

Implementación de la Inteligencia Artificial en las Evaluaciones de Diseño de los Programas Presupuestarios en México

Semana de la Evaluación 2025

Junio 2025



Índice	
1. Introducción	3
2. Retos Normativos, Regulatorios, Éticos y de Implementación de la IA	4
2.1 Desafíos Normativos y Regulatorios	4
2.2 Consideraciones Éticas en la Implementación de IA	7
3. Problemas Metodológicos en la Implementación de la IA	9
3.1 Definición del problema	9
3.2 Dificultades en la recopilación y estructuración de datos	10
3.3 Interpretación y validación de resultados generados por IA	12
3.4 Limitaciones de modelos de IA	13
4. Comparación de Metodologías para la Implementación de IA en Evaluaciones Presupuestarias	15
4.1 Elementos centrales para introducir en la IA	15
4.2 Selección de Programas presupuestarios: B001 y P002	15
4.3 Metodologías de implementación	16
4.4 Selección de los elementos	17
4.4.1 Selección de los TdR	17
4.4.2 Selección de las herramientas de IA	19
4.5 Evaluación Comparativa de Metodologías en Aplicaciones de IA	19
4.5.1 Escenario sin conocimiento de los elementos centrales para la evaluación	20
4.5.2 Escenario con conocimiento	37
4.5.3 Escenario con conocimiento	55
4.6 Aplicación de la IA en los Modelos de TdR 2025 (Evaluación de Consistencia y Resultados, e Indicadores)	80
4.6.1 Evaluación de Consistencia y Resultados	80
4.6.2 Evaluación de Indicadores	85
5. Discusión de Resultados	114
6. Conclusiones y Recomendaciones	116
6.1 Resumen de hallazgos	116
6.2 Propuesta de mejores prácticas	116
6.3 Sigüientes pasos en la investigación y desarrollo de IA en la evaluación de programas presupuestarios	117
7. Referencias	119

1. Introducción

La inteligencia artificial (IA), en los últimos años, ha surgido como una herramienta que permite realizar tareas complejas en una menor cantidad de tiempo, que normalmente requieren de la inteligencia y actividad humana, como el aprendizaje y el razonamiento (Gobierno de España, 2023).

Los Modelos de Lenguaje de Gran Tamaño (LLM, por sus siglas en inglés), englobados dentro de la Inteligencia Artificial, son sistemas capaces de procesar información y generar contenido textual con características similares a las humanas; estos modelos fundamentan su aprendizaje a través de grandes repositorios de información, lo que les permite desarrollar una comprensión contextual de lo que se les solicita, y generar en lenguaje natural una amplia variedad de respuestas (AWS, 2025).

La popularización de los LLM ha sido impulsada por el auge de interfaces accesibles y públicas, como ChatGPT, Claude, Gemini, o la más reciente DeepSeek, solo por mencionar algunas; estas han permitido que usuarios sin conocimientos técnicos puedan beneficiarse de herramientas basadas en IA, por ejemplo, en el campo académico y de investigación, estos modelos funcionan como asistentes que sintetizan información, sugieren ideas y agilizan procesos de revisión bibliográfica.

De acuerdo con Ocaña et al. (2021), la IA tiene el potencial para optimizar procesos administrativos, prever tendencias y facilitar la personalización de políticas públicas; mediante el análisis de grandes volúmenes de datos y la automatización de tareas repetitivas, permitiendo a los gobiernos tomar decisiones, así como ofrecer respuestas rápidas y adaptadas a las necesidades de la sociedad.

La IA generativa, como los LLM, se han establecido como herramientas que, además de reconfigurar nuestra interacción con la tecnología y el acceso a la información, está transformando las industrias y sectores mediante la optimización de procesos y ampliación de posibilidades de colaboración entre humanos y sistemas inteligentes, consolidando así los cimientos para una era de inteligencia artificial aplicada global.

La evaluación de programas es uno de las muchas áreas que se ha visto beneficiada en el uso de la IA. El presente trabajo tiene como objetivo central analizar los desafíos y oportunidades que surgen de la integración de la IA en la evaluación de programas presupuestarios, específicamente en programas no sociales de la Secretaría de Hacienda y Crédito Público. De igual manera se abordarán los problemas metodológicos inherentes al uso de estas herramientas, las implicaciones normativas que deben considerarse, los dilemas éticos que emergen, y finalmente se explorarán las metodologías viables para una implementación efectiva de las herramientas de IA para la evaluación de programas.

2. Retos Normativos, Regulatorios, Éticos y de Implementación de la IA

2.1 Desafíos Normativos y Regulatorios

En México no existe una ley específica que prohíba o regule el uso de inteligencia artificial (IA) en la evaluación de política pública, sin embargo, el uso de la IA en estos ámbitos está sujeto a la normatividad general en materia de transparencia, rendición de cuentas, y protección de datos personales.

El marco legal actual establece principios éticos y legales que deben observarse en la implementación de cualquier tecnología en el sector público, pero no se han establecido regulaciones específicas que prohíban o restrinjan de forma explícita el uso de la IA en estos procesos.

Se han hecho múltiples esfuerzos institucionales y legislativos relacionados con el tema; en 2018, la Coordinación de la Estrategia Digital Nacional publicó el documento “Principios Generales y Guía de Análisis de Impacto para el desarrollo y uso de sistemas con elementos de Inteligencia Artificial en la Administración Pública Federal en México”.

Dentro de sus propuestas también se encontraba:

- a) Establecer una guía y principios generales para el desarrollo de la IA en beneficio de la APF
- b) Implementar en las distintas agencias de gobierno el uso de la IA, ya sea en procesos administrativos o en la entrega de servicios a la ciudadanía
- c) Desarrollar un marco de gobernanza adecuado para fomentar el diálogo multisectorial, a través de la creación de una Comisión.
- d) Mapear los usos y necesidades en la industria e identificar mejores prácticas en gobierno.
- e) Impulsar el liderazgo internacional de México en la materia.
- f) Fortalecer la coordinación y sinergias entre sectores

El Artículo 6 del acuerdo para impulsar el uso y aprovechamiento de la informática, el gobierno digital, las tecnologías de la información y comunicación, y la seguridad de la información en la Administración Pública Federal (2021); señala que todos los proyectos de Tecnologías de la Información y Comunicación (TIC) que las Instituciones pretendan ejecutar a través de prácticas de desarrollo, implementación propia, o a través de contrataciones, deben apegarse a un proceso de planeación estratégica alineado a las disposiciones del Plan Nacional de Desarrollo y programas que de él deriven, la Estrategia Digital Nacional, así como a la legislación en materia de desarrollo nacional, presupuesto, austeridad y transparencia.

La Estrategia Digital Nacional 2021 - 2025 (EDN) expone el conjunto de acciones para hacer posible el uso y desarrollo eficiente, democrático e incluyente de las TIC para aprovechar el potencial de éstas, mediante su incorporación a la vida cotidiana de las personas y a la Administración Pública Federal, con el fin de alcanzar un Estado de Bienestar.

Los principios de la EDN son el conjunto de conceptos que guían y respaldan las diferentes acciones y decisiones de la política tecnológica del Gobierno de México, aplicables a todas las iniciativas de proyectos y uso de TIC en el gobierno.

- **Principio de Austeridad:** Principio de bien común, relativo a lograr servicios de alta calidad con el máximo aprovechamiento de recursos y disminución del gasto.
- **Principio de Combate a la Corrupción:** Acabar con prácticas injustas, desleales, leoninas y perversas que benefician intereses particulares perjudicando al Estado o a sus integrantes.
- **Principio de Eficiencia en los procesos Digitales:** Implica la simplificación operativa y atención focalizada de los procedimientos gubernamentales.
- **Principio de Seguridad de la Información:** Concepto que hace referencia a la estabilidad, protección y certidumbre de la información generada resguardada en sistemas o plataformas digitales. El desarrollo de dichas plataformas y sistemas también deben brindar estabilidad y certidumbre de su operación.
- **Principio de Soberanía Tecnológica:** Atribución única de la nación para tomar decisiones sin interferencias externas sobre qué políticas y estrategias se deben seguir en torno al ámbito digital y tecnológico.

En materia de protección de datos e información, de conformidad con el artículo 110 de la Ley Federal de Transparencia y Acceso a la Información Pública, se considera información reservada aquella que compromete la seguridad nacional, la seguridad pública o la defensa nacional; que ponga en riesgo la vida, seguridad o salud de una persona física; que obstruya la prevención o persecución de los delitos; que vulnere la conducción de los Expedientes judiciales o de los procedimientos administrativos seguidos en forma de juicio; que se encuentre contenida dentro de las investigaciones de hechos que la ley señale como delitos; y las que por disposición expresa de una ley tengan tal carácter.

Mientras que el artículo 113, considera información confidencial aquella que contenga datos personales concernientes a una persona física identificada o identificable; los secretos bancario, fiduciario, industrial, comercial, fiscal, bursátil y postal; y aquella que presenten los particulares a los sujetos obligados.

Bajo este marco, dentro de los lineamientos de clasificación y valuación de la información, la Secretaría de Hacienda y Crédito Público, clasifica la información de acuerdo con su nivel de confidencialidad en sus Lineamientos de Seguridad de la Información, como se muestra a continuación:

- **Confidencial:** Información que contenga datos personales, así como, secretos bancarios, fiscal y bursátil; dicha información no está sujeta a temporalidad alguna.

- **Reservada:** Información que se encuentra temporalmente fuera del acceso público, y además pueda comprometer los objetivos o actividades de la Secretaría, la seguridad nacional, poner en riesgo la estabilidad de instituciones o sistema financiero del país, pueda comprometer la seguridad de la moneda nacional u obstruir las actividades relativas al cumplimiento de leyes o afecte la recaudación de contribuciones.
- **Pública:** Información destinada al público en general y que en caso de publicación externa no existe impacto.

Es importante mencionar que, de acuerdo a los Lineamientos, todo servidor público de la SHCP debe de custodiar y cuidar la documentación y recursos de información que tengan bajo su resguardo, así como impedir o evitar su uso indebido.

Por último, el 15 de abril del 2025 se publicó el Plan Nacional de Desarrollo (2025 – 2030), donde el Eje Transversal 2: Innovación Pública para el Desarrollo Tecnológico Nacional, tiene como objetivo dotar al Estado mexicano de capacidades físicas, tecnológicas y humanas para modernizar las funciones del Gobierno en sus tres niveles, mediante una transformación digital y tecnológica, que permita un ahorro significativo de tiempo y recursos, todo ello en un marco de transparencia y eficacia en la gestión pública. Dentro de sus estrategias se encuentran:

- **T2.1.2:** Establecer un modelo nacional que modernice las operaciones gubernamentales mediante la simplificación y digitalización de trámites administrativos, cumpliendo con reglas de gestión documental, facilitando el acceso a derechos, promoviendo la inversión y eliminando prácticas de corrupción.
- **T2.2.4:** Incrementar la productividad del sector público mediante la capacitación continua de sus funcionarios en áreas clave como tecnología pública, desarrollo de software y ciberseguridad, asegurando que el personal esté preparado para afrontar los desafíos digitales actuales y contribuir al fortalecimiento de los servicios públicos.
- **T2.4.4:** Implementar mecanismos de vinculación, modernización y transferencia tecnológica e innovación entre los sectores académico, productivo, público y social, alineados con las prioridades nacionales y orientados al bienestar social.

En materia de evaluación, el artículo 110, párrafo cuarto, fracción I de la LFPRH señala que se efectuarán las evaluaciones por sí mismas o a través de personas físicas y morales especializadas y con experiencia probada en la materia que corresponda evaluar, que cumplan con los requisitos de independencia, imparcialidad, transparencia y los demás que se establezcan en las disposiciones aplicables.

De igual manera, el artículo 27, fracción V, inciso a) del DPEF 2025, señala que las evaluaciones de programas presupuestarios podrán ser efectuadas en los siguientes términos:

- a) Por sí mismas
- b) A través de personas físicas y morales especializadas y con experiencia probada en la materia que corresponda evaluar, instituciones académicas y de investigación,

u organismos especializados, de carácter nacional o internacional, que cuenten con reconocimiento y experiencia en las respectivas materias de los programas.

En concordancia con los párrafos, el PAE 2025 señala en su numeral 20 que en los casos en los que las dependencias y entidades de la APF opten por realizar evaluaciones en los términos que establece el artículo 110, párrafo cuarto, fracción I de la LFPRH y el artículo 27, fracción V, inciso a) del DPEF 2025, deberán notificarlo a la UPER, observando lo establecido en este PAE, así como en las "Disposiciones para que las Dependencias y Entidades de la Administración Pública Federal evalúen por sí mismas los Programas Presupuestarios para 2025".

2.2 Consideraciones Éticas en la Implementación de IA

Diversos autores (Ausín, 2021; Sebio, 2020; García, 2019), han mostrado su preocupación por la privacidad y protección de datos, los sesgos, la manipulación del comportamiento, la transparencia y la seguridad que representan el uso de la IA, los cuales pueden comprometer la calidad y equidad en la formulación y evaluación de programas si no se gestionan de forma adecuada.

Privacidad y Protección de Datos

Las diferentes herramientas basadas en IA, además de manejar grandes volúmenes de información, requieren de la información que se le proporcione para poder brindar respuestas apropiadas con lo que se le está preguntando, información que es almacenada en los servidores para mejorar la calidad del servicio y adaptarlo a las necesidades individuales de los usuarios.

La recopilación y análisis de datos personales, económicos y sociales, así como de las entidades gubernamentales para su uso en herramientas de uso libre, como Chat GPT, representa un riesgo a la privacidad de la información sensible de las Unidades. Para atenuar este riesgo, es esencial implementar protocolos en el manejo de la información de carácter privado y compartir con la IA aquella que es de carácter público, sitios web gubernamentales.

