

CONTEXTO Y ESTADO DEL ARTE



INTRODUCCIÓN

Este texto es un extracto del TFM de Carlos Rossique Delmas.
Puede usarse citando al autor, según la licencia CC-BY-NC-SA

La toma de decisiones colectivas ha sido una preocupación central en disciplinas tan diversas como la filosofía política, la psicología social, la economía conductual y la teoría de sistemas. Con la expansión de internet y el desarrollo de tecnologías colaborativas, han surgido múltiples plataformas y metodologías que buscan mejorar la deliberación y la agregación de preferencias en grupos. En paralelo, la inteligencia artificial (IA), especialmente en su vertiente basada en modelos de lenguaje y razonamiento automatizado, está transformando radicalmente los procesos de análisis y toma de decisiones.

El presente capítulo se centra en dar noticia de los fundamentos teóricos y contextuales alrededor de este tema y en revisar el estado del arte de tecnologías y aplicaciones relacionadas, identificando vacíos y oportunidades que justifican la novedad del presente trabajo.

Se detalla también en qué algoritmos y técnicas metodológicas se basa el trabajo; como el método Delphi, el Frente de Pareto o la introducción de un nuevo álgebra donde se toman las propuestas como operandos sobre los que se definen diversas operaciones.

CONTEXTO DEL PROBLEMA: DELIBERACIÓN Y DECISIONES COLECTIVAS

La toma de decisiones colectivas implica coordinar las opiniones, intereses y conocimientos de múltiples individuos con el fin de alcanzar resultados informados, legítimos y eficaces. Desde una perspectiva epistemológica, la inteligencia colectiva parte de la premisa de que, bajo ciertas condiciones, los grupos pueden superar en capacidad cognitiva a los individuos que los componen (Surowiecki, 2004).

Sin embargo, las formas tradicionales de decisión y deliberación colectivas presentan importantes limitaciones, como la presencia de sesgos cognitivos, la falta de estructura argumentativa, la polarización ideológica o la dificultad de escalar la participación cuando los grupos son numerosos (Sunstein, 2002; Landemore, 2012).

Históricamente, la teoría del voto ha constituido el marco de referencia para la agregación de preferencias individuales. Un voto expresa la elección de una persona entre un conjunto limitado de opciones, y los mecanismos de votación se han diseñado tradicionalmente para

producir una única decisión colectiva. En los modelos más simples, como los referéndums binarios o las elecciones uninominales a una vuelta, se adopta la opción que supera el umbral de mayoría. Cuando existen más de dos alternativas, pueden emplearse sistemas mayoritarios simples (escoger una entre varias), o fórmulas proporcionales para distribuir representación según los votos recibidos (como los métodos de D'Hondt, Hare o Saint-Laguë).

Sin embargo, estos mecanismos presentan múltiples deficiencias cuando se desea capturar la riqueza de las preferencias individuales en contextos complejos. A menudo, los sistemas mayoritarios tienden a excluir opciones minoritarias, y los proporcionales, aunque más inclusivos, pueden distorsionar la voluntad colectiva según el diseño del sistema. Por ello, se han propuesto métodos alternativos de votación que permiten una expresión más matizada, como el voto aprobatorio (approval vote), en el que cada persona selecciona todas las opciones aceptables sin ordenarlas, o el voto preferencial, que permite ordenar las alternativas según el grado de preferencia, utilizando algoritmos como Borda, Dowdall o el voto único transferible (VUT) para determinar el resultado.

No obstante, el célebre teorema de imposibilidad de Arrow (1952) demostró que ningún sistema de votación puede cumplir simultáneamente ciertos principios básicos de equidad (criterios de mayoría, de Condorcet, monotonidad e independencia de alternativas irrelevantes) cuando hay tres o más opciones. Esta imposibilidad teórica sugiere que toda técnica de agregación conlleva compromisos entre justicia, simplicidad y estabilidad.

Frente a estas limitaciones, se han desarrollado enfoques que permiten una expresión cuantitativa o multidimensional de las preferencias, como los sistemas de votación ponderada. En estos, los participantes pueden distribuir puntos entre opciones, calificarlas en escalas numéricas (e.g., 0 a 10), o evaluar múltiples atributos de una misma propuesta (por ejemplo, su calidad, factibilidad o equidad). Este tipo de mecanismos es ampliamente utilizado en contextos digitales (valoraciones en redes sociales, rankings en plataformas audiovisuales), y se beneficia de técnicas estadísticas que permiten corregir sesgos, identificar polarización o estimar valoraciones agregadas más robustas (e.g., mediante medias bayesianas ponderadas).



