; O'Neil, 2016). El 2021, la revista *Forbes* va informar de [biaixos algorítmics en les sol·licituds d'hipoteques](#), de manera que al 80 % dels sol·licitants negres se'ls va denegar. De la mateixa manera, *The Markup* (2021) va descobrir que els sol·licitants de color tenien un [40-80 % més de probabilitats que els deneguessin els préstecs](#), la qual cosa emfasitza l'impacte discriminatori de la IA.

A més, la concentració de la riquesa a les grans ciutats, degut a la IA urbana, comporta gentrificació urbana (Sanchez et al., 2024). Especialment, la població amb rendes baixes té cada vegada més difícil, si no impossible,

A diferència de les formes tradicionals de discriminació, la discriminació automatitzada generada pels algoritmes resulta més abstracta o opaca i poc intuïtiva, subtil, intangible, difícil de detectar i a gran escala.

Degut a la seva opacitat i magnitud, és difícil per a les persones afectades fer valer els seus drets [...] perquè els sistemes de presa de decisions algorítmics alteren els recursos i els procediments jurídics tradicionals que detecten, investiguen, prevenen i corregeixen la discriminació.

accedir a un habitatge. Així doncs, el desplegament de polítiques d'IA urbana contribueix a reforçar les asimetries entre territoris i les desigualtats urbanes.

Per tant, els sistemes d'IA tenen efectes importants sobre les ciutats i les societats. Aquesta ambivalència planteja la necessitat d'una governança efectiva. A més, degut a la seva opacitat i a la magnitud del seu impacte, és difícil per a les persones afectades defensar-se o fer valer els seus drets. Això posa entredit el dret a la no-discriminació del qual gaudeixen els ciutadans perquè els sistemes de presa de decisions algorítmics alteren els recursos i els procediments jurídics tradicionals que detecten, investiguen, prevenen i corregeixen la discriminació (Wachter et al., 2021).

## 4. Recomanacions polítiques

Segons les recomanacions de la UNESCO, els agents implicats en la IA han d'adoptar un enfocament inclusiu per tal que els beneficis de les tecnologies de la IA arribin a tothom, tenint en compte les necessitats concretes dels diferents grups (UNESCO, 2023). Des del punt de vista de les ciutats, la implementació de l'equitat i la no-discriminació en els sistemes d'IA urbana requereix l'articulació dels diversos papers que els ajuntaments assumeixen com a desenvolupadors de solucions internes (tot i que relativament infreqüents degut a limitacions financeres i tècniques), responsables del desplegament i reguladors. A més, millorar l'equitat dels sistemes d'IA implica tenir en compte tot el seu cicle de vida, abordant els diversos aspectes al llarg dels processos de disseny, desenvolupament i implementació. Així mateix, oferir solucions eficaces per solucionar les desigualtats en els resultats del sistema d'IA passa per identificar-ne les causes subjacents.

### 4.1. Recomanacions generals

- **Definir una estratègia:** les ciutats han d'implementar estratègies d'IA que girin al voltant dels principis d'equitat i no-discriminació. Aquests documents estratègics permeten a les ciutats establir uns fonaments robustos i dur a terme una avaluació precisa dels seus objectius relacionats amb la IA. Aquest enfocament és fonamental per planificar la integració de la IA amb l'objectiu de maximitzar els beneficis i alhora mitigar els riscos potencials. Aquestes estratègies s'han de desenvolupar a través d'un procés participatiu juntament amb plans d'acció que defineixin mesures concretes per garantir la integració equitativa de la IA que no deixi enrere ningú.
- **Establir nivells de risc segons les aplicacions:** les ciutats han d'identificar les aplicacions d'IA de risc alt dins el marc de les seves competències, tenint en compte les disparitats i les desigualtats existents al territori. Un cop identificades aquestes aplicacions de risc alt cal implementar mecanismes de protecció. Les aplicacions relacionades amb serveis socials essencials s'han de considerar de risc alt i se n'ha de prohibir el funcionament de manera totalment autònoma. Per exemple, la [ciutat californiana de San José](#) ha posat en marxa un registre d'IA estructurat al voltant d'una avaluació rigorosa dels sistemes d'IA. Aquest procés inclou l'anàlisi de riscos i una posterior avaluació més exhaustiva de l'impacte, en funció del nivell de risc, tot documentat en

un «Full d'impacte» i una «Fulla informativa d'IA».