Sesgos

Se entiende por sesgo a un prejuicio en la forma de procesar la información, con tendencia a percibir de forma distorsionada la realidad (Ausín, 2021).

Los algoritmos de IA no son neutrales; reflejan y amplifican los sesgos presentes en sus datos de entrenamiento, traducándose en decisiones que pueden favorecer a ciertos grupos, o áreas, sobre otros, generando inequidades y reforzando la exclusión social y la estratificación.

Es importante mencionar que, en la evaluación de programas no sociales, estos sesgos llegan a presentarse en menor frecuencia que en programas sociales, donde los problemas suelen estar ligados a cuestiones históricas, políticas, sociales, con más de un actor involucrado.

La importancia de los sesgos en temas no sociales y la IA, de acuerdo con Austín (2021), radica en la posibilidad de caer en una *dictadura de datos*, donde no se toman acciones con base a juicios de acciones y resultados reales, sino sobre la base de lo que los datos y la IA indiquen serán acciones probables, reforzando patrones previos sin cuestionar su validez.

Desafortunadamente, como particulares, no se puede rediseñar o modificar el algoritmo y la base de datos con las que funcionan las herramientas de IA, sin embargo, la participación activa en la revisión de la información generada por la IA para mitigar sesgos, puede garantizar respuestas objetivas y equitativas.

Manejo en la Toma de Decisiones

Las capacidades de la IA para analizar y predecir situaciones o comportamientos, pueden ser utilizadas para influir en las decisiones de las personas, lo que plantea el riesgo de un mal manejo en la toma de decisiones. En la administración pública, los sistemas de IA pueden influir en la percepción de prioridades de inversión, manipulando la opinión pública y condicionando la toma de decisiones. Al igual que en el punto anterior, las recomendaciones generadas por IA deben ser revisadas y validadas por expertos en política pública antes de tomarse como definitivas.

Manipulación, Seguridad y Vulnerabilidad

Al igual que otros sistemas informáticos, la IA se compone de elementos de software y hardware susceptibles de funcionar erróneamente, y representar riesgos de seguridad, como ciberataques, manipulación de datos y vulnerabilidades en los sistemas de toma de decisiones.

Si una IA utilizada para evaluar programas y proyectos es comprometida, podría generar resultados erróneos o sesgados, afectando la efectividad de las políticas públicas. Para mitigar estos riesgos, es importante implementar medidas de ciberseguridad, así como contar con protocolos de respuesta ante incidentes de seguridad.

Cuidado del Medio Ambiente

Los grandes centros de computación y de almacenamiento de datos en la nube, usados por las IA, requieren enormes cantidades de energía y tienen una alta huella por emisiones de CO₂, con un impacto medioambiental muy elevado, además el alto consumo eléctrico genera una gran cantidad de emisiones de carbono (Winfield 2019), es por esto que el desarrollo de infraestructuras vinculadas a la IA deberá buscar la sostenibilidad medioambiental, sin embargo la tendencia en el uso de estas herramientas va a la alza.

Transparencia

La tendencia de asumir la información generada por la IA como propia, sin cuestionar su origen o validez, generar un exceso de confianza en las recomendaciones algorítmicas, reduciendo el análisis crítico en la toma de decisiones.

Para garantizar transparencia, es necesario el desarrollo de metodologías, así como la publicación de códigos y parámetros utilizados, que contribuyan a una mayor claridad y confianza en el uso de estas tecnologías.

3. Problemas Metodológicos en la Implementación de la IA

3.1 Definición del problema

Los avances en inteligencia artificial, incluidos los Modelos de Lenguaje de Gran Tamaño (LLM), representan una oportunidad para transformar los procesos de evaluación, a través de una mayor precisión, profundidad y capacidad analítica de los programas y el contexto en el que operan.

Entre los aportes del uso de IA en la evaluación de programas se encuentra el análisis documental y procesamiento de información, que tradicionalmente requiere del análisis de una gran cantidad de información y documentación, como evaluaciones anteriores, diagnósticos, normatividad, matrices de indicadores, manuales de procedimientos, por parte de los evaluadores.

Los LLM destacan por su capacidad de procesar grandes cantidades de texto, logrando extraer, categorizar y relacionar información relevante con una rapidez superior a las capacidades humanas. Además, estos modelos facilitan la integración de datos de diversas fuentes y formatos, tales como estadísticas oficiales, encuestas, registros administrativos, datos geoespaciales, permitiendo un análisis integral sobre el desempeño programático.

Estas herramientas además permiten optimizar tareas que exigen una gran cantidad de recursos humanos, económicos y de tiempo, logrando facilitar el análisis de información, la identificación de patrones y la generación de insumos para la toma de decisiones basada en evidencia, en un lapso de tiempo mínimo y con una menor cantidad de recursos.

Un aspecto de gran importancia en la evaluación de los programas, es el análisis de la coherencia interna del programa y diseño programático, donde se debe verificar si los indicadores reflejan de manera precisa el cumplimiento de los objetivos propuestos, así como la alineación entre la problemática identificada y las actividades implementadas. Los LLM tienen la capacidad de realizar un análisis de estas relaciones, identificando posibles desalineaciones e inconsistencias entre la planificación y la ejecución que podrían comprometer la efectividad de los programas.

De igual manera, los LLM destacan por su capacidad para simular escenarios y proyectar resultados; con la información contextual necesaria, estos modelos pueden simular el impacto potencial de modificaciones realizadas al diseño del programa, permitiendo a los tomadores de decisiones, y evaluadores, anticipar las posibles consecuencias previo al cambio programático.

Otro aspecto que puede beneficiarse del uso de la IA en la evaluación, es la generación de informes de evaluación, pues los LLM pueden agilizar este proceso generando borradores estructurados que sinteticen hallazgos clave y propongan recomendaciones basadas en patrones identificados. Estos borradores pueden luego ser revisados y refinados por los evaluadores, optimizando el proceso y los tiempos de entrega, y reservando criterios particulares a los evaluadores.

Como se puede observar, la integración de la IA en la evaluación de programas, representa una oportunidad transformadora para la gestión y administración pública; desde el análisis documental hasta la creación de informes, la IA ofrece herramientas que pueden elevar la calidad de estas evaluaciones. Sin embargo, plasmar este potencial requiere un enfoque

estratégico que aborde los desafíos técnicos, institucionales y éticos; las administraciones que logren implementar estas tecnologías, manteniendo un equilibrio adecuado entre innovación y responsabilidad, lograrán una mejor posición para optimizar el uso de los recursos públicos y maximizar el valor agregado.

3.2 Dificultades en la recopilación y estructuración de datos

En la aplicación y uso de tecnologías basadas en IA, para la evaluación de programas presupuestarios, es importante reconocer los desafíos relacionados con la recopilación y estructuración de los datos que alimentarán estos sistemas, pues representan limitaciones que pueden afectar la efectividad estas herramientas.

Datos Abiertos y su Disponibilidad en México

El artículo 4 de la Ley General de Transparencia y Acceso a la Información Pública (2021), señala que toda la información generada, obtenida, adquirida, transformada o en posesión de los sujetos obligados, es pública y accesible a cualquier persona; y sólo podrá ser clasificada como reservada temporalmente por razones de interés público y seguridad nacional.

Bajo la misma ley, el artículo 6 señala que el Estado es quien garantizará el efectivo acceso de toda persona a la información en posesión de cualquier entidad, órgano, partido político, fideicomisos y fondos públicos; así como de cualquier persona física, moral o sindicato que reciba y ejerza recursos públicos.

En México se ha experimentado un desarrollo importante en el fortalecimiento de políticas de datos abiertos; por ejemplo, iniciativas como la Política Nacional de Datos Abiertos, o la plataforma transparenciapresupuestaria.gob.mx, son herramientas que promueven no solo el acceso a la información gubernamental, sino que también promueven la rendición de cuentas.

A pesar de estos esfuerzos, la disponibilidad de los datos enfrenta múltiples desafíos, como la fragmentación institucional, donde la información gubernamental se encuentran distribuida entre múltiples plataformas e instituciones (Escalante, 2023), tales como la Secretaría de Hacienda y Crédito Público (SHCP), la Auditoría Superior de la Federación (ASF), el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), los portales de transparencia, entre otros; dificultando los pasos a seguir para saber dónde consultar los datos más actualizados (Escalante, 2023). Para las diferentes herramientas de IA con acceso de información en tiempo real, esta fragmentación institución representa una barrera en la recopilación y agrupación de los datos relevantes para la evaluación.

Otro factor que se observa es que los datos públicos, frecuentemente carecen del nivel de detalle necesario para análisis específicos, pues la información se encuentra localizada en amplias bases de datos proporcionadas por múltiples dependencias y órganos de gobierno, las cuales son complejas para herramientas de IA que no cuentan con la programación para leer hojas de cálculo, así como extensos documentos en lenguaje técnico (Díaz, 2023).

Si bien la manipulación humana de estos documentos puede hacerlos comprensibles, las características técnicas actuales de la IA impiden que estos documentos sean comprensibles por sí solos.

Otro problema que se detecta, es la falta de continuidad en las series temporales de datos; cambios en metodologías de registro, reclasificaciones presupuestarias, o reorganizaciones administrativas sexenales o trienales, hacen que datos entre diferentes períodos no sean directamente comparables (Rosales, 2013).

Por último, a pesar de los esfuerzos por estandarizar la información, existen una gran variedad de formatos en la publicación de información, los cuales van desde archivos PDF u hojas de cálculo en diferentes formatos, como .xlsx, .xml, .txt; hasta flujogramas en formato. vds, es por esto que esta falta de homogeneidad en los formatos en los que se reporta la información, complica el ingreso automatizado de datos que requieren los LLM.

Pese al marco normativo existente, que promueve la transparencia y el acceso a la información pública, las barreras en la recopilación y estructuración de datos abiertos en México enfrenta múltiples desafíos que afectan la calidad de la información disponible, influyendo negativamente de forma indirecta la calidad de los datos captados por las diferentes herramientas de IA.

Es importante mencionar que las diferentes herramientas basadas en IA, además de manejar grandes volúmenes de información, requieren de la información que se le proporcione para poder brindar respuestas más acertadas a lo que requiere la instancia evaluadora, información que después es almacenada en los servidores para mejorar la calidad del servicio y adaptarlo a las necesidades futuras de los usuarios.

La recopilación y análisis de datos personales, económicos y sociales, así como de las entidades gubernamentales para su uso en herramientas de uso libre, como Chat GPT, representa un riesgo a la privacidad de la información sensible de las Unidades. Para atenuar este riesgo, es esencial implementar protocolos en el manejo de la información de carácter privado y compartir con la IA aquella que es de carácter público, sitios web gubernamentales.

Sesgos en los Datos y Calidad de la Información

Se entiende por sesgo a un prejuicio en la forma de procesar la información, con tendencia a percibir de forma distorsionada la realidad (Ausín, 2021).

Los algoritmos de IA no son neutrales; reflejan y amplifican los sesgos presentes en sus datos de entrenamiento, traducándose en decisiones que pueden favorecer a ciertos grupos, o áreas, sobre otros, generando inequidades y reforzando la exclusión social y la estratificación.

El principal sesgo que se identifica es la pérdida y/o degradación de la información, generando desinformación intencional o accidental dentro de los repositorios de información; se visibilizada que los monopolios de la IA, ejercen una influencia de valores e ideales con una carga principalmente occidental, afectando así la neutralidad en la investigación (da Mota, 2024).

La importancia de los sesgos en temas no sociales y la IA, de acuerdo con Austín (2021), radica en la posibilidad de caer en una dictadura de datos, donde no se toman acciones con base a juicios de acciones y resultados reales, sino sobre la base de lo que los datos y la IA indiquen serán acciones probables, reforzando patrones previos sin cuestionar su validez.

Es importante reconocer que los datos disponibles suelen ser una sobre representación de lo medible cuantitativamente, dejando fuera aspectos cualitativos igualmente importantes, por lo que esta selección sesgada de variables modifica el panorama presentado por la IA.

A pesar de que en la evaluación de programas no sociales, estos sesgos llegan a presentarse en menor frecuencia que en programas sociales, donde los problemas suelen estar ligados a cuestiones históricas, políticas, sociales, con más de un actor involucrado; como particulares, no se puede rediseñar o modificar los algoritmos y la base de datos con las que funcionan las herramientas de IA, sin embargo, la participación activa de las partes interesadas en la revisión de la información generada por la IA, puede garantizar respuestas objetivas y equitativas, con la finalidad de equilibrar cualquier sesgo que pueda existir (da Mota, 2024).

3.3 Interpretación y validación de resultados generados por IA

A medida que los sistemas basados en IA se vuelven parte de las actividades cotidianas y académicas, surge la necesidad de desarrollar métodos para interpretar y validar los resultados que estas generan, y saber si se debe confiar en los resultados o si se debe prescindir de ellas.

Transparencia y trazabilidad de la IA en la toma de decisiones

En la administración pública, los sistemas de IA pueden influir en la percepción de prioridades de inversión, manipulando la opinión pública y condición de la toma de decisiones; en la implementación de la inteligencia artificial en la administración pública, determinar quién es responsable de las acciones de un sistema de inteligencia artificial resulta complicado, lo que puede dificultar el proceso de reclamación de responsabilidades.

La integración de la inteligencia artificial en la evaluación de programas presupuestarios, enfrenta desafíos relacionados con la transparencia y trazabilidad de la IA en la toma de decisiones, pues la tendencia de asumir la información generada por la IA como cierta, sin cuestionar su origen o validez, generar un exceso de confianza en las recomendaciones algorítmicas, reduciendo el análisis crítico en la toma de decisiones.