Figura 1: Jerarquía de decisión colectiva

Sin embargo, aún con esas mejoras la elección se propone como algo unidireccional y en gran parte pasivo, es decir, entre opciones ya dadas y usualmente muy limitado en el tiempo.

En esta línea, Speroni (2011) propone una crítica estructural a la democracia representativa tradicional: en un mundo de comunicaciones lentas y estructuras jerárquicas, el modelo de voto es adecuado, pero en un entorno global, digital y en red, donde la comunicación es rápida y multilateral ("muchos a muchos"), se abre la posibilidad de incorporar preguntas abiertas y propuestas generadas por los propios participantes. La capacidad expresiva del votante en estos nuevos escenarios es mucho mayor, y su "tasa de información" —medida en bits por segundo— supera con creces la de los sistemas de elección cerrada tradicionales, que resultan ineficientes en términos informativos si se comparan con interacciones digitales cotidianas como, por ejemplo, la declaración de impuestos o el consumo de contenidos en línea.

Este cambio de paradigma abre la puerta a nuevas formas de participación digital como la democracia líquida, donde el voto no solo se ejerce de forma puntual, sino que puede modificarse o delegarse de manera dinámica. Este enfoque permite combinar representatividad y participación directa, ajustando el grado de implicación ciudadana a las circunstancias y conocimientos de cada individuo. Del mismo modo, se cuestiona la máxima de "una persona, un voto" cuando el grado de afectación, el conocimiento o el mérito podrían justificar ponderaciones diferenciadas.

En paralelo, surgen plataformas que permiten a los participantes proponer directamente soluciones o respuestas abiertas, más allá de simplemente elegir entre opciones predeterminadas. Estas propuestas pueden someterse a ciclos de evaluación, mejora, fusión o reformulación, generando dinámicas deliberativas estructuradas que se asemejan a procesos evolutivos o incluso a algoritmos genéticos, donde las propuestas "mutan" o se recombinan en busca de soluciones más consensuadas.

El concepto de voluntad general, introducido por Rousseau en el siglo XVIII, y hasta ahora difícil de operacionalizar, podría hoy encontrar un tratamiento computacional gracias a las herramientas de análisis de datos, razonamiento automático y aprendizaje automático. Bajo esta visión, el consenso no es simplemente la mayoría, sino un óptimo social que maximiza simultáneamente múltiples criterios de valor colectivo.

Speroni (2011) propone para ello el uso del Frente de Pareto como marco para identificar soluciones dominantes en espacios de decisión multidimensionales. Hay que advertir que el consenso no significa unanimidad, es decir, que todos piensen igual sobre determinado asunto, ni gregarismo, ni uniformidad, ni autocensura. Más que en un resultado concreto, se centra en un enfoque y un proceso (figura 2). En realidad, el objetivo de dicho proceso va más allá, hacia cierta unidad emotiva. Tampoco significa un proceso de negociación donde todos tienen que ceder para quedar en un descontento general y equidistante, ni una mayoría cualificada, por amplia que sea. La base del consenso es un acto conjunto de cooperación, no una lucha de intereses o puntos de vista particulares. Al contrario, es deliberación, aporte, cooperación, pensamiento crítico y requiere de cierta flexibilidad y empatía para ver puntos de vista propios y ajenos, que no significa que uno deba acatar o ceder sin comprender.