#### 4.2. Recomanacions concretes per a les ciutats com a desenvolupadores de solucions internes

- **Fomentar la innovació socioecològica inclusiva.** Cal comptar amb diverses parts interessades de perfil no tècnic al llarg del cicle de vida de la IA. Segons ONU-Habitat, aquest cicle de vida de la IA consta de cinc fases: contextualització, disseny, implementació, desplegament i manteniment. Si les decisions en aquestes etapes les prenen majoritàriament tècnics o grups homogenis, hi ha un risc important que els seus biaixos s'integrin en el sistema d'IA. Aquest risc és especialment pronunciat si l'eina s'aplica posteriorment de manera generalitzada a sectors més amplis de població. Els governs locals han de posar més èmfasi en la interdisciplinarietat i la multidisciplinarietat que integren els grups socials en el cicle de vida de la IA urbana.
- **Aplicar tècniques d'equitat,** com ara el pretractament de dades (que implica identificar i abordar biaixos en les dades abans de l'entrenament del model), la selecció del model (que utilitza mètodes de selecció de models que prioritzen l'equitat) i les decisions de postprocessament (que consisteix en ajustar el resultat dels models de IA per mitigar els biaixos i garantir l'equitat) (Ferrara, 2023).
- **Fomentar la diversitat a l'hora de construir bases de dades en tres aspectes: equips, dades i models.** Crear equips diversos i interdisciplinaris i proporcionar formació contínua en equitat i ètica és fonamental per reduir els biaixos. Pel que fa a les dades, millorar la recollida d'atributs sensibles (per exemple, sexe, raça, ètnia) i documentar les decisions relacionades amb les dades fomenta la transparència i ajuda a fer front a les desigualtats del món real. Per als models, proporcionar accés obert a la comunitat per al testatge, garantir la documentació transparent i utilitzar IA explicable pot ajudar a identificar i reduir els biaixos, garantint uns resultats equitatius (González-Sendino et al., 2023).
- **Integrar la correlació compensatòria en els sistemes d'IA.** Tal com indiquen Giovanola i Tiribelli (2022), garantir la igualtat d'oportunitats en els sistemes d'IA no es pot limitar a eliminar els biaixos discriminatoris en les dades d'entrenament. Cal dissenyar els sistemes d'IA urbana per considerar les desigualtats existents en el seu context i incorporar-hi mecanismes per compensar-les. Per exemple, en una ciutat amb diferències entre comunitats o grups socials, les IA han de tenir en compte aquestes disparitats i aplicar-hi mesures compensatòries, cosa que es pot materialitzar en contingut personalitzat, per exemple.
- **Integrar tècniques de mitigació en el cicle de vida de la IA.** Neutralitzar els efectes discriminatoris de les dades durant la fase de preentrenament a través de mètodes com el remostreig (és a dir, modificar la mida del conjunt de dades que afecta la distribució sense transformar les dades), la representació justa (que s'aconsegueix eliminant informació que permeti associar un individu amb un grup protegit) i la reponderació (utilitzada per transformar les dades modificant el seu pes dins el conjunt de dades). Durant la

fase d'entrenament, cal utilitzar la regularització i l'entrenament adversari, que són els mètodes més comuns per a aquest objectiu. Altres enfocaments emergents són l'aprenentatge descentralitzat, la regressió lineal justa, el DeepFair, els models multimodals i la clusterització de fairlets. Durant la fase de postentrenament, cal implementar probabilitats igualades, probabilitats igualades calibrades i rebutjar la classificació d'opcions.

#### **4.3. Recomanacions concretes per a les ciutats com a responsables de desplegament i reguladores**

- **Determinar normes locals per a la IA equitativa.** La discriminació i les desigualtats es poden manifestar de maneres diferents segons el context, afectant individus o grups socials de maneres diferents i a diferents escales. Per tant, les ciutats han d'aplicar normes d'equitat per a la IA urbana que tinguin en compte aquestes particularitats locals. Aquestes normes han d'incloure principis generals i alhora qüestions d'abast local amb l'objectiu de garantir que la IA urbana no reforça la discriminació ja existent ni crea nous biaixos que afectin de manera desproporcionada els més vulnerables. Cal consultar les comunitats locals a l'hora de desenvolupar aquestes normes i tenir en compte tot el cicle de vida de la IA.
- **Establir normes d'adquisició pública per a una IA equitativa.** Les ciutats han de garantir que les entitats que els proporcionen serveis compleixin els principis de la IA equitativa, cosa que requereix establir mecanismes d'adquisició pública que obliguin els proveïdors de serveis a complir les normes d'IA equitativa de la ciutat. Els proveïdors han de complir els requisits de conformitat relatius als seus algorismes si els ha d'usar la ciutat. Per exemple, la [Coalició GOV AI](#), encapçalada per la ciutat de San José, ha adoptat i introduït l'esmentada «Fulla informativa d'IA» per a sistemes de tercers, una plantilla unificada perquè els proveïdors facilitin informació detallada sobre els productes d'IA, que inclou, entre altres aspectes, la finalitat del sistema, les dades d'entrenament, els detalls del model, els indicadors per mesurar el rendiment, la gestió de biaixos, la robustesa i la interacció persona-ordinador.
- **Implementar una legislació urbana que garanteixin el dret a la justificació.** Aquest dret permet a les persones afectades per un sistema d'IA entendre el raonament subjacent a una decisió algorítmica, cosa que possibilita als ciutadans entendre i controlar com els tracten aquests sistemes. Quan aquest dret no es respecta de manera adequada, s'ha de tenir la capacitat per qüestionar i modificar els paràmetres subjacents de la decisió. Per tant, les ciutats han de considerar al llarg del procés si despleguen o retiren un sistema d'IA, especialment si no es pot atendre la petició d'explicacions per part d'algú. Aquesta mesura permet a les persones que pateixen discriminació fer valer els seus drets digitals.
- **Crear òrgans assessors per investigar, prevenir i mitigar els potencials usos maliciosos de la IA.** Els governs municipals han de crear òrgans assessors multidisciplinars que incloguin organitzacions comunitàries, el món acadèmic, empreses i altres parts interessades de la comunitat. Aquests organismes tindran un paper de control per limitar la discriminació relacionada amb la IA. Avaluaran els models d'IA de la