Para garantizar transparencia, es necesario el desarrollo de metodologías, así como la publicación de códigos y parámetros utilizados, que contribuyan a una mayor claridad y confianza en el uso de estas tecnologías. Bajo esta lógica, las instituciones requieren implementar niveles de transparencia, brindando explicaciones técnicas detalladas para auditores y analistas presupuestarios, así como resúmenes ejecutivos para los tomadores de decisiones.

La OCDE emitió en 2023 una serie de recomendaciones y principios para un correcto desarrollo, seguimiento y regulación de la inteligencia artificial en favor de la sociedad. Dentro de sus recomendaciones se encuentra:

- Integridad de la información: Abordar el problema de la desinformación y la manipulación de los datos en el contexto de la IA generativa, evitando la generación de información falsa o engañosa.
- Transparencia y la divulgación responsable: Informar adecuadamente sobre los objetivos, las características, las limitaciones y los riesgos de los sistemas de IA.

Para seguir el rastro de decisiones de un sistema de IA, es importante almacenar las diferentes versiones del modelo y los cambios realizados durante la evaluación, así como monitorear las decisiones tomadas por los evaluadores. Esta implementación permite no solo identificar el origen de posibles errores o sesgos, sino también facilitar auditorías independientes que validen el funcionamiento de estos sistemas.

Comparación con métodos tradicionales de evaluación

Los métodos tradicionales de evaluación están basados en principios establecidos, como la capacidad de obtener resultados similares al repetir un procedimiento bajo las mismas condiciones. En comparación con los sistemas basados en IA, los métodos tradicionales, han sido refinados durante y cuentan con marcos teóricos sólidos que facilitan su aplicación sistemática.

La evaluación de resultados generados por IA presenta desafíos únicos, pues las diferentes herramientas incorporan elementos de aleatoriedad que pueden producir resultados diferentes en cada ejecución, dificultando la reproducibilidad exacta, además el rendimiento de los sistemas de IA puede variar dependiendo del contexto de aplicación, lo que complica la generalización de las evaluaciones. A pesar de esto, las evaluaciones que se apoyan en inteligencia artificial, tienen la capacidad de brindar un panorama mucho más amplio en relación a las causas y consecuencias del programa, así como el análisis de la lógica y sus indicadores, que como se mencionaba anteriormente, este análisis se realiza en un tiempo mucho más corto, y con menor recurso económico y humano.

Se debe reconocer que los métodos tradicionales de evaluación y los nuevos enfoques para IA no son mutuamente excluyentes, sino complementarios, pues los principios fundamentales de validez, fiabilidad y reproducibilidad siguen siendo relevantes, pero requieren adaptaciones. La combinación de métodos clásicos con técnicas de evaluación específicas para IA puede proporcionar una visión más holística del rendimiento y la confiabilidad de estos sistemas.

3.4 Limitaciones de modelos de IA

A pesar del potencial transformador del uso de herramientas basadas en IA para la evaluación de programas presupuestarios, es importante reconocer las limitaciones inherentes de la IA para lograr implementar estas tecnologías de manera efectiva y obtener los resultados que se esperan.

IA generativa vs. Modelos predictivos: Utilidad comparada

Ambos enfoques ofrecen ventajas diferenciales para la evaluación de programas:

Los modelos predictivos, se destacan por su capacidad para identificar patrones en datos históricos y proyectar tendencias futuras con un margen de error cuantificable; para la evaluación de programas, estos modelos resultan valiosos para estimar impactos, detectar desviaciones respecto a tendencias esperadas, establecer umbrales óptimos de trabajo, entre otros.

Su mayor fortaleza radica en la precisión y capacidad para establecer relaciones causales entre variables, lo que facilita la toma de decisiones basadas en evidencia. Sin embargo, estos modelos suelen limitarse a análisis estructurados dentro de parámetros predefinidos, además de que suponen un alto costo de operatividad.

Por otro lado, la IA generativa, como los LLM, ofrece capacidades complementarias que se mencionaron anteriormente, como la generación de informes, la síntesis de grandes volúmenes de información para identificar patrones no evidentes, la formulación de hipótesis alternativas, o el análisis de documentos presupuestarios no estructurados para extraer información relevante para la evaluación.

Adaptabilidad de modelos de IA a metodologías preexistentes: el caso de la Metodología de Marco Lógico

La Metodología de Marco Lógico (MML) constituye un enfoque estructurado utilizado en la planificación y evaluación de programas presupuestarios, su integración con herramientas de IA enfrenta obstáculos metodológicos importantes, siendo la capacidad de análisis cualitativo el principal de ellos, pues los supuestos y elementos contextuales de la MML tienen naturaleza cualitativa, presentando desafíos para su incorporación en modelos de IA que requieren insumos cuantificables.

Es importante mencionar que la IA constituye una herramienta complementaria, no sustitutiva, del proceso de evaluación por parte del evaluador; el mayor potencial de la IA reside en amplificar la capacidad analítica, permitiendo fundamentar las decisiones de los evaluadores en evidencia más robusta y análisis más comprehensivos.

4. Comparación de Metodologías para la Implementación de IA en Evaluaciones Presupuestarias

4.1 Elementos centrales para introducir en la IA

La correcta implementación de la Inteligencia Artificial (IA) para la evaluación de programas presupuestarios, requiere de un enfoque estructurado que garantice resultados alineados con los objetivos que se pretenden alcanzar. Los dos elementos requeridos para un completo entendimiento por parte de la IA en relación a lo que se busca obtener, son la Metodología del Marco de Medición Lógica (MML) y los Términos de Referencia (TdR).

La MML constituye una herramienta que permite estructurar los componentes clave de un programa, a través de la cual se puede examinar el diseño de un programa, así como su desempeño en sus diferentes etapas (Coneval, 2012). Al introducir las características de esta metodología en herramientas como los LLM, se logra obtener un análisis del programa que se busca evaluar, poniendo énfasis en la identificación del problema y objetivo del programa, la correcta estructuración de los árboles del problema y objetivos, la coherencia vertical y horizontal de la MIR, por mencionar algunos, logrando así identificar inconsistencias lógicas, redundancias o vacíos en la estructura programática con base a los requerimientos específicos que dicta la MML.

Por otro lado, los TdR, que deben entenderse como los criterios técnicos que deben cumplir las evaluaciones de acuerdo a su tipo (Transparencia, 2024), constituyen el marco metodológico, al cual deben apegarse todos los evaluadores que orienta las evaluaciones de los programas presupuestarios, garantizando que estas sean imparciales y orientadas a resultados concretos. Al incorporar parámetros específicos de los TdR a las herramientas de IA, se logra estandarizar el proceso evaluativo, mientras se especifican los criterios para cada tipo de evolución.

Como se observa, ambos elementos proporcionan un marco estructural que proporciona unanimidad en el proceso de evaluación, transformando así la implementación de IA de un proceso aleatorio, a una iniciativa con resultados medibles y alineados con los objetivos de la evaluación de programas en la Administración Pública.

4.2 Selección de Programas presupuestarios: B001 y P002

La selección de los programas presupuestarios B001 (Producción de impresos valorados, no valorados, numerados y de seguridad) y P002 (Diseño y conducción de la política de gasto público) de Ramo 6 Hacienda y Crédito Público para este proyecto de investigación, responde a criterios estratégicos y metodológicos pues ambos programas fueron evaluados en 2023 y 2024, respectivamente, por la subdirección de evaluación de la Dirección General de Recursos Financieros de las Secretaría de Hacienda y Crédito Público.

Bajo este contexto, se tiene un punto de comparación entre las definiciones previas a la evaluación, las redefinidas por la instancia evaluadora, y las proporcionadas por la IA, permitiendo valorar la calidad de éstas últimas, para valorar su futura aplicación en la evaluación de programas presupuestarios.

4.3 Metodologías de implementación

La Inteligencia Artificial hoy en día, está condicionada a obtener muchos datos como entradas (big data) y de aprender a predecir un resultado gracias al constante entrenamiento y pruebas realizadas con anterioridad.

De acuerdo con Alex Pentland, la magia del Big Data y de la Inteligencia Artificial, es su función de asignación de crédito, entendido como la capacidad de asignar características y combinación de características, que le permiten a la máquina obtener las respuestas correctas y acertadas.

Otra manera de generar valor agregado a los datos, es mediante la aplicación del modelo para explotación de datos, *Cross-industry Standard Process for Data Mining* (CRISP-DM), con una metodología que incluye descripciones de las fases de un proyecto, las tareas necesarias en cada fase y una explicación de las relaciones entre las tareas, para desarrollar modelos analíticos.

El modelo CRISP-DM tiene una naturaleza cíclica pues es un proceso constante de aprendizaje y más cuando se requiere que la explotación de datos arroje resultados estables y replicables. Este modelo se divide en 6 fases: Entender el problema (preparación), comprensión de datos, preparar los datos, modelamiento, evaluación e implementación.

El modelo de CRISP-DM es flexible y se pueden personalizar fácilmente, donde el trabajo se centra en explorar y visualizar datos para descubrir patrones y crear un modelo que se adapte a necesidades concretas. De esta manera, las fases de modelamiento, evaluación e implementación, pueden ser menos relevantes que las fases de preparación y comprensión de datos.

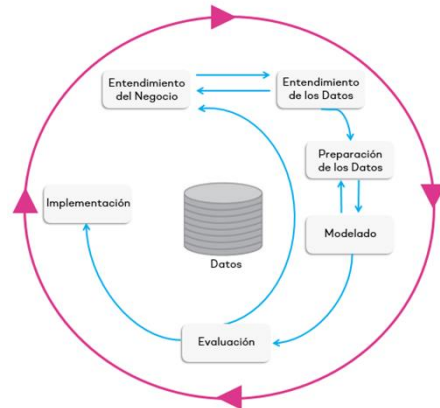
Los chatbots basados en inteligencia artificial tales como ChatGPT, Copilot o Gemini, utiliza Aprendizaje por Refuerzo a partir de la Retroalimentación Humana (RLHF por sus siglas en inglés), una categoría de técnica de aprendizaje automático basado en minería de datos, utilizada para mejorar las capacidades conversacionales; estos chatbots están entrenados para utilizar RLHF a través de un ciclo de retroalimentación de capacitadores manuales, entrenado para separar una respuesta correcta de una incorrecta.

En el área de evaluación, para poder utilizar herramientas basadas en inteligencia artificial, es necesario un enfoque estructurado que permita analizar las preguntas relacionadas con la evaluación de los programas, basado en la Metodología del Marco Lógico.

Para este proyecto se retomaron las fases que comprende el CRISP-DM para ser aplicado en las herramientas de LLM y obtener las respuestas lo más cercano y congruente a lo que se está buscando en materia de evaluación.

1. **Comprensión del proyecto:** Se definió el objetivo de la evaluación y el papel que juega la IA en ésta.
2. **Compresión de los datos (información):** Se definió la alineación de cada una de las partes que se busca evaluar con la metodología del marco lógico.

3. **Preparación de los datos (información):** Se agregó la información a un documento para ser adjuntado y leído por la IA.
4. **Modelado:** Se modificó la información de forma que fuera sencilla de comprender.
5. **Evaluación:** Se evaluó si las respuestas dadas por la IA están alineadas con lo que se busca obtener.
6. **Implementación:** Se requirió modificar el documento en cinco ocasiones hasta que las respuestas fueran congruentes, acertadas y efectivas, permitiendo así tener un modelo que pudiera ser replicado en diversas evaluaciones.



4.4 Selección de los elementos

4.4.1 Selección de los TdR

La optimización en la ejecución de las evaluaciones conforme a los Términos de Referencia (TdR); mediante la automatización de tareas operativas tales como el análisis de documentos programáticos, o la estructuración preliminar de los hallazgos con base a los requerimientos del evaluador, permite concentrar los esfuerzos en actividades de mayor valor analítico, asegurando el cumplimiento de los entregables con base en los criterios de calidad.

Los criterios para determinar aplicabilidad:

- 1) La implementación implica una cantidad de información para la cual las herramientas de uso gratuito son insuficientes.
- 2) La implementación implica la utilización de un formato para la cual las herramientas de uso gratuito son incompatibles.
- 3) La información requerida para contestar la pregunta no requiere de ningún tipo de análisis en el cual la implementación de IA presente una ventaja.
- 4) El análisis necesario para responder la pregunta proviene de un ejercicio colaborativo y de acumulación de resultados donde las herramientas de uso gratuito no arrojan resultados adecuados.

Tabla 1. Porcentaje de implementación de IA en los TdR para la evaluación de Diseño

Sección	Preguntas por Sección	Total de Incisos por Sección	Número de incisos donde se puede implementar IA	Porcentaje de la sección donde se puede implementar IA
I	N/A	0	0	0%
II	5	12	10	83%
III	5	8	7	88%
IV	10	32	2	6%

V	2	0	0	50%*
VI	2	0	0	100%*
VII	1	0	0	0%
VIII	4	13	5	38%
IX	N/A	0	0	0%
X	N/A	0	0	0%
XI	N/A	0	0	0%
Total		65	24	37%

*Si bien el TdR no contempla incisos que permitan asignar un porcentaje, la redacción de la pregunta y los criterios asociados permiten una valoración.

Con base los TdR, se analizaron las once secciones que comprenden este documento con la finalidad de obtener una estimación en relación al porcentaje de las preguntas donde se puede implementar la IA. Como se observa en la tabla, el 37% del TdR, puede ser respondido por herramientas basadas en IA, lo que corresponde a 12 de las 29 preguntas.