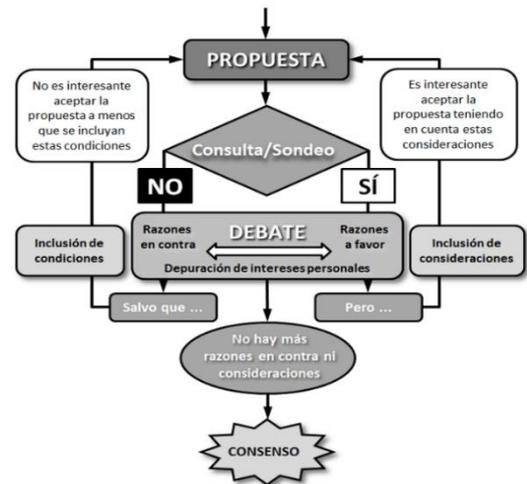


Figura 2. Esquema básico de un posible ciclo orientado al consenso

En suma, la evolución de los mecanismos de decisión colectiva, desde el voto binario hasta los modelos deliberativos multietapa (figura 1), encuentra hoy un terreno fértil en las tecnologías digitales. Estas permiten no solo recoger preferencias de forma más rica, sino también estructurar el debate, facilitar la deliberación constructiva y modelar soluciones algorítmicamente tratables que reflejen mejor el interés colectivo. Esta perspectiva fundamenta el desarrollo del sistema propuesto en este trabajo, que busca asistir, mediante inteligencia artificial, los procesos complejos de toma de decisiones colectivas.

HERRAMIENTAS TIC DE FACILITACIÓN DE INTELIGENCIA COLECTIVA

Diversas plataformas han sido desarrolladas para canalizar la deliberación y la decisión colectivas. No se pretende aquí un estudio exhaustivo y tampoco se incluyen aplicaciones que desarrollan solo mejoras relacionadas con la teoría del voto o encuestas de respuestas a opciones cerradas, sino que se incluyen herramientas que incluyen respuestas a preguntas abiertas y procesos deliberativos. A continuación se analizan las más representativas, prestando atención a su diseño, proceso deliberativo, alcance y aplicación o no de técnicas IA.

2.3.1 ThoughtExchange



ThoughtExchange es una plataforma de participación colectiva fundada en 2009, centrada en captar el pensamiento colectivo mediante dos funcionalidades principales: los intercambios (Exchanges) y las encuestas (Surveys). En los intercambios, los participantes responden a una pregunta abierta y se califica anónimamente y de forma aleatoria las ideas aportadas por otros. Las encuestas pueden combinar varios tipos de preguntas como elecciones entre una o varias opciones, escalas cuantitativas, preferenciales e incluso respuestas de texto abiertas.

ThoughtExchange se ha utilizado sobre todo en el entorno educativo y menos en negocios y gobierno, para facilitar la toma de decisiones colaborativas. Aunque inicialmente no incorporaba inteligencia artificial, desde 2023 ThoughtExchange ha integrado capacidades de IA generativa y análisis automatizado, incluyendo un asistente conversacional (integrado con OpenAI) que identifica temas emergentes y proporciona visiones estructuradas sobre los aportes recogidos, incluyendo puntos en común y en desacuerdo, y que pueden verse en sus paneles de resultados como gráficos, resúmenes, etc.

2.3.2 LiquidFeedback



LiquidFeedback es una plataforma de código abierto, de deliberación y toma de decisiones colectivas desarrollada en 2009 por el Public Software Group e.V. en Alemania.

Su funcionalidades permiten a los participantes proponer ideas, fase de discusión con posibilidad de enmiendas, valoración de estas, que el iniciador puede incluir o no, debatirlas públicamente y finalmente votar incluyendo la posibilidad de voto preferencial y también de delegación (líquidez) del voto en otra persona, por tema. Este sistema de delegación flexible permite que las decisiones reflejen de manera más precisa las preferencias colectivas, adaptándose a las competencias y conocimientos de los participantes. LiquidFeedback ha sido utilizada por diversos partidos políticos, organizaciones no gubernamentales y comunidades en línea para promover procesos de toma de decisiones más inclusivos y transparentes.

A diferencia de otras más recientes, LiquidFeedback no incorpora técnicas de IA, apoyándose en mecanismos más formales, estructurados y transparentes como base de su legitimidad.

2.3.4 Mesydel



Mesydel (Multi-round E-Survey for DELphi) es una plataforma de consulta participativa en línea desarrollada en 2008 por la Universidad de Lieja (Bélgica). Basada en el método Delphi, estructura procesos de consulta en múltiples rondas anónimas. Los participantes (usualmente paneles de expertos) responden a preguntas abiertas y cerradas; luego, sus respuestas se analizan cualitativa y cuantitativamente y se realimentan al grupo para nuevas iteraciones.

Se ha usado en contextos como políticas públicas, evaluaciones participativas, investigación científica y construcción de consenso interdisciplinar. Actualmente, no incorpora técnicas de inteligencia artificial, aunque automatiza parcialmente las tareas logísticas y analíticas del proceso Delphi.

2.3.3 Loomio



Loomio es una plataforma de código abierto (AGPLV3) de toma de decisiones colaborativa desarrollada en 2011 en Nueva Zelanda por miembros del movimiento Occupy y Enspiral. Su objetivo es facilitar procesos deliberativos horizontales en organizaciones, colectivos y comunidades. (Boyte, 2014).

El funcionamiento de Loomio se basa en tres elementos clave: grupos, hilos de discusión y propuestas. Los participantes pueden iniciar debates, proponer acciones y votar utilizando opciones como "de acuerdo", "en desacuerdo", "abstención" o "bloqueo". El sistema permite modificar el voto durante la deliberación y visualizar gráficamente el consenso alcanzado. (Ver figura 3)

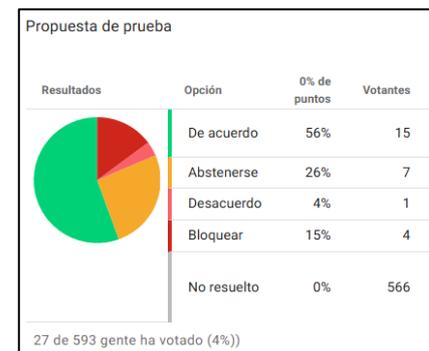


Figura 3. Votación de una propuesta en Loomio (Fuente: Loomio)

Propone varios tipos de decisión (consenso, consejo, consentimiento, votación simple) en los que se definen criterios, p.ej., de inexistencia de bloqueos, porcentaje mínimo de participación o de acuerdo, y tiempo mínimo de vida de la encuesta/votación. La plataforma ha sido utilizada por organizaciones como cooperativas, colectivos y empresas autogestionadas para mejorar sus prácticas de gobernanza y fomentar una cultura de toma de decisiones más inclusiva y transparente. Tiene una escasa escalabilidad ya que se limita. No utiliza técnicas avanzadas de IA, aunque permite cierta integración con APIs externas, Su enfoque se centra en la transparencia y la trazabilidad y la participación estructurada sin automatismos algorítmicos.

2.3.5 Appgree



Appgree fue una plataforma de participación colectiva desarrollada en España y lanzada en 2013. Su objetivo principal era facilitar la expresión y priorización de ideas dentro de grandes grupos. Su funcionamiento se basa en un proceso en varias etapas: primero, los participantes responden a una pregunta abierta formulando sus propias propuestas; posteriormente, estas propuestas son sometidas a una evaluación anónima y aleatoria por parte de otros usuarios. Las mejores posicionadas vuelven a distribuirse para su evaluación por conjuntos aleatorios de participantes y así sucesivamente hasta escoger la respuesta más apoyada.

Aunque su enfoque es ágil y escalable, su capacidad deliberativa es limitada, pues no permite explorar en profundidad las razones detrás de las preferencias (Redondo, 2015). Appgree fue empleada en campañas de marketing, eventos en vivo y procesos participativos a gran escala, buscando facilitar la interacción y las opiniones mayoritarias en grandes grupos. A partir de 2014 incorporó agrupamiento automático con IA, pero hoy está inactiva.

2.3.6 iWarsM'aps / CiBUC



iWarsM'aps (Innovation Wars Mobile App) fue una herramienta desarrollada en España lanzada en 2014, como aplicación de mapeo de la argumentación. Su objetivo era facilitar a empresas e instituciones la detección colectiva de problemas y la generación de soluciones mediante preguntas abiertas, estructurando las aportaciones de los usuarios en torno a afirmaciones, razones, objeciones y contraargumentos, de forma similar al método Delphi.

Su evolución, la plataforma CiBUC, introduce técnicas de proceso de lenguaje natural e IA en la deliberación colectiva. Tras hacer una pregunta abierta a un grupo de participantes, las



Figura 4. Algoritmos de IA en CiBUC (Fuente: cibuc.com)

respuestas anónimas son procesadas mediante varios algoritmos de clusterización e hibridación semántica, para posteriormente pasar estas respuestas sintéticas un proceso de votación. Esta combinación de PLN y heurísticas deliberativas permite detectar patrones emergentes, organizar el pensamiento colectivo y facilitar el consenso. Estas iniciativas fueron un caso pionero en aplicar argumentación estructurada + IA al ámbito público, aunque su impacto quedó limitado a proyectos locales y hoy en día ambas aplicaciones no están activas.