ciutat basant-se en indicadors que mesurin l'equitat. Aquesta avaluació (1) identificarà els potencials biaixos que poden afectar l'equitat, (2) seleccionarà indicadors per mesurar l'equitat dels sistemes, i (3) mitigarà l'impacte generat per aquests biaixos. A més, actuaran com a organismes assessors per orientar les ciutats a l'hora d'emprendre accions i dur a terme polítiques relacionades amb la IA equitativa.

## 5. Limitacions

Assolir l'equitat en la IA és un procés complex, ja que les intervencions que tenen aquesta meta en contextos d'IA urbana poden entrar en conflicte amb els objectius mateixos dels algoritmes. Per aquest motiu, les ciutats han d'adoptar solucions de compromís per arribar a un equilibri entre els guanys i els beneficis i alhora prioritzar el benestar de les persones i de les comunitats. Tanmateix, aquest principi pot semblar abstracte i pot donar lloc a interpretacions divergents, cosa que complica l'aplicació de mesures d'èxit i impacte (Sadek et al., 2024). Per això, les ciutats han d'adoptar un enfocament local per posar en pràctica l'equitat i la no-discriminació en la IA urbana. Aquest enfocament integral considera la configuració socioeconòmica i cultural de la ciutat al llarg de tot el cicle de vida de la IA.

Des del punt de vista tècnic, la IA urbana equitativa requereix recursos humans diversos i infraestructures adaptades (Du et al., 2023; Marvin et al., 2022; Yigitcanlar et al., 2020, 2023), per a la qual cosa es necessiten considerable inversions econòmiques (Bettoni et al., 2021). També es generen més despeses per la capacitat i la formació del personal i les comunitats (Sadek et al., 2024; Varanasi, 2023). Les ciutats també han de preveure els costos jurídics i de conformitat, com ara auditories i ajustos del sistema per complir les normes reglamentaries, inversions que poden suposar despeses importants, especialment per a les ciutats petites i mitjanes.

Per superar aquestes limitacions, les ciutats poden recórrer al treball en xarxa a través de xarxes que proporcionen oportunitats per a l'intercanvi de coneixement, la innovació política i les respostes coordinades a problemes globals. En són exemples:

**Coalició de Ciutats per als Drets Humans Digitals:** plataforma per fomentar un desenvolupament inclusiu i democràtic de les noves tecnologies a les ciutats.

**City AI Connect:** comunitat d'aprenentatge global i plataforma digital perquè les ciutats provin i promoguin l'ús de IA generativa per millorar els serveis públics.

**GovAI:** coalició formada per més de 1.000 membres i més de 350 entitats locals, estatals i federals unides amb l'objectiu de promoure una IA responsable i amb un propòsit clar en el sector públic.

**AI4Cities:** projecte que va permetre a Hèlsinki, Amsterdam, Copenhagen, el Gran París, Stavanger i Tallinn desafiar el mercat a trobar solucions d'IA per reduir les emissions de CO<sub>2</sub> en els camps de l'energia i la mobilitat.



## Referències bibliogràfiques

Barocas, S., Hardt, M., i Narayanan, A. «Classification». Dins *Fairness and Machine Learning : Limitations and Opportunities*. MIT Press, 2023a

Barocas, S., Hardt, M., i Narayanan, A. «Relative notions of fairness». Dins *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press, 2023b

Bettoni, A. et al. «An AI adoption model for SMEs: A conceptual framework». *IFAC-PapersOnLine*, 2021, vol. 54, n.º 1, p. 702-708.

Calzada, I. «The right to have digital rights in smart cities». *Sustainability (Switzerland)*, 2021, vol. 13, n.º 20.