Tabla 2. Relación de preguntas de los TdR donde se puede implementar IA

Sección	Pregunta	
II	2	¿El problema o necesidad pública que busca atender el Pp cuenta con las características siguientes?
II	3	¿El problema o necesidad pública que busca atender el Pp se sustenta en los siguientes elementos?
II	5	¿En el diseño del programa se identifican las causas y efectos del problema o necesidad pública que atiende, considerando las siguientes características?
III	6	¿El objetivo central del Pp cuenta con las características señaladas a continuación?
III	8	¿Las poblaciones potencial, objetivo y atendida del Pp se encuentran correctamente identificadas?
V	21	¿La modalidad presupuestaria del Pp es consistente con el objetivo que éste persigue, con los bienes y/o servicios que genera, con sus actividades sustantivas y, en conjunto, con su mecanismo de intervención?
VI	23	¿El objetivo central del Pp contribuye al cumplimiento de alguno de los objetivos o estrategias que se definen en los programas que se derivan del Plan Nacional de Desarrollo (PND) vigente?
VI	24	¿El objetivo central del Pp se vincula con los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030?
VIII	26	¿El ISD vigente del Pp permite obtener información relevante sobre los siguientes elementos de diseño del Pp?
VIII	27	¿Los indicadores que integran el ISD vigente del Pp cumplen con los siguientes criterios?
VIII	29	¿Las metas de los indicadores que integran el ISD vigente del Pp, cumplen con los siguientes criterios?
X	N/A	Análisis FODA
Anexos	2	Árbol del Problema
Anexos	3	Árbol de Objetivos
Anexos	10	Alineación a los ODS
Anexos	12	Instrumentos de Seguimiento del Desempeño

4.4.2 Selección de las herramientas de IA

Para analizar cuáles son las mejores herramientas, se realizaron múltiples pruebas, donde se le solicitó a las diferentes herramientas identificar las deficiencias en la definición de la población potencial, población objetivo, y definición del problema.

En la siguiente tabla se puede observar que, de las seis herramientas consultadas, cuatro identificaron uno o más errores en la definición de la población objetivo y el problema. La herramienta de Copilot, desarrollada por Microsoft, fue la única que consideró que tanto la población, como el problema, estaban correctamente definidos y no necesitaban ser corregidos.

Tabla 3. Relación de criterios cumplidos por herramienta de IA

	Población Potencial	Población Objetivo	Problema
ChatGPT	✓	✗	✗
Claude	✓	✗	✗
DeepSeek	✓	✗	✗
Perplexity	✓	✗	✗
Gemini	✓	✓	✗
Copilot	✓	✓	✓

✓ Criterio que la IA considera como cumplido / ✗ Criterio que la IA considera como no cumplido

Esta segunda tabla muestra los criterios que se le solicitó a la IA para considerar si el problema público estaba correctamente definido. ChatGPT y Claude fueron las que tomaron en consideración todos los criterios descritos en el documento; a pesar de que Copilot consideró cuatro de los cinco criterios, señaló que no existía ninguna deficiencia en la definición del problema.

Perplexity y Gemini únicamente tomaron dos de los cinco criterios, y DeepSeek, a pesar de considerar 2 criterios, mencionó que la definición del problema únicamente no cumplía con la formulación en negativo y estaba formulado como la ausencia de un bien o servicio.

	Formulado como un hecho negativo	No formulado como ausencia de un bien o servicio	Sin enunciados complejos o ambiguos	No se debe mencionar más de una problemática	Sintaxis: "Población objetivo + problema público"
ChatGPT	✗	✗	✗	✓	✗
Claude	✗	✓	✓	✓	✓
DeepSeek	✗	✗			
Perplexity	✗				
Gemini	✗				
Copilot	✓		✓	✓	✓

✓ Criterio que la IA considera como cumplido / ✗ Criterio que la IA considera como no cumplido / Información no generada por la IA

4.5 Evaluación Comparativa de Metodologías en Aplicaciones de IA

4.5.1 Escenario sin conocimiento de los elementos centrales para la evaluación

En este apartado, se solicitó a la IA responder las preguntas de los TdR's identificados en la sección 4.4.1, viables para ser respondidos mediante IA.

4.5.1.1 Análisis programa presupuestario B001

Prompt: “Con base en el documento “Información B001”, responde las preguntas e incisos de estas, del documento “TdR Diseño””.

Chat GPT

Pregunta 2: Claridad del problema público

Valoración: Parcialmente cumple.

- **Definición del problema:** Está formulado como una afirmación negativa (“documentos... susceptibles de ser falsificados y alterados”), sin embargo, se trata de una descripción amplia que agrupa distintas problemáticas: falsificación, vulneración y manejo inadecuado.
- **Carácter único:** No cumple totalmente. Se identifican varias dimensiones del problema (identidad, evasión fiscal, salud pública, etc.).
- **Población afectada:** Mencionada de forma institucional (Dependencias y Entidades de la APF), sin análisis desagregado.
- **Cambio esperado:** No se explicita el resultado en términos del efecto sobre la población, sino la solución directa (documentos seguros).

Recomendación: Reformular el problema como: *“Las Dependencias y Entidades de la APF enfrentan riesgos operativos y legales debido al uso de documentos oficiales vulnerables a falsificación y alteración, afectando la certeza jurídica, la recaudación fiscal y la confianza ciudadana”*.

Esto permitiría focalizar la intervención en el cambio deseado y delimitar mejor la población afectada.

Pregunta 3: Sustento del problema

Valoración: No cumple.

- **Antecedentes:** No se mencionan acciones previas ni su efectividad.
- **Contextualización:** No se apoya en estadísticas oficiales, ni se presentan evidencias sobre la incidencia del problema (falsificaciones detectadas, daños estimados).
- **Evolución:** No se evidencia si el problema ha aumentado o cambiado.
- **Territorialización:** No aplica diferenciación geográfica o por tipo de dependencia.

Recomendación: Incluir estadísticas del SAT, SRE, PROFECO u otras fuentes institucionales que documenten el número de incidentes de falsificación. Justificar la ausencia de caracterización territorial con base en la centralización del proceso.

Pregunta 5: Árbol de causas y efectos

Valoración: Parcialmente cumple.

- **Causas:** Se alude a condiciones técnicas (falta de medidas de seguridad, maquinaria especializada).
- **Efectos:** Se mencionan impactos relevantes (robo de identidad, evasión fiscal, afectaciones a la salud), sin una lógica causal formalizada.
- **Diferenciación:** No hay análisis por grupo poblacional o territorial.

Recomendación: Vincular no las causas técnicas (falta de resguardo, tecnología obsoleta) con efectos sistémicos. Justificar la no diferenciación poblacional por el carácter transversal del Pp.

Pregunta 6: Objetivo central del Pp

Valoración: Cumple.

- **Población objetivo:** Correctamente identificada como las Dependencias y Entidades de la APF.
- **Cambio esperado:** Claramente formulado: uso de documentos seguros y confiables.
- **Unidad:** El objetivo es único y se alinea con la necesidad pública.
- **Consistencia:** Coherente con el problema identificado.

Recomendación: Incluir una referencia explícita al efecto esperado sobre el desempeño institucional y la legalidad.

Pregunta 8: Identificación de poblaciones

Valoración: Cumple.

- **Población potencial y objetivo:** Ambas coinciden (20 dependencias y 210 entidades), están bien delimitadas y cuantificadas.
- **Población atendida:** No está especificada directamente, pero puede inferirse que corresponde a la población objetivo en función del cumplimiento del componente.

Recomendación: Incorporar información histórica del número de entidades efectivamente atendidas por año para seguimiento y análisis de cobertura.

Pregunta 21: Modalidad presupuestaria

Valoración: Cumple.

- **Justificación:** La modalidad "B" (Provisión de Bienes Públicos) es congruente con el objetivo del programa, ya que su función principal es la producción de impresos con medidas de seguridad.

Pregunta 26: ISD y procesos operativos

Valoración: Cumple.

- El ISD cubre los procesos clave: adquisición de insumos, mantenimiento de maquinaria, producción, entrega, y satisfacción del cliente. La trazabilidad de los procesos es clara y alineada al propósito del programa.

Pregunta 27: Indicadores del ISD

Valoración: Cumple con áreas de mejora.

- **Claridad:** Los nombres de indicadores son entendibles.
- **Relevancia y adecuación:** Reflejan componentes clave del desempeño.
- **Monitoreo:** Son medibles y trazables.
- **Eficiencia:** No se documenta el costo de recolección, aunque es esperable que la información esté disponible operativamente.

Recomendación: Complementar con fichas técnicas que documenten metodología, periodicidad, fuentes y responsables de cálculo.

Pregunta 29: Metas de los indicadores

Valoración: Parcialmente cumple.

- Las fórmulas están documentadas, pero no se especifican metas cuantitativas ni justificación metodológica.
- No hay evidencia de análisis de factibilidad ni de orientación hacia mejora progresiva.

Recomendación: Establecer metas anuales con base en históricos, capacidad instalada y mejoras esperadas, en congruencia con principios de mejora continua.

DeepSeek

2. ¿El problema o necesidad pública que busca atender el Pp cuenta con las características siguientes?

a) Claridad y unicidad del problema: Sí, el problema se define de manera clara y concreta: la falsificación y vulneración de documentos oficiales utilizados por las Dependencias y Entidades de la APF, lo que genera delitos como robo de identidad, evasión fiscal y daños a la salud. No se identifican múltiples problemáticas, sino una central con consecuencias específicas.

b) Formulación como hecho negativo o situación reversible: El problema se formula como un hecho negativo (falsificación y vulneración de documentos) que puede ser revertido mediante la producción de impresos seguros y confiables.

c) Población afectada identificada: La población afectada se identifica claramente como las 20 Dependencias y 210 Entidades de la APF que utilizan documentos oficiales susceptibles de falsificación.

d) Cambio esperado: El cambio esperado no solo es la ausencia de falsificación, sino la reducción de riesgos mediante la provisión de documentos seguros y confiables, lo que impacta positivamente en la sociedad y el Estado.

3. ¿El problema o necesidad pública que busca atender el Pp se sustenta en los siguientes elementos?

a) Antecedentes de acciones: El documento menciona que la impresión de documentos oficiales ya se realiza con maquinaria e insumos especializados, lo que indica antecedentes de acciones para mitigar el problema.

b) Contexto socioeconómico: El problema se contextualiza en la situación actual del país, donde la falsificación de documentos facilita delitos como evasión fiscal y fraudes, afectando la economía y la seguridad pública.

c) Consistencia con cambios en la problemática: No se proporcionan estadísticas específicas, pero se describe cómo la falsificación ha evolucionado y cómo el Pp busca contrarrestarla con medidas de seguridad.

d) Caracterización territorial: El problema se enfoca en el ámbito federal, abarcando a todas las Dependencias y Entidades de la APF, lo que delimita su alcance territorial.

5. ¿En el diseño del programa se identifican las causas y efectos del problema?

a) Causas principales: Las causas incluyen la falta de medidas de seguridad adecuadas en documentos oficiales y la accesibilidad a técnicas de falsificación.

b) Efectos principales: Los efectos son delitos como robo de identidad, evasión fiscal y daños a la salud, que perjudican a la sociedad y al Estado.

c) Lógica causal: Existe una relación clara entre las causas (falsificación) y los efectos (delitos), y el Pp aborda esta cadena causal mediante la producción de impresos seguros.

d) Análisis diferenciado: No se mencionan diferencias por grupos de población o territorios, ya que el problema es homogéneo en el ámbito federal.

6. ¿El objetivo central del Pp cuenta con las características señaladas?

a) Población objetivo identificada: Sí, son las 20 Dependencias y 210 Entidades de la APF.

b) Cambio esperado: El objetivo busca que los documentos sean seguros y confiables, reduciendo los riesgos de falsificación.

c) Unicidad del objetivo: El objetivo es único: garantizar la seguridad de los documentos oficiales.

d) Correspondencia con la solución: El objetivo está alineado con la solución de producir impresos con medidas de seguridad.

8. ¿Las poblaciones potencial, objetivo y atendida del Pp se encuentran correctamente identificadas?

a) Población potencial: Sí, son las 20 Dependencias y 210 Entidades de la APF.

b) Población objetivo: Coincide con la población potencial, ya que todas son elegibles para recibir los impresos seguros.

c) Población atendida: No se especifica en el documento, pero se asume que es un subconjunto de la población objetivo en un ejercicio fiscal.

d) Consistencia entre documentos: No se mencionan inconsistencias, por lo que se asume que las poblaciones son consistentes en los documentos estratégicos.

21. ¿La modalidad presupuestaria del Pp es consistente con el objetivo que éste persigue?

La modalidad del Pp es "Provisión de Bienes Públicos" (Modalidad B), lo cual es consistente con su objetivo de producir y suministrar impresos seguros para las Dependencias y Entidades de la APF. Esta modalidad se ajusta a las actividades sustantivas del programa.

26. ¿El ISD vigente del Pp permite obtener información relevante sobre los elementos de diseño del Pp?

El ISD (Instrumento de Seguimiento del Desempeño) incluye indicadores como el porcentaje de eficiencia de los procesos productivos y el porcentaje de satisfacción del cliente, los cuales son relevantes para monitorear la gestión de recursos y el cumplimiento del propósito del Pp.

27. ¿Los indicadores que integran el ISD vigente del Pp cumplen con los siguientes criterios?

a) Claridad: Los indicadores son entendibles, como "Porcentaje de satisfacción del cliente".

b) Relevancia: Proveen información valiosa, como la eficiencia de los procesos y la calidad de los impresos.

c) Economía: No se mencionan costos, pero se asume que la información es accesible.

d) Monitoreabilidad: Los métodos de cálculo están documentados, lo que permite su verificación.

e) Adecuación: Los indicadores miden adecuadamente el desempeño del Pp.

29. ¿Las metas de los indicadores que integran el ISD vigente del Pp cumplen con los siguientes criterios?

a) Método de cálculo documentado: Sí, cada indicador incluye una fórmula clara.

b) Unidad de medida congruente: Las unidades (porcentajes) son congruentes con los indicadores.

c) Orientación a la mejora: Las metas buscan mejorar la eficiencia y satisfacción, no son laxas.

d) Factibilidad: Se consideran los recursos disponibles, como maquinaria y materias primas.

Conclusión:

El programa cumple con la mayoría de los criterios evaluados. Las áreas de mejora podrían incluir la incorporación de estadísticas oficiales para sustentar el contexto socioeconómico y un análisis más detallado de la población atendida en ejercicios fiscales específicos.

Claude

Basándome en el documento "Información B001", responderé las preguntas del documento "TdR Diseño":

Pregunta 2: ¿El problema o necesidad pública que busca atender el Pp cuenta con las siguientes características?

a) Se define de manera clara, concreta, acotada y es único ✓ Cumple. El problema se define como: "Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, susceptibles de ser falsificados y alterados." Esta definición es clara, concreta, acotada y presenta una única problemática.

b) Se formula como un hecho negativo o como una situación que puede ser revertida ✓ Cumple. El problema se expresa como un hecho negativo (documentos susceptibles de falsificación y alteración) que puede ser revertido mediante la producción de documentos seguros.

c) Identifica a la población que enfrenta la problemática de manera clara, concreta y delimitada ✓ Cumple. La población está claramente identificada como "Las Dependencias y Entidades de la Administración Pública Federal".

d) Identifica un cambio (resultado) sobre la población que enfrenta la problemática ✓ Cumple parcialmente. El problema identifica la situación negativa (vulnerabilidad de documentos), pero podría fortalecer la identificación del cambio esperado sobre cómo afecta esto a las dependencias y entidades, más allá de la ausencia de seguridad en los documentos.

Pregunta 3: ¿El problema o necesidad pública que busca atender el Pp se sustenta en los siguientes elementos?

a) Considera los antecedentes de acciones que se hayan realizado para su atención ✓ Cumple parcialmente. El documento menciona que existen medidas de seguridad, pero no detalla antecedentes específicos de acciones previas para atender la falsificación.

b) Se encuentra contextualizado en la situación socioeconómica actual del país ✓ Cumple parcialmente. Menciona consecuencias como evasión fiscal, daños a la salud, etc., pero no proporciona datos estadísticos específicos sobre la magnitud del problema.

c) Es consistente con los cambios que ha experimentado la problemática según información oficial X No cumple. El documento no presenta información o estadísticas que muestren la evolución de la problemática a lo largo del tiempo.

d) Se encuentra caracterizado en un contexto territorial ✓ Cumple. Se delimita al ámbito nacional (México) y menciona explícitamente que "En México, la adquisición de maquinaria y de materia prima especializada para estos fines está restringida..."

Pregunta 5: ¿En el diseño del programa se identifican las causas y efectos del problema o necesidad pública que atiende?

a) Es resultado directo de las principales causas identificadas ✓ Cumple parcialmente. Se mencionan algunas causas implícitas como la vulnerabilidad de documentos por falta de elementos de seguridad, pero no hay un análisis estructurado de causas.

b) Se vincula directamente con los principales efectos identificados ✓ Cumple. Se identifican efectos como: "robo de identidad, evasión fiscal, daños a la salud en los casos de bebidas alcohólicas adulteradas, usurpación de profesiones, fraudes financieros, entre otros."

c) Existe lógica causal entre sus causas y efectos ✓ Cumple parcialmente. Existe una lógica implícita, pero no se presenta un análisis estructurado de la cadena causal completa.

d) Presenta un análisis diferenciado por grupos de población, territorios o medio ambiente X No cumple. No se presenta un análisis diferenciado por grupos de población o territorios específicos.

Pregunta 6: ¿El objetivo central del Pp cuenta con las características señaladas?

a) Identifica a la población objetivo ✓ Cumple. El objetivo identifica claramente a "Las Dependencias y Entidades de la Administración Pública Federal" como población objetivo.

b) Identifica el cambio que el Pp busca generar ✓ Cumple. El cambio es que las dependencias "utilizan documentos probatorios de identidad, de carácter fiscal y social, seguros y confidenciales que permiten reducir los riesgos de falsificación y alteración."

c) Es único ✓ Cumple. El objetivo es único y no define múltiples objetivos.

d) Corresponde a la solución del problema ✓ Cumple. El objetivo responde directamente al problema identificado, buscando revertir la vulnerabilidad de los documentos oficiales.

Pregunta 8: ¿Las poblaciones potencial, objetivo y atendida del Pp se encuentran correctamente identificadas?

a) Delimita y cuantifica la población potencial ✓ Cumple. Se identifica y cuantifica como "20 Dependencias y 210 Entidades de la Administración Pública Federal".

b) Delimita y cuantifica la población objetivo ✓ Cumple. Se identifica y cuantifica igual que la población potencial: "20 Dependencias y 210 Entidades de la Administración Pública Federal".

c) Delimita y cuantifica la población atendida X No cumple. El documento no proporciona información específica sobre la población atendida en un ejercicio fiscal concreto.

d) Las poblaciones son consistentes entre diversos documentos estratégicos ✓ Cumple parcialmente. Dentro del documento analizado, la definición de población potencial y objetivo es consistente, pero no se puede verificar la consistencia con otros documentos estratégicos no proporcionados.

Pregunta 21: ¿La modalidad presupuestaria del Pp es consistente con el objetivo?

✓ Cumple. El Pp está clasificado en la Modalidad B "Provisión de Bienes Públicos", lo cual es consistente con su objetivo de producir y suministrar impresos con elementos y medidas de seguridad para las dependencias de la APF. El programa se enfoca en crear, fabricar y elaborar bienes (impresos de seguridad) que son competencia del Sector Público Federal.

Pregunta 26: ¿El ISD vigente del Pp permite obtener información relevante sobre los elementos de diseño del Pp?

✓ Cumple parcialmente. La MIR contiene indicadores relacionados con los principales procesos operativos como la adquisición de materias primas, insumos y refacciones (actividad A1) y la ejecución del programa de mantenimiento preventivo (actividad B2). Sin embargo, no incluye indicadores sobre otros procesos relevantes como la recepción, registro y supervisión de solicitudes de impresión.

Pregunta 27: ¿Los indicadores que integran el ISD vigente del Pp cumplen con los criterios?

a) Es claro ✓ Cumple parcialmente. Algunos indicadores son claros, como "Porcentaje de satisfacción del cliente", pero otros como "Porcentaje de eficiencia de los procesos productivos" podrían resultar ambiguos sin mayor contexto.

b) Es relevante ✓ Cumple. Los indicadores proveen información valiosa sobre la producción, satisfacción y eficiencia, que son aspectos centrales del programa.

c) Es económico ✓ Cumple. La información para calcular los indicadores parece estar disponible como parte de la operación normal del programa.

d) Es monitoreable ✓ Cumple. Los indicadores tienen fórmulas de cálculo definidas que permiten su verificación independiente.

e) Es adecuado ✓ Cumple parcialmente. Los indicadores abordan aspectos importantes del programa, pero podrían complementarse con otros que midan la calidad de los elementos de seguridad y la efectividad para prevenir falsificaciones.

Pregunta 29: ¿Las metas de los indicadores que integran el ISD vigente del Pp cumplen con los criterios?

a) Se establecen con base en un método de cálculo documentado ✓ Cumple. Para cada indicador se define un método de cálculo claro y documentado.

b) Cuentan con unidad de medida y son congruentes con el sentido del indicador ✓ Cumple. Los indicadores muestran porcentajes como unidad de medida, lo cual es coherente con la naturaleza de los métodos de cálculo.

c) Se orientan a la mejora del desempeño X No cumple. No se proporciona información sobre los valores de las metas para determinar si son ambiciosas o laxas.

d) Son factibles ✓ Cumple parcialmente. Los indicadores parecen factibles de medir con los recursos disponibles, pero no se proporciona información sobre plazos y recursos específicos para su cumplimiento.

4.5.1.2 Análisis programa presupuestario P002

Prompt: "Con base en el documento "Información P002", responde las preguntas e incisos de estas, del documento "TdR Diseño".

Chat GPT

Pregunta 2: Claridad del problema público

Valoración: Parcialmente cumple.

- **Definición clara y concreta:** El problema se formula de manera general como una necesidad de asignaciones presupuestarias sujetas a criterios de eficiencia y eficacia. No obstante, se presenta como un hecho negativo, aunque no se delimita con suficiente precisión ni se describe como único, dado que integra múltiples causas subyacentes (ineficiencias normativas, desarticulación, etc.).
- **Población afectada:** Se identifica de forma general como los ejecutores del gasto, aunque no se caracteriza de forma diferenciada.
- **Cambio esperado:** No se expresa explícitamente como un resultado sobre la población afectada.

Recomendación: Reformular el problema como: *“Los ejecutores del gasto de la APF, Poderes, Entes Autónomos y EPE enfrentan restricciones para acceder a asignaciones presupuestarias alineadas con criterios de eficiencia y eficacia, lo que limita su contribución al cumplimiento de los objetivos de desarrollo nacional”.*

Pregunta 3: Sustento del problema

Valoración: Parcialmente cumple.

- **Antecedentes:** No se presenta un análisis histórico de acciones pasadas.
- **Contextualización socioeconómica:** No se ancla con datos oficiales actuales (ej. evolución del gasto, brechas en eficiencia).
- **Caracterización territorial:** No se incorpora enfoque territorial.
- **Evolución del problema:** No se documenta.

Recomendación: Integrar un análisis contextual utilizando información del CONEVAL, INEGI o el SHCP para evidenciar la persistencia y evolución de los problemas de eficiencia del gasto. Incorporar contexto regional en caso de existir diferencias entre tipos de entes o zonas geográficas.

Pregunta 5: Árbol de problemas (causas y efectos)

Valoración: Cumple con áreas de mejora.

- **Causas y efectos:** Se identifican múltiples causas y efectos, y se establece una lógica causal general.
- **Diferenciación por grupos:** No se realiza análisis diferenciado por población, territorio o medio ambiente.

Recomendación: Integrar justificación explícita de por qué no aplica una diferenciación por grupos poblacionales o territoriales, considerando que el Pp es de naturaleza normativa y transversal.

Pregunta 6: Objetivo central

Valoración: Cumple.

- **Población objetivo:** Correctamente identificada como ejecutores del gasto.
- **Cambio esperado:** Se define como acceso a asignaciones presupuestarias con criterios de eficiencia.
- **Objetivo único:** Está alineado a la necesidad pública.

Recomendación: Mantener la coherencia con el árbol del problema y la MIR, fortaleciendo la medición del cambio esperado mediante indicadores sensibles al uso eficiente del gasto.

Pregunta 8: Identificación de poblaciones

Valoración: Cumple.

- **Población potencial y objetivo:** Bien delimitadas y cuantificadas (54 ejecutores del gasto), y ambas consistentes entre sí.
- **Población atendida:** No se especifica, se asume implícitamente como coincidente.

Recomendación: Incluir información empírica anual sobre la población atendida para efectos de trazabilidad y monitoreo en el ISD y documentos normativos.

Pregunta 21: Modalidad presupuestaria

Valoración: Cumple.

- **Consistencia:** La modalidad "P" (Planeación, seguimiento y evaluación de políticas públicas) es coherente con los bienes/servicios, objetivos y mecanismo de intervención del Pp P002.

Pregunta 26: ISD y procesos operativos

Valoración: Parcialmente cumple.

- El ISD considera algunos procesos de gestión (formulación del PPEF, adecuaciones presupuestarias, emisión de informes), pero no explicita el desglose completo de gestión operativa.

Recomendación: Ampliar el ISD con actividades de gestión de recursos (humanos, financieros, tecnológicos) y mecanismos de control que soportan el cumplimiento del propósito.

Pregunta 27: Indicadores del ISD

Valoración: Parcialmente cumple.

- **Claridad y relevancia:** Los nombres son claros y reflejan el desempeño, pero no todos cuentan con definiciones operacionales completas.
- **Monitoreo y trazabilidad:** Es posible, pero algunos indicadores presentan dificultades para evaluar de forma independiente su cálculo.
- **Económicos y adecuados:** Generalmente sí, aunque no se documenta el costo ni el esfuerzo de recolección.

Recomendación: Revisar y fortalecer la ficha técnica de cada indicador conforme a la metodología de CONEVAL y la MML.

Pregunta 29: Metas de los indicadores

Valoración: Parcialmente cumple.

- Se presentan fórmulas de cálculo, pero falta información sobre la base metodológica para definir metas, unidades de medida y su congruencia con el sentido del indicador.

Recomendación: Documentar el método de cálculo de metas, especificar su unidad de medida y validar su factibilidad frente a capacidades institucionales.

DeepSeek

Respuestas a los criterios del documento "TdR Diseño" basadas en el documento "Información P002":

2. ¿El problema o necesidad pública que busca atender el Pp cuenta con las siguientes características?

a) Definición clara, concreta y única: Sí. El problema se define como: "Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles". Es claro y acotado, sin múltiples problemáticas.

b) Formulación como hecho negativo o situación reversible: Sí. El problema se plantea como una carencia (falta de asignaciones eficientes) que puede revertirse mediante acciones del programa.

c) Población identificada: Sí. La población es clara: ejecutores del gasto (dependencias, entidades de la APF, Poderes, Entes Autónomos y Empresas Productivas del Estado).

d) Cambio esperado: Sí. El objetivo busca que los ejecutores del gasto "reciban asignaciones presupuestarias sujetas a criterios de eficiencia y eficacia", lo que implica un cambio positivo.

3. ¿El problema se sustenta en los siguientes elementos?

a) Antecedentes de acciones: No se mencionan acciones previas específicas, pero se alude a problemas como desarticulación entre necesidades y recursos, lo que sugiere un historial de ineficiencias.

b) Contexto socioeconómico: Sí. Se menciona la política de austeridad 2019-2024 y la necesidad de equilibrio entre austeridad e inversión estratégica.

c) Consistencia con estadísticas oficiales: No se citan datos específicos, pero se alude a problemas como "ejercicio ineficiente del gasto público", que suelen documentarse en informes gubernamentales.

d) Contexto territorial: No aplica directamente, ya que el problema es nacional y administrativo.

Propuesta de mejora: Incluir datos estadísticos sobre ineficiencias presupuestarias históricas y referencias a informes como el MSD para fortalecer la justificación.

5. ¿En el diseño del programa se identifican causas y efectos del problema?

a) Causas principales: Sí. Se listan causas como desarticulación entre necesidades y recursos, y ambigüedad normativa.

b) Efectos principales: Sí. Efectos como incumplimiento de objetivos nacionales y ejercicio ineficiente del gasto.

c) Lógica causal: Sí. Las causas (ej. desarticulación) explican los efectos (ej. ineficiencia).

d) Análisis diferenciado: No. El problema es homogéneo para todos los ejecutores del gasto.

Propuesta de mejora: Profundizar en cómo afecta a sectores específicos (ej. dependencias vs. entes autónomos).

6. ¿El objetivo central del Pp cumple con las características?

a) Población objetivo identificada: Sí. Los ejecutores del gasto están claramente definidos.

b) Cambio esperado: Sí. Busca que reciban asignaciones eficientes.

c) Objetivo único: Sí. No hay múltiples objetivos.

d) Solución al problema: Sí. El objetivo aborda directamente la falta de eficiencia en asignaciones.

8. ¿Las poblaciones potencial, objetivo y atendida están correctamente identificadas?

a) Población potencial: Sí. Todos los ejecutores del gasto (54 entidades).

b) Población objetivo: Sí. Coincide con la población potencial.

c) Población atendida: No se especifica en el documento.

d) Consistencia entre documentos: No se evalúa por falta de información.

Propuesta de mejora: Incluir datos de población atendida en ejercicios fiscales recientes.

21. ¿La modalidad presupuestaria es consistente con el objetivo?

Sí. La modalidad "P" (Planeación) se alinea con el objetivo de diseñar políticas de gasto eficientes.

26. ¿El ISD vigente permite obtener información relevante sobre procesos operativos?

Sí. La MIR incluye indicadores como porcentaje de gasto programable con desempeño satisfactorio, que monitorean procesos clave.

27. ¿Los indicadores del ISD cumplen con los criterios?

- a) **Claridad:** Sí. Los nombres son entendibles (ej. "Porcentaje del gasto programable")
- b) **Relevancia:** Sí. Miden resultados clave (ej. cumplimiento de metas).
- c) **Economía:** Sí. Usan datos disponibles (ej. MSD).
- d) **Monitoreabilidad:** Sí. Son trazables (ej. fórmulas definidas).
- e) **Adecuación:** Sí. Cubren dimensiones críticas del desempeño.

29. ¿Las metas de los indicadores cumplen con los criterios?

- a) Método de cálculo documentado: Sí. Cada indicador incluye su fórmula.
- b) Unidad de medida congruente: Sí. Usan porcentajes y tasas adecuadas.
- c) Orientación a la mejora: Sí. Las metas exigen desempeño satisfactorio (ej. valoración ≥ 85 en MIR).
- d) Factibilidad: Sí. Consideran recursos y normatividad vigente.

Conclusiones: El Pp P002 cumple en general con los criterios evaluados, pero podría mejorar en la inclusión de datos estadísticos y análisis diferenciado por sectores. La MIR y los indicadores son robustos y alineados con los objetivos del programa.

Claude

Analizaré cada pregunta del documento "Criterios para Prompt" y responderé basándome en la información disponible en "Información P002".

Pregunta 2. ¿El problema o necesidad pública que busca atender el Pp cuenta con las características siguientes?

- a) **Se define de manera clara, concreta, acotada y es único: Cumplimiento parcial.** El problema está definido como: "Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles." Es relativamente claro y único, pero su formulación no es completamente acotada ya que

menciona varios elementos (asignaciones presupuestarias, criterios de eficiencia y eficacia, objetivos del desarrollo nacional).

b) Se formula como un hecho negativo o como una situación que puede ser revertida: No cumple. El problema no se formula claramente como un hecho negativo o situación a revertir. Está redactado como una necesidad ("requieren de asignaciones") más que como una problemática existente. No indica una carencia o situación negativa actual que deba ser modificada.

c) Identifica a la población que enfrenta la problemática de manera clara, concreta y delimitada: Cumple. El problema identifica claramente a "los ejecutores del gasto" como la población que enfrenta la problemática. Posteriormente el documento detalla que estos son: dependencias y entidades de la APF, los Poderes y Entes Autónomos, así como las Empresas Productivas del Estado.

d) Identifica un cambio sobre la población que enfrenta la problemática: No cumple. La definición del problema no identifica claramente un cambio o resultado sobre la población, sino que se define principalmente como la necesidad de un servicio (asignaciones presupuestarias). No se especifica el cambio concreto que se espera lograr en la condición de los ejecutores del gasto.

Pregunta 3. ¿El problema o necesidad pública que busca atender el Pp se sustenta en los siguientes elementos?

a) Considera los antecedentes de acciones que se hayan realizado para su atención: Cumple parcialmente. El documento menciona que el Pp P002 inició operaciones en 2008 y menciona las unidades responsables de su operación. Sin embargo, no detalla acciones específicas anteriores que se hayan realizado para atender la problemática.

b) Se encuentra contextualizado en la situación socioeconómica actual del país: Cumple. El documento contextualiza el problema en el marco de la política de gasto público del Gobierno de México para el periodo 2019-2024, mencionando principios como la disciplina fiscal y medidas de austeridad definidas en el artículo 61 de la LFPRH.

c) Es consistente con los cambios que ha experimentado la problemática: Cumple parcialmente. El documento menciona el reto de la política de gasto que busca el balance entre la austeridad y la inversión en sectores estratégicos, lo que sugiere adaptación a cambios en la problemática. Sin embargo, no proporciona datos específicos sobre cómo ha evolucionado el problema.

d) Se encuentra caracterizado en un contexto territorial: Cumple. El problema está contextualizado a nivel nacional, haciendo referencia al gasto público federal en México y mencionando las directrices establecidas en el PPEF y el PEF, así como el marco normativo aplicable (LOAPF, LFPRH).

Pregunta 5. ¿En el diseño del programa se identifican las causas y efectos del problema o necesidad pública que atiende, considerando las siguientes características?

a) Es resultado directo de las principales causas identificadas: Cumple. El problema es resultado directo de las causas principales identificadas en el árbol del problema: la desarticulación entre necesidades, prioridades y recursos; la necesidad de ajustes en estructuras y calendarios; y la ambigüedad normativa.

b) Se vincula directamente con los principales efectos identificados: Cumple. El problema se vincula directamente con los efectos identificados, principalmente el incumplimiento de los objetivos del desarrollo nacional y el ejercicio ineficiente del gasto público.

c) Existe lógica causal entre sus causas y efectos: Cumple. Hay una lógica causal coherente entre las causas identificadas (como la desarticulación entre necesidades y recursos, ajustes en estructuras y ambigüedad normativa) y los efectos (incumplimiento de objetivos y ejercicio ineficiente del gasto).

d) Presenta un análisis diferenciado de las causas y efectos por grupos de población: No cumple. El documento no presenta un análisis diferenciado de causas y efectos por grupos específicos de ejecutores de gasto (por ejemplo, diferenciando entre dependencias, poderes autónomos y empresas productivas del Estado) ni por territorios.

Pregunta 6. ¿El objetivo central del Pp cuenta con las características señaladas a continuación?

a) Identifica a la población objetivo del Pp: Cumple. El objetivo central identifica claramente a "los ejecutores del gasto" como población objetivo, y el documento especifica que estos incluyen a las dependencias y entidades de la APF, Poderes y Entes Autónomos, y Empresas Productivas del Estado.

b) Identifica el cambio que el Pp busca generar en la población: Cumple parcialmente. El objetivo indica que los ejecutores del gasto "reciben asignaciones presupuestarias sujetas a criterios de eficiencia y eficacia", pero no detalla claramente el cambio específico que esto generará en ellos más allá de recibir las asignaciones.

c) Es único, es decir, no se definen múltiples objetivos: Cumple. El objetivo es único y se centra en una sola finalidad: que los ejecutores del gasto reciban asignaciones presupuestarias con criterios específicos.

d) Corresponde a la solución del problema que origina la acción gubernamental: Cumple. El objetivo corresponde directamente a la solución del problema planteado, ya que aborda la necesidad de asignaciones presupuestarias eficientes y eficaces para los ejecutores del gasto.

Pregunta 8. ¿Las poblaciones potencial, objetivo y atendida del Pp se encuentran correctamente identificadas?

a) El Pp delimita y cuantifica a la población total que presenta el problema: Cumple. El documento delimita claramente la población potencial como "los ejecutores del gasto" que incluyen a las dependencias y entidades de la APF, los Poderes y Entes Autónomos, así como a las Empresas Productivas del Estado.

b) El Pp delimita y cuantifica a la población que tiene planeado atender: Cumple. La población objetivo está definida igual que la potencial, indicando que el programa planea atender a todos los ejecutores del gasto identificados (54 en total).

c) El Pp delimita y cuantifica a la población atendida en un ejercicio fiscal: No se puede determinar. El documento no menciona específicamente la población atendida en un ejercicio fiscal concreto, solo define la población objetivo que se planea atender.

d) Las poblaciones son consistentes entre diversos documentos del programa: Cumple parcialmente. En el documento analizado, las definiciones de población potencial y objetivo son consistentes, pero no se proporciona información sobre otros documentos estratégicos del programa para verificar la consistencia en todos ellos.

Pregunta 21. ¿La modalidad presupuestaria del Pp es consistente con el objetivo que éste persigue, con los bienes y/o servicios que genera, con sus actividades sustantivas y, en conjunto, con su mecanismo de intervención?

Sí, es consistente. El programa P002 tiene la modalidad "P" (Planeación, seguimiento y evaluación de políticas públicas), lo cual es coherente con su objetivo de diseñar y conducir la política de gasto público. El programa se enfoca en integrar el PPEF y atender adecuaciones en materia de programación, presupuesto y seguimiento del gasto público federal, actividades que corresponden a la planeación y conducción de políticas públicas. No se trata de un programa que entregue subsidios o realice actividades de producción directa, sino que se centra en la conducción de la política pública de gasto, lo que es consistente con la clasificación "P".

Pregunta 26. ¿El ISD vigente del Pp permite obtener información relevante sobre los elementos de diseño del Pp?

Parcialmente. La MIR incluye indicadores relacionados con los procesos operativos del programa, particularmente en los componentes: "Cartera de Programas y Proyectos de Inversión", "Información de desempeño valorada", "Informes de Contabilidad Gubernamental e Informes de Gestión" y "MIR con valoración satisfactoria". Sin embargo, no se observan indicadores específicos sobre la gestión de recursos humanos, tecnológicos y financieros internos del Pp, ni sobre procesos de adquisición de insumos o supervisión. Los indicadores se centran más en los resultados que en los procesos de gestión interna.

Pregunta 27. ¿Los indicadores que integran el ISD vigente del Pp cumplen con los siguientes criterios?

a) Es claro: Cumple parcialmente. Algunos indicadores son claros en su denominación (como "Porcentaje del gasto de inversión en infraestructura productiva"), pero otros contienen acrónimos o términos técnicos sin definición clara en el documento (como MSD - Modelo Sintético de Información de Desempeño).

b) Es relevante: Cumple. Los indicadores proporcionan información valiosa sobre aspectos clave del programa, como la eficiencia del gasto, la calidad de la información y el desempeño de los programas presupuestarios.

c) Es económico: No se puede determinar completamente. El documento no proporciona información suficiente sobre el costo de obtención de los datos para cada indicador, aunque varios parecen basarse en información administrativa que probablemente esté disponible como parte de la operación normal.

d) Es monitoreable: Cumple parcialmente. Los indicadores incluyen métodos de cálculo que permiten su estimación, pero para varios de ellos (como los relacionados con el MSD) no se detalla completamente cómo verificar de manera independiente la información.

e) Es adecuado: Cumple. Los indicadores proporcionan información suficiente para medir el desempeño del programa en sus distintos niveles: fin, propósito y componentes.

Pregunta 29. ¿Las metas de los indicadores que integran el ISD vigente del Pp cumplen con los siguientes criterios?

a) Se establecen con base en un método de cálculo documentado: Cumple. Los indicadores presentan métodos de cálculo documentados que permiten establecer metas de manera clara.

b) Cuentan con unidad de medida y son congruentes con el sentido del indicador: Cumple parcialmente. Los indicadores cuentan con métodos de cálculo que implican unidades de medida (principalmente porcentajes), pero el documento no especifica explícitamente el sentido de cada indicador (ascendente o descendente).

c) Se orientan a la mejora del desempeño: No se puede determinar completamente. El documento no proporciona información sobre los valores históricos o metas específicas que permitan evaluar si estas son laxas o ambiciosas.

d) Son factibles: No se puede determinar completamente. Sin conocer los valores específicos de las metas y los recursos disponibles, no es posible evaluar completamente su factibilidad, aunque los métodos de cálculo sugieren que son técnicamente viables.

4.5.2 Escenario con conocimiento

En este apartado se proveyó contexto detallado a la IA en relación a los programas presupuestarios y las evaluaciones, atendiendo de manera puntual la información que se requiere para “Problema Público”, “El Objetivo Central”, la “Población Objetivo”, la “Matriz de Indicadores del Desempeño” y la información que el evaluador considere relevante para cada apartado.