2.3.7 Vilfredo goes to Athens

Vilfredo goes to Athens es un proyecto de eDemocracy orientado a facilitar la construcción de consensos en grupos pequeños, típicamente de 5 y 20 personas, que fue desarrollado por el Dr. Pietro Speroni di Fenizio hacia 2015. Vilfredo emplea explícitamente el concepto de Frente de Pareto en su algoritmo de deliberación colectiva. En cada ronda, las propuestas son evaluadas por los participantes, y aquellas que no son dominadas por otras forman el frente de Pareto (una propuesta A domina a una B si los votantes de B son un subconjunto de los votantes de A). Este conjunto de propuestas representa las opciones más aceptables y que reflejan fielmente la diversidad de opiniones dentro del colectivo.

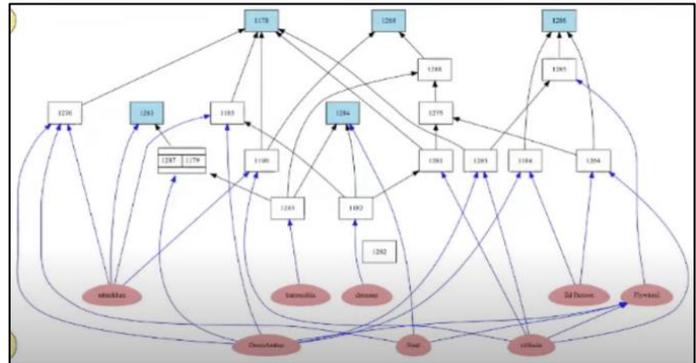


Figura 5. Jerarquía de dominación en Vilfredo (Fuente: Vilfredo goes to Athens)

Además, el sistema realiza un análisis matemático de los votos para sugerir acciones específicas que podrían facilitar el logro de un consenso, como indicar quiénes deberían reconsiderar su voto en una propuesta particular, o entablar diálogo entre sí, o quiénes podrían reformular ciertas propuestas para hacerlas más inclusivas. Estas sugerencias buscan optimizar el proceso deliberativo, orientándolo hacia el consenso aunque los participantes conservan plena autonomía para decidir.

2.3.8 Polis



Pol.is es una plataforma open source de deliberación colectiva lanzada en 2012 por el Computational Democracy Project. Estructura el proceso deliberativo a partir de una pregunta abierta, a la que los participantes envían declaraciones breves y votan sobre las de otros indicando acuerdo, desacuerdo o neutralidad. A través de técnicas IA de aprendizaje automático y análisis estadístico, el sistema agrupa automáticamente a los usuarios en clústeres de opinión, identifica zonas de consenso y disenso. No hay votación final clásica.

Ha sido utilizada en varios países, destacando su implementación en vTaiwan, una iniciativa del gobierno taiwanés para la elaboración colaborativa de políticas públicas. Aunque no analiza semánticamente el contenido (PLN), su enfoque estadístico permite gestionar eficazmente grandes volúmenes de participación. (Tang & Kalih, 2020)

2.3.9 Decidim



Decidim es una plataforma digital de participación ciudadana de código abierto, desarrollada en 2016 por el Ayuntamiento de Barcelona y mantenida por una comunidad global. Permite configurar procesos participativos complejos como presupuestos participativos, consultas, normativas colaborativas y asambleas, mediante fases personalizables que incluyen propuestas, debates, votaciones y seguimiento. Los participantes pueden enviar propuestas, comentarlas y votar con distintos modos (limitado, ponderado, por coste), con trazabilidad y transparencia del proceso y auditoración de resultados. (Fuster Morell, 2017). Aunque no incorpora IA, existen iniciativas para explorar su integración. Decidim está plenamente activa; ha sido usada por más de 400 instituciones públicas y organizaciones en todo el mundo.

2.3.10 Kialo



Kialo es una plataforma en línea lanzada en 2017 que facilita debates estructurados. El proceso deliberativo en Kialo comienza con la formulación de una tesis central. Los participantes añaden argumentos "pro" o "contra", que a su vez pueden ser respaldados o refutados por otros, creando un mapa detallado de la discusión. Cada nodo es votado por los usuarios en función de su relevancia y persuasión.