Cugurullo, F. et al. «Urban AI in China: Social control or hyper-capitalist development in the post-smart city? ». *Frontiers in Sustainable Cities*, 2022

Dauvergne, P. «Facial recognition technology for policing and surveillance in the Global South: A call for bans». Dins *Third World Quarterly*. Routledge, 2022, vol. 43, n.º 9, p. 2325-2335.

Du, J. et al. «Artificial intelligence enabled participatory planning: A review». *International Journal of Urban Sciences*, 2023

Fainstein, S. «Planning and the Just City». Dins P. Marcuse (Ed.), *Searching for the just city: Debates in urban theory and practice*. Routledge, 2009

Ferrara, E. «Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies». *Sci*, 2023, vol 6, n.º 1, p. 3.

Fincher, R., i Iveson, K. «Justice and injustice in the city». *Geographical research*, 2012, vol 50, n.º.3, p. 231-241.

Giovanola, B., i Tiribelli, S. «Weapons of moral construction? On the value of fairness in algorithmic decision-making». *Ethics and Information Technology*, 2022, vol. 24, n.º 1, p. 3.

González-Sendino, R. et al. «A review of Bias and Fairness in Artificial Intelligence». *International Journal of Interactive Multimedia and Artificial Intelligence* (en premsa), 2023

Hall, P. «Cities of tomorrow: An intellectual history of urban planning and design since 1880 (4ª edició)». *Wiley-Blackwell*, 2014

Harvey, D., i Potter, C. «The right to the Just City». Dins P. Marcuse (Ed.), *Searching for the just city: Debates in urban theory and practice*. Routledge, 2009

Holdsworth, J. «What is AI bias?», 2023

Kleinberg, J. et al. «Discrimination in the Age of Algorithms». *Journal of Legal Analysis*, 2018, vol.10, 2005, p. 113-174.



- Lefebvre, H. «[The Right to the City](#)», 1968
- Lyu, Y. *et al.* «IF-City: Intelligible fair city planning to measure, explain and mitigate inequality». *IEEE Transactions on Visualization and Computer Graphics*, 2023
- Martinez, E., i Kirchner, L. «[The secret bias hidden in Mortgage-Approval algorithms](#)», 2021
- Marvin, S. *et al.* «Urban AI in China: Social control or hyper-capitalist development in the post-smart city? ». *Frontiers in Sustainable Cities*, 2022, vol. 4, p. 1030318.
- O'Neil, C. «Weapons of math destruction: How big data increases inequality and threatens democracy (First edition) ». *Crown*, 2016
- Rezende, I. N. «Facial recognition in police hands: Assessing the 'Clearview case' from a European perspective». *New Journal of European Criminal Law*, 2020, vol. 11, n.º 3, p. 375-389.
- Sadek, M. *et al.* «Challenges of responsible AI in practice: Scoping review and recommended actions». *AI & SOCIETY*, 2024
- Sanchez, T. W., Brenman, M., i Ye, X. «The Ethical Concerns of Artificial Intelligence in Urban Planning». *Journal of the American Planning Association*, 2024, vol. 0, n.º 0.
- Sherman, S. «The Polyopticon: A diagram for urban artificial intelligences». *AI and Society*, 2023, vol. 38, n.º 3, p. 1209-1222.
- UNESCO. «Readiness assessment methodology: A tool of the recommendation on the ethics of artificial intelligence». *UNESCO*, 2023
- UN-Habitat. «[Mainstreaming human rights in the digital transformation of cities-A guide for local governments](#)». *United Nations Human Settlements Programme*, 2020.
- UN-Habitat. «[Human rights in the digital era](#)». *United Nations Human Settlements Programme*, 2023, p. 1-56.
- Varanasi, R. A. «“It is currently hodgepodge”: Examining AI/ML Practitioners' Challenges during Co-production of Responsible AI Values», 2023
- Wachter, S., Mittelstadt, B., i Russell, C. «Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI». *Computer Law & Security Review*, 2021, vol. 41, p. 105567.
- Weinkauf, D. «[Privacy Tech-Know Blog: When Worlds Collide – The Possibilities and Limits of Algorithmic Fairness \(Part 1\)](#)». *Office of the Privacy Commissioner of Canada*, 2023.
- Yigitcanlar, T. *et al.* «Artificial intelligence in local governments: Perceptions of city managers on prospects, constraints and choices». *AI and Society*, 2023, vol. 38, n.º 3, p. 1135-1150.

Yigitcanlar, T. *et al.* «Responsible urban innovation with local government artificial intelligence (Ai): A conceptual framework and research agenda». *Journal of Open Innovation: Technology, Market, and Complexity*, 2021, vol. 7, n.º 1, p. 1-16.

Yigitcanlar, T. *et al.* «Contributions and Risks of Artificial Intelligence (AI) in Building Smarter Cities: Insights from a Systematic Review of the Literature», 2020