4.5.2.1 Análisis programa presupuestario B001

Prompt: “Con base en el documento “Información Evaluación”, verifica si la información de documento “Información B001” cumple con los parámetros descritos, el análisis debe ser preciso y alineado con los lineamientos de la Metodología del Marco Lógico. Si consideras que no es así, presenta las modificaciones correspondientes”.

Chat GPT

Se presenta a continuación un análisis detallado de la información, evaluando su cumplimiento conforme a la Metodología del Marco Lógico (MML), así como las propuestas de modificación necesarias para su alineación metodológica.

1. Definición del problema

Evaluación:

- **Redacción actual:**
“Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, susceptibles de ser falsificados y alterados.”
- **Cumplimiento:**
 - ✓ Está formulado como un hecho negativo.
 - ✓ La población afectada está identificada.
 - ✗ No presenta el cambio que experimenta la población por la existencia del problema.
 - ✗ No está redactado en la forma recomendada: *“Población objetivo + problema público”*.
 - ✗ Utiliza una descripción del objeto del programa como problema (ausencia de seguridad documental), lo que contradice los lineamientos de MML.

Propuesta de redacción:

“Las Dependencias y Entidades de la Administración Pública Federal enfrentan riesgos operativos y legales debido al uso de documentos oficiales susceptibles de ser falsificados o vulnerados, lo que compromete la seguridad institucional y la confianza ciudadana.”

2. Población potencial y objetivo

Evaluación:

- **Redacción actual:**
20 Dependencias y 210 Entidades de la Administración Pública Federal como población potencial y objetivo.
- **Cumplimiento:**
 - ✓ La población está cuantificada.
 - ✗ No se justifica por qué estas entidades son elegibles.
 - ✗ No se evidencia si esta población corresponde a quienes enfrentan el problema público.

Recomendación:

- Incluir justificación metodológica que delimite la relación entre estas dependencias y el problema identificado (por ejemplo, análisis de riesgo de falsificación documentada por tipo de entidad).
- Alinear los criterios de elegibilidad con las competencias legales y operativas de cada dependencia.

3. Definición del objetivo

Evaluación:

- **Redacción actual:**
“Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, seguros y confidenciales que permiten reducir los riesgos de falsificación y alteración.”
- **Cumplimiento:**
 - ✓ Está vinculado directamente con la solución del problema.
 - ✓ Es único.
 - ✗ No identifica explícitamente el cambio en la población (mejora de seguridad, operatividad, confianza, etc.).
 - ✗ No es cuantificable ni suficientemente concreto.

Propuesta de redacción:

“Las Dependencias y Entidades de la Administración Pública Federal reducen los riesgos operativos y legales mediante el uso de documentos oficiales seguros y confiables, con elementos técnicos de alta seguridad que previenen su falsificación o vulneración.”

4. Matriz de Indicadores para Resultados (MIR)

Evaluación:

- La MIR está formulada de manera estructurada, con niveles de objetivos, indicadores y métodos de cálculo.
- **Áreas de mejora:**
 - ✗ No se identifican claramente los **medios de verificación** ni los **supuestos**.
 - ✗ Algunos indicadores están orientados a procesos y no a resultados (ej. adquisición de insumos, mantenimiento de maquinaria).
 - ✗ El Fin está redactado con una narrativa abstracta (“regeneración ética”), lo cual puede dificultar su medición.

Recomendaciones:

- **Fin:** Reformular para enfocarlo en el impacto sectorial concreto, por ejemplo:

“Contribuir a la mejora de la seguridad documental en los servicios de la Administración Pública Federal mediante el suministro de documentos oficiales con altos estándares de seguridad.”

- **Incluir Medios de Verificación:** Reportes de calidad, acuses de entrega, encuestas de satisfacción.
- **Incluir Supuestos Críticos:** Ej. Disponibilidad oportuna de insumos especializados, continuidad operativa de las impresoras certificadas.

Conclusión Ejecutiva

La información contenida en "Información B001" cumple parcialmente con los parámetros establecidos en la Metodología del Marco Lógico. Las principales áreas de mejora se concentran en:

1. **Precisión técnica en la definición del problema y objetivo:** se recomienda un ajuste que incorpore de forma explícita la población afectada y el resultado esperado.
2. **Justificación y delimitación de la población objetivo:** se debe establecer la relación causal y metodológica entre la población y el problema público.
3. **Fortalecimiento de la MIR:** es necesario incorporar medios de verificación, supuestos y reformular indicadores con enfoque en resultados.

DeepSeek

Análisis y Recomendaciones para el Programa Presupuestario B001

1. Población Potencial y Objetivo

- Población Potencial: 20 Dependencias y 210 Entidades de la APF.
- Población Objetivo: 20 Dependencias y 210 Entidades de la APF.

Evaluación:

- Cumple con la delimitación y cuantificación requerida. Sin embargo, se recomienda especificar si la población objetivo es un subconjunto de la potencial (ejemplo: priorización por riesgo de falsificación).

Recomendación:

- Aclarar si la población objetivo es igual a la potencial o si hay criterios de elegibilidad adicionales.

2. Definición del Problema Público

"Las Dependencias y Entidades de la APF utilizan documentos probatorios de identidad, de carácter fiscal y social, susceptibles de ser falsificados y alterados."

Evaluación:

- Cumple con:
 - Población objetivo + problema público.
 - Formulado como hecho negativo ("susceptibles de ser falsificados").
- Falta:
 - No identifica un cambio deseado en la población afectada (ejemplo: reducción de delitos por falsificación).
 - Usa términos como "susceptibles" (no prohibidos, pero podría mejorarse).

Recomendación:

- Reformular: "Las Dependencias y Entidades de la APF utilizan documentos probatorios falsificados o alterados, lo que facilita delitos como robo de identidad y evasión fiscal, sin que existan mecanismos suficientes para garantizar su integridad."

3. Definición del Objetivo**

"Las Dependencias y Entidades de la APF utilizan documentos probatorios seguros y confidenciales que permiten reducir los riesgos de falsificación y alteración."

Evaluación:

- Es la contraparte positiva del problema y es único.
- Falta: No especifica el cambio concreto en la población

Recomendación:

- Añadir una meta cuantificable: "...seguros y confidenciales, reduciendo en un 30% los delitos asociados a falsificación para 2025."

4. Árbol de Problemas

Recomendación:

- **Raíces (causas):**
 - Insuficientes medidas de seguridad en la producción.
 - Acceso no regulado a maquinaria especializada.
 - Falta de capacitación en manejo de documentos.
- **Copa (efectos):**
 - Aumento de delitos financieros.
 - Pérdida de confianza en instituciones públicas.
 - Daños a la salud (ejemplo: bebidas adulteradas).

5. Árbol de Objetivos

Recomendación:

- Medios (soluciones):
 - Implementar medidas de seguridad avanzadas.
 - Restringir el acceso a maquinaria especializada.
 - Capacitar al personal en manejo seguro.
- Fines:
 - Reducción de delitos financieros.
 - Restauración de confianza pública.

6. Matriz de Indicadores de Resultados (MIR)

Evaluación:

- **Fin:** Bien alineado con el Plan Nacional de Desarrollo, pero el indicador (% eficiencia) no mide directamente el impacto ético.
- **Propósito:** Indicador adecuado (satisfacción del cliente), pero falta vincularlo a reducción de falsificaciones.
- **Componentes/Actividades:** Coherentes, pero los indicadores son de proceso, no de resultado (ejemplo: "impresos producidos" vs. "impresos que redujeron fraudes").

Recomendaciones:

- **Fin:** Cambiar indicador a "Porcentaje de reducción de delitos vinculados a documentos falsos (fuente: SSP)."
- **Propósito:** Añadir "Porcentaje de impresos sin reportes de falsificación (fuente: auditorías)."

7. Indicadores y Supuestos

- **Problemas Identificados:**
 - Indicadores de gestión (ejemplo: % mantenimiento) no vinculados a resultados.
 - Supuestos no mencionados (ejemplo: "Estabilidad en el presupuesto asignado").
- **Recomendaciones:**
 - Añadir indicadores de impacto (ejemplo: "Número de casos de fraude detectados").
 - Incluir supuestos clave:
 - Disponibilidad continua de insumos de seguridad.
 - Cooperación interinstitucional para reportar falsificaciones.

Propuesta de Mejora Integrada

1. Problema Público Reformulado:

"Las Dependencias y Entidades de la APF utilizan documentos probatorios falsificados o alterados, lo que facilita delitos como robo de identidad y evasión fiscal, sin que existan mecanismos suficientes para garantizar su integridad."

2. Objetivo Ajustado:

"Las Dependencias y Entidades de la APF utilizan documentos probatorios seguros y confidenciales, reduciendo en un 30% los delitos asociados a falsificación para 2025."

3. MIR Modificada:

- Fin: "Contribuir a la regeneración ética de las instituciones, reduciendo en un 30% los delitos vinculados a documentos falsos (indicador: reportes de la SSP)."
- Propósito: "Documentos entregados con 0% de reportes de falsificación en auditorías anuales."
- Componentes: Incluir "Número de impresos con elementos de seguridad validados."

Nota: Las modificaciones propuestas aseguran alineación con la Metodología de Marco Lógico, enfocándose en resultados medibles y claros.

Claude

Analizaré el programa presupuestario B001 siguiendo los lineamientos de la Metodología del Marco Lógico (MML) presentados en el documento "Información Evaluación", identificando fortalezas y áreas de mejora.

1. Definición de la Población

Estado Actual:

- **Población Potencial:** "20 Dependencias y 210 Entidades de la Administración Pública Federal"
- **Población Objetivo:** "20 Dependencias y 210 Entidades de la Administración Pública Federal"

Análisis:

- La población está claramente delimitada y cuantificada.
- Sin embargo, hay completa igualdad entre población potencial y objetivo, lo que sugiere que no existen criterios de priorización o focalización.
- No se menciona la población atendida en un ejercicio fiscal específico.

Recomendación: Establecer criterios de priorización para distinguir la población objetivo de la potencial, considerando factores como nivel de riesgo de falsificación, volumen de documentos, o importancia estratégica.

2. Definición del Problema

Estado Actual: "Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, susceptibles de ser falsificados y alterados."

Análisis:

- El problema identifica claramente la población afectada (Dependencias y Entidades de la APF).
- Se expresa como un hecho negativo (documentos "susceptibles de ser falsificados").
- Sin embargo, está formulado más como una condición permanente que como un problema específico que requiere intervención.

Recomendación: Reformular el problema para reflejar mejor la situación negativa: "Las Dependencias y Entidades de la Administración Pública Federal enfrentan altos riesgos de falsificación y alteración en sus documentos probatorios de identidad, de carácter fiscal y social."

3. Análisis del Problema (Árbol de Problemas)

Estado Actual: Se identifica implícitamente algunas causas como:

- Vulnerabilidad de los documentos por falta de elementos de seguridad.
- Inadecuado tratamiento de la información.

- Carencia de medidas de seguridad físicas, administrativas o técnicas.

Y algunos efectos como:

- Robo de identidad
- Evasión fiscal
- Daños a la salud (bebidas alcohólicas adulteradas)
- Usurpación de profesiones
- Fraudes financieros

Recomendación: Desarrollar un árbol de problemas explícito que muestre claramente la relación causal entre:

Causas directas:

- Insuficientes elementos de seguridad en los documentos oficiales
- Tecnologías de falsificación cada vez más sofisticadas
- Falta de estándares unificados para documentos de seguridad

Causas indirectas:

- Limitada capacidad técnica para la producción de documentos seguros
- Alto costo de tecnologías de seguridad documental
- Débil coordinación interinstitucional para estandarización

Efectos directos:

- Incremento de documentos falsos en circulación
- Pérdida de confianza en documentos oficiales
- Aumento de fraudes y delitos asociados

Efecto final:

- Afectación a la seguridad jurídica y credibilidad del Gobierno Federal

4. Definición del Objetivo

Estado Actual: "Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, seguros y confidenciales que permiten reducir los riesgos de falsificación y alteración."

Análisis:

- Identifica el área de enfoque (Dependencias y Entidades de la APF).
- Señala el cambio deseado (documentos seguros que reducen riesgos).
- Es un único objetivo.

- Corresponde a la solución del problema identificado.

Fortaleza: La definición del objetivo cumple adecuadamente con los criterios establecidos.

5. Análisis de la MIR

5.1 Lógica Vertical

Fin: "Contribuir a la Regeneración Ética de las Instituciones, mediante la producción de impresos que requieran las Dependencias y Entidades de la Administración Pública Federal bajo los principios de austeridad, honestidad y transparencia."

Análisis:

- Se formula como una contribución a un objetivo superior.
- Sin embargo, no se conecta directamente con la problemática de falsificación de documentos.

Recomendación: Reformular el Fin: "Contribuir a la seguridad documental y la confianza en las instituciones públicas mediante la reducción de los riesgos de falsificación y alteración de documentos oficiales."

Propósito: "Las Dependencias y Entidades de la Administración Pública Federal reciben los impresos solicitados de conformidad con las características técnicas que requieren."

Análisis:

- No refleja el cambio esperado en la población objetivo.
- Se enfoca en la entrega conforme de los impresos y no en la reducción de falsificación.

Recomendación: Ajustar el Propósito para alinearlo con el objetivo: "Las Dependencias y Entidades de la APF utilizan documentos oficiales con elementos de seguridad que reducen significativamente su vulnerabilidad a la falsificación y alteración."

Componentes:

- A: "Impresos valorados, numerados y de seguridad producidos"
- B: "Impresos diversos producidos"

Análisis:

- Los componentes representan los bienes producidos.
- Sin embargo, no reflejan aspectos como la calidad o características de seguridad específicas.