Su principal fortaleza es el orden argumentativo que impone, ayudando a evitar redundancias y desviaciones (Kialo Technologies, 2021). Cada argumento puede ser comentado, evaluado y vinculado a evidencia externa, facilitando una comprensión profunda de temas complejos. Su formato visual, con mapas interactivos (figura 6), ayuda a los usuarios a entender cómo se relacionan diferentes ideas, promoviendo una comprensión más profunda de los temas discutidos.

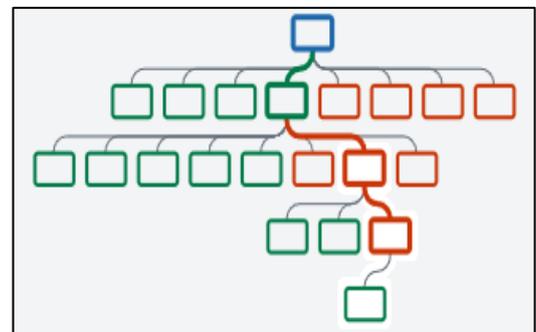


Figura 6. Jerarquía de argumentos a favor y en contra en Kialo (Fuente: Kialo)

Aunque Kialo no utiliza inteligencia artificial para analizar los contenidos, su estructura facilita el análisis por parte de los investigadores en áreas como el procesamiento del lenguaje natural y la minería de argumentos. Kialo se utiliza en contextos educativos, empresariales y gubernamentales, y cuenta con una versión específica para el ámbito educativo llamada Kialo Edu, que permite a profesores y estudiantes participar en debates privados y estructurados.

2.3.10 Dlbrt



dlbrt (Deliberate) es una plataforma de deliberación colectiva desarrollada en 2023 por el Deliberative Democracy Lab de la Universidad de Stanford en colaboración con la empresa Novelcore. (Novelcore, 2023). De código propietario, la plataforma no utiliza inteligencia artificial para analizar el contenido de las discusiones, pero incorpora herramientas automatizadas para facilitar la moderación y garantizar una participación equitativa.

Ante una determinada temática o problema, va recogiendo aportes agrupados en tendencias sobre los que los participantes deliberan y que a su vez son comentadas y evaluadas como a favor/en contra/neutro. dlbrt parece centrarse principalmente en ofrecer una experiencia visual



Figura 7. Votación de una tendencia en dlbrt (Fuente: dlbrt)

más rica que otras herramientas como Loomio, pero sin demostrar claramente las capacidades de análisis automatizado. Ha sido utilizada en diversos contextos, incluyendo eventos nacionales de deliberación en países como Chile, Canadá y EEUU, y está diseñada para ser escalable.

2.3.11 RESUMEN

En resumen, las herramientas TIC de facilitación de la inteligencia colectiva están evolucionando desde simples sistemas de votación hacia entornos complejos de deliberación estructurada, generación de propuestas y búsqueda de consenso asistido por IA. Desde modelos centrados en el orden argumentativo (como Kialo o iWarsM'ap) hasta soluciones que incorporan procesamiento estadístico y aprendizaje automático (como Polis), se evidencia una convergencia entre los principios de la democracia deliberativa y las posibilidades tecnológicas contemporáneas (Ver tabla 1 y figura 9).

Tabla 1. Tabla de Herramientas TIC relevantes

| Plataforma/Aplicación | Año | Open Source | Activa | Uso de IA | Algoritmo/Foco principal | Escalabilidad (participantes) | Carga de deliberación | Dominio principal / Observaciones |
|-------------------------|------|-------------|--------|-----------|---|-------------------------------|-----------------------|---|
| ThoughtExchange | 2009 | No | Sí | Sí (2023) | IA generativa + agrupamiento de ideas | Alta | Media | Educativo (K-12). Integrado con OpenAI (2023) |
| LiquidFeedback | 2009 | Sí | Sí | No | Votación preferencial, delegación, liquidez | Alta | Baja | Democracia líquida, trazabilidad completa |
| Mesydel | 2008 | No | Sí | No | Método Delphi digital, análisis cualitativo | Media | Baja | Académico, institucional, universidades |
| Loomio | 2011 | Sí | Sí | No | Consenso por discusión estructurada | Media | Baja | Colectivos, ONGs, diseño accesible, cooperativo |
| Appgree | 2013 | No | No | No | Valoración binaria (sí/no), ranking de ideas | Alta | Baja | Medios, campañas. Inspiración en ELO |
| iWarsM'aps / CiBUC | 2014 | No | No | Sí | Delphi + por fases: clusterización, enjambre | Media | Alta | Consultoría, innovación. IA supervisada |
| Vilfredo Goes to Athens | 2015 | Parcial | No | No | Comparación por pares + frente de Pareto | Media | Alta | Experimental / académico, Delphi sin votación |
| Pol.is | 2015 | Sí | Sí | Sí | Análisis PCA, clustering de opiniones | Alta | Alta | Gobierno, part. Ciudadana. Usado en vTaiwan |
| Decidim | 2016 | Sí | Sí | No (aún) | Modulos para procesos participativos | Alta | Media | Gobiernos locales, ONGs. Transparencia |
| Kialo | 2017 | No | Sí | No | Árbol argumental estructurado | Alta | Alta | Educación, debate online, trazabilidad argumental |
| dlbrt | 2023 | No | Sí | No | Deliberative Polling, Moderación automatizada | Alta | Alta | Grupos simultáneos con evaluación post-debate |

Un patrón común es el intento de superar la lógica binaria o reduccionista del voto tradicional, permitiendo una representación más rica de las preferencias, así como una mejora en la calidad del diálogo. No obstante, muchas de estas herramientas aún adolecen de problemas de escalabilidad, interoperabilidad y autonomía semántica. Algunas excepciones notables, como Polis, han demostrado viabilidad en procesos participativos a gran escala, pero en general, pocas han integrado todavía capacidades avanzadas de razonamiento automático o PLN capaces de intervenir activamente en la síntesis de propuestas.

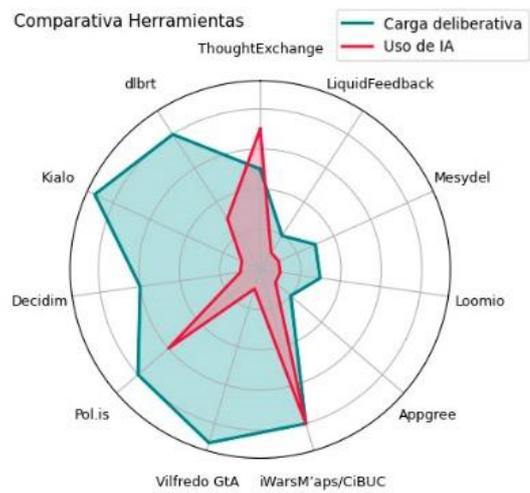


Figura 8. Grafico radar de herramientas

Este panorama justifica la necesidad de explorar nuevos enfoques híbridos, como el que se plantea en este trabajo, que combinen la estructuración deliberativa con la capacidad analítica de los sistemas inteligentes, en busca de procesos colectivos más robustos, inclusivos y consensuales.

2.4 INTEGRACIÓN EMERGENTE DE IA EN PROCESOS DELIBERATIVOS

La aplicación de IA, el procesamiento del lenguaje natural (PLN) y el uso de modelos de lenguaje de gran escala (LLMs) en contextos deliberativos está en una etapa inicial, aunque emergen investigaciones y prototipos que muestran su potencial transformador en la estruc-

turación del discurso colectivo. Diversos estudios recientes abordan este desafío desde enfoques complementarios, algunos centrados en la representación lingüística, otros en la automatización de tareas asociadas a la deliberación o la mediación argumentativa.

El trabajo de Dong y Herrera-Viedma (2020) constituye una referencia fundamental en este ámbito. Proponen un enfoque computacional lingüístico para apoyar procesos de consenso en la toma de decisiones grupales. Aunque no utilizan LLMs modernos como GPT, emplean representaciones lingüísticas difusas y técnicas de decisión multicriterio que permiten modelar la incertidumbre y la vaguedad en contextos deliberativos.

Serrano et al. (2021) desarrollan técnicas de aprendizaje automático, incluyendo clustering y resumen automático, para mejorar plataformas de participación ciudadana. Si bien no integran LLMs, su trabajo sienta las bases para futuras incorporaciones de modelos más avanzados en la automatización de tareas deliberativas.