Recomendación: Especificar mejor los componentes:

- A: "Impresos valorados, numerados y de seguridad producidos con estándares técnicos que garantizan su inviolabilidad"
- B: "Impresos diversos producidos conforme a requerimientos de confidencialidad y seguridad establecidos"

Actividades:

- A1: "Adquisición de materias primas, insumos y refacciones"
- B2: "Ejecución del programa de mantenimiento preventivo de maquinaria de producción"

Análisis:

- Las actividades son limitadas y no cubren todo el proceso necesario.
- Faltan actividades relacionadas con desarrollo de elementos de seguridad, control de calidad, etc.

Recomendación: Ampliar las actividades:

- A1: "Adquisición de materias primas e insumos especializados para impresión segura"
- A2: "Desarrollo e implementación de elementos de seguridad innovadores"
- A3: "Implementación de controles de calidad en procesos de impresión de seguridad"
- B1: "Ejecución del programa de mantenimiento preventivo de maquinaria de producción"
- B2: "Capacitación técnica especializada al personal en nuevas tecnologías de seguridad documental"

5.2 Lógica Horizontal (Indicadores)

Análisis:

- Los indicadores están presentes para todos los niveles de la MIR.
- Sin embargo, varios indicadores se enfocan en eficiencia y producción, no en seguridad documental.

Recomendaciones para indicadores:

Fin:

- Actual: "Porcentaje de eficiencia de los procesos productivos"
- Propuesta: "Índice de confiabilidad de documentos oficiales" o "Tasa de falsificaciones detectadas en documentos oficiales"

Propósito:

- Actual: "Porcentaje de satisfacción del cliente"
- Propuesta: "Porcentaje de documentos oficiales que cumplen con estándares de seguridad internacional" y "Tasa de vulneración exitosa en pruebas controladas"

Componentes:

- Mantener los indicadores actuales pero complementarlos con indicadores de calidad y seguridad.

Actividades:

- Complementar con indicadores que midan la calidad de insumos adquiridos y la efectividad de la capacitación.

6. Conclusiones y Recomendaciones Finales

1. **Fortalecer la distinción entre población potencial y objetivo** mediante criterios de priorización.
2. **Reformular el problema** para reflejar mejor la situación negativa que se pretende resolver.
3. **Desarrollar un árbol de problemas explícito** que muestre claramente las relaciones causales.
4. **Revisar la lógica vertical de la MIR**, especialmente el Fin y el Propósito, para alinearlos mejor con la problemática de falsificación documental.
5. **Ampliar el conjunto de actividades** para cubrir todos los procesos necesarios.
6. **Reformular los indicadores** para que midan no solo la eficiencia productiva sino también la efectividad en términos de seguridad documental.
7. **Fortalecer la conexión entre el árbol de problemas y el árbol de objetivos** para garantizar que cada causa tenga un medio correspondiente y cada efecto un fin.

Estas mejoras permitirían que el programa presupuestario B001 tenga mayor coherencia interna y se alinee mejor con la Metodología del Marco Lógico, fortaleciendo su capacidad para abordar efectivamente la problemática de falsificación documental.

4.5.2.2 Análisis programa presupuestario P002

Prompt: *“Con base en el documento "Información Evaluación", verifica si la información de documento "Información P002" cumple con los parámetros descritos, el análisis debe ser preciso y alineado con los lineamientos de la Metodología del Marco Lógico. Si consideras que no es así, presenta las modificaciones correspondientes”.*

Chat GPT

A continuación, se detallan los hallazgos, evaluaciones y propuestas de mejora.

1. Definición del problema

Redacción actual:

“Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles.”

Evaluación:

- **✗** No está redactado como un hecho negativo (no se identifica un problema existente, sino una condición deseable).
- **✗** No se apega a la estructura “Población objetivo + problema público”.
- **✗** No identifica un cambio perjudicial que afecta actualmente a los ejecutores del gasto.
- **✗** No está acotado ni caracterizado territorial ni temporalmente.

Propuesta de redacción:

“Los ejecutores del gasto público presentan deficiencias en la asignación presupuestaria debido a una planeación, seguimiento y evaluación ineficaces, lo que limita el cumplimiento eficiente de los objetivos del desarrollo nacional.”

2. Población potencial y objetivo

Evaluación:

- ✓ La población está identificada y cuantificada (54 ejecutores del gasto).
- ✗ Falta justificación sobre cómo esta población enfrenta el problema identificado.
- ✗ No se explicitan los criterios de elegibilidad ni su vinculación con el problema público.

Recomendación:

- Incluir justificación técnica que vincule a los 54 ejecutores con problemáticas específicas de asignación, eficiencia y desempeño presupuestario.

3. Árbol de Problemas

Evaluación:

- ✓ Presenta causas y efectos diferenciados.
- ✓ Utiliza redacción en términos negativos.
- ✓ Incluye más de un nivel causal.
- ✗ Algunas causas exceden el ámbito de control directo del Pp (e.g., “divergencia en ingresos previstos”, “ambigüedad normativa”).
- ✗ No se prioriza un único efecto final.

Recomendaciones:

- Priorizar una sola línea causal directa que derive del mandato institucional de la SHCP y de las UR participantes.
- Asegurar que todas las causas estén dentro del ámbito de competencia del programa, y dejar como supuestos las variables externas.

4. Definición del objetivo

Redacción actual:

“Los ejecutores del gasto reciben de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles.”

Evaluación:

- ✗ El objetivo no presenta un cambio observable en la población.

- ✗ Está formulado como una condición deseable, no como una transformación de la situación problemática.
- ✗ No está suficientemente enfocado ni es cuantificable.

Propuesta de redacción:

“Los ejecutores del gasto mejoran la eficacia y eficiencia en la planeación, programación y ejercicio del gasto público mediante asignaciones presupuestarias alineadas a objetivos de desarrollo nacional.”

5. Matriz de Indicadores para Resultados (MIR)

Evaluación:

- ✓ Se identifican niveles de objetivos (Fin, Propósito, Componentes, Actividades).
- ✓ Incluye indicadores e instrumentos metodológicos.
- ✗ El objetivo de nivel Fin está sobredimensionado (“contribuir a la paz y Estado democrático”), no es medible ni atribuible al Pp.
- ✗ Se omiten **medios de verificación** y **supuestos**, requeridos por la MML.
- ✗ Algunos indicadores miden procesos indirectos o no atribuibles de forma clara al programa.

Recomendaciones:

- **Fin:** reformular a un nivel de desarrollo alineado, pero atribuible, por ejemplo:

“Contribuir a la mejora de la calidad del gasto público mediante la articulación eficiente entre planeación, programación y evaluación del desempeño de los programas presupuestarios.”

- **Agregar medios de verificación:** informes del MSD, reportes de auditoría, resultados de evaluaciones externas.
- **Incluir supuestos clave:** continuidad normativa, disponibilidad de información oportuna, colaboración institucional.

Conclusión Ejecutiva

El documento "**Información P002**" presenta avances importantes en cuanto a la estructuración general de la información y alineación funcional del programa presupuestario. Sin embargo, **incumple elementos críticos de la Metodología del Marco Lógico**, particularmente en la definición del problema, objetivo, delimitación de población, y estructura completa de la MIR.

Se recomienda proceder con una revisión técnica integral del diagnóstico y la MIR, con especial énfasis en:

- Reformular el problema desde una lógica causal negativa y clara.
- Redefinir el objetivo en términos de cambio verificable y atribuible.
- Priorizar y depurar el árbol de problemas.

- Reestructurar la MIR con medios de verificación, supuestos y objetivos coherentes.

DeepSeek

Análisis y Modificaciones al Documento "Información P002" según los Lineamientos de la Metodología del Marco Lógico (MML)

1. Definición del Problema

Cumplimiento: Parcial.

Observaciones:

El problema actual: "Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia..." no cumple con los criterios de redacción exigidos:

- Falta negatividad clara: No usa términos como "no", "ningún", etc.
- No evita la ausencia de bienes/servicios: Usa "requieren", que implica carencia.
- Población objetivo no está claramente delimitada en términos de afectación directa.

Propuesta de Modificación:

"Los ejecutores del gasto (54 entidades) no logran cumplir los objetivos del desarrollo nacional debido a asignaciones presupuestarias ineficientes e ineficaces, en el marco de los recursos públicos disponibles."

Justificación:

Cumple con:

- Población objetivo cuantificada (54 entidades).
- Problema redactado como hecho negativo ("no logran").
- Enfoque en resultado (incumplimiento de objetivos).

2. Análisis del Problema (Árbol del Problema)

Cumplimiento: Parcial.

Observaciones:

- **Causas:**
 - Correctamente identificadas pero incluyen más de tres niveles (ej. "a, b, c, d").
 - Falta contextualización territorial/socioeconómica explícita.
- **Efectos:**
 - Priorizados correctamente, pero el efecto final ("Incumplimiento de objetivos del desarrollo nacional") debe ser único.

Propuesta de Modificación:

- **Causas (Raíces):**
 1. Desarticulación entre necesidades y recursos disponibles en la APF.
 2. Deficiente planeación de programas de inversión pública.
 3. Incumplimiento de medidas de disciplina presupuestaria.

- **Efectos (Ramas):**
 - Efecto final único: "Frenado del crecimiento sostenido e incluyente del país.

Justificación:

- Reduce a tres niveles de causas.
- Vincula el efecto final al Plan Nacional de Desarrollo.

3. Definición del Objetivo

Cumplimiento: Parcial.

Observaciones:

- El objetivo actual es un reflejo positivo del problema, pero:
- No especifica el cambio concreto en la población.
- Falta cuantificación (ej. porcentaje de mejora esperada).

Propuesta de Modificación:

"Garantizar que el 100% de los ejecutores del gasto (54 entidades) reciban asignaciones presupuestarias eficientes y eficaces, logrando un 90% de cumplimiento en los objetivos del desarrollo nacional para 2024."

Justificación:

- Incluye:
- Población objetivo clara.
- Cambio medible (90% de cumplimiento).
- Alineación con el problema.

4. Matriz de Indicadores para Resultados (MIR)

Cumplimiento: Parcial.

Observaciones:

- **Fin:** Bien alineado con el PND, pero falta claridad en la contribución específica del programa.
- **Componentes:** Indicadores son técnicos pero no todos tienen metas cuantificadas (ej. "Porcentaje de MIR con valoración satisfactoria").

Propuesta de Modificación:

Nivel	Indicador	Meta	Medios de Verificación
Fin	% de programas presupuestarios alineados al PND con logro satisfactorio (MSD ≥ 4).	80% para 2024	Informes anuales del MSD y SIIPPG.
Propósito	% de gasto programable ejecutado con eficiencia (MSD ≥ 4).	75% para 2024	Reportes de la SHCP y evaluaciones externas.

Componentes	% de proyectos de inversión aprobados con evaluación de impacto positiva.	70% para 2024	Reportes de la SHCP y evaluaciones externas.
--------------------	---------------------------------------------------------------------------	---------------	----------------------------------------------

Justificación:

- Metas específicas, cuantificables y con fuentes de verificación claras.
- Alineación con los niveles de la MML.

Recomendaciones Finales

1. Contextualizar el problema: Añadir datos socioeconómicos (ej. impacto del PIB por ineficiencia del gasto).
2. Simplificar el árbol de causas: Mantener máximo tres niveles jerárquicos.

Nota: Las modificaciones propuestas aseguran el cumplimiento estricto de la MML y los lineamientos del CONEVAL, priorizando claridad, medición y alineación con el PND.

Claude

He revisado detalladamente ambos documentos para evaluar si el programa presupuestario P002 "Diseño y Conducción de la Política de Gasto" cumple con los lineamientos de la Metodología del Marco Lógico (MML). A continuación, presento mi análisis:

1. Definición del problema

Criterios de evaluación:

- Formulación clara, concreta, acotada y sin ambigüedades
- Identificación clara de la población afectada
- Planteamiento como hecho negativo
- Identificación de un cambio sobre la población
- No definido como ausencia de un bien o servicio
- Estructura "Población objetivo + problema público"

Problema actual:

"Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles."

Observaciones:

- No está planteado como un hecho negativo
- Utiliza la palabra "requieren" que sugiere ausencia o necesidad
- No identifica claramente un cambio sobre la población

Propuesta de mejora:

"Los ejecutores del gasto enfrentan dificultades para recibir asignaciones presupuestarias eficientes y eficaces que contribuyan al cumplimiento de los objetivos del desarrollo nacional."

2. Análisis del problema

Criterios de evaluación:

- Causas definidas en términos negativos
- Máximo tres niveles de causas
- Alineación con el ámbito de competencia
- Priorización de un único efecto final
- Contextualización socioeconómica y territorial

Observaciones:

- Las causas están definidas correctamente en términos negativos
- Respeta la estructura de tres niveles
- Los efectos culminan en un único efecto principal: "Incumplimiento de los objetivos del desarrollo nacional"
- Falta contexto socioeconómico específico del país
- Carece de caracterización territorial específica

Mejoras sugeridas:

- Agregar contexto sobre la situación fiscal actual de México
- Incluir referencias a disparidades territoriales en la asignación y ejercicio del gasto público

3. Definición del objetivo

Criterios de evaluación:

- Identificación del área de enfoque
- Identificación del cambio que busca generar
- Objetivo único
- Correspondencia con la solución del problema

Objetivo actual:

"Los ejecutores del gasto reciben de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles."

Observaciones:

- Identifica correctamente el área de enfoque (ejecutores del gasto)
- El cambio que busca generar no es completamente claro
- Es un objetivo único
- Hay un error de redacción: "reciben de asignaciones"

Propuesta de mejora:

"Los ejecutores del gasto reciben asignaciones presupuestarias eficientes y eficaces que contribuyen al cumplimiento de los objetivos del desarrollo nacional, optimizando los recursos públicos disponibles."

4. Elaboración de la MIR

Criterios de evaluación:

- Estructura en cuatro filas y cuatro columnas
- Lógica vertical (actividades → componentes → propósito → fin)
- Lógica