Binns et al. (2023) exploran la co-creación de sistemas de IA responsables mediante la participación ciudadana. Esta contribución, de carácter más conceptual, tampoco se basa en la implementación directa de LLMs, más bien reflexiona críticamente sobre su rol potencial en la construcción de herramientas auditables, explicables y centradas en el usuario. Se destaca la importancia de diseñar sistemas deliberativos auditables y explicables, aspectos en los que los LLMs podrían desempeñar un papel crucial, pero con participación de los usuarios.

En América Latina, Negoc-IA, una herramienta chilena, incorpora un asistente virtual basado en modelos generativos de lenguaje para apoyar negociaciones colectivas, ofreciendo sugerencias y síntesis en tiempo real. Aunque no se especifica públicamente el modelo exacto empleado, se infiere el uso de un LLM de tipo GPT-3 o GPT-4, posiblemente accedido a través de una API comercial. Su implementación práctica demuestra la viabilidad de aplicar IA generativa en escenarios complejos de interacción estratégica y argumentación.

Prometea, desarrollado por la Fiscalía General en Buenos Aires, es un sistema de IA aplicado a tareas jurídico-deliberativas. Inicialmente basado en procesamiento de lenguaje natural (PLN) estructurado y técnicas de automatización incluyendo árboles de decisión, recientes implementaciones han incorporado el uso de IA generativa con ChatGPT para la redacción de resoluciones judiciales, evidenciando la transición hacia IA generativa en el ámbito legal.

El ejemplo más avanzado de integración efectiva de LLMs generativos en procesos deliberativos es el experimento desarrollado por Google DeepMind y la Universidad de Oxford, publicado en Science (Stuhlmüller et al., 2024). En este trabajo se presentó un sistema informalmente denominado “La Máquina de Habermas”, que facilitó la deliberación de más de 5.000 ciudadanos sobre temas controvertidos como la privatización del sistema de salud o la edad mínima para votar. El sistema utilizó un modelo comparable a GPT-4, entrenado y ajustado finamente para sintetizar declaraciones grupales a partir de opiniones individuales. Estas síntesis eran refinadas en ciclos iterativos con retroalimentación humana, y en evaluaciones a ciegas, los participantes prefirieron en un 56 % las declaraciones generadas por la IA frente a las humanas, valorando su claridad, imparcialidad y capacidad de integrar puntos de vista diversos. Esto constituye una demostración empírica de que los LLMs pueden actuar como mediadores deliberativos, ayudando a generar consensos amplios a partir de visiones plurales.

Tabla 2. Cuadro resumen de uso incipiente de IA en procesos deliberativos.

| Trabajo / Proyecto | Año | Tipo de modelo o técnica IA empleada | Uso de LLMs | Integración deliberativa |
|-----------------------|------|---|-------------|--------------------------|
| Dong & Herrera-Viedma | 2020 | Representaciones lingüísticas difusas, MCDM | No | Medio (decisión) |
| Serrano et al. | 2021 | ML clásico y PLN para agrupamiento y resumen | No | Bajo-medio |
| Binns et al. | 2023 | Marco para IA deliberativa con co-diseño ciudadano | No | Alto (conceptual) |
| Stuhlmüller et al. | 2024 | LLM tipo GPT-4 con fine-tuning y retroalimentación humana | Sí | Muy alto |
| Negoc-IA | 2024 | LLM generativo (probablemente GPT-3 o GPT-4 vía API) | Sí | Alto |
| Prometea (vigente) | 2018 | PLN estructurado con reglas. Sin LLMs actualmente | No | Bajo |

Estas experiencias, resumidas en la tabla 2, evidencian una tendencia creciente hacia el diseño de sistemas híbridos, donde la inteligencia artificial no reemplaza la deliberación humana, sino que la amplifica y estructura. No obstante, la integración de modelos de lenguaje en la facilitación de decisiones colectivas plantea desafíos significativos en torno a la transparencia algorítmica, la trazabilidad del razonamiento, la mitigación de sesgos y la validación ética de las decisiones automatizadas. Aun así, los avances recientes apuntan hacia un horizonte en el que la IA no sólo puede aumentar la eficiencia de los procesos deliberativos, sino también mejorar su inclusividad, calidad argumentativa y capacidad de síntesis colectiva, facilitando procesos deliberativos humanos, ampliando el alcance y la calidad de la participación democrática.

Este texto es un extracto del TFM de Carlos Rossique Delmas.
Puede usarse citando al autor, según la licencia CC-BY-NC-SA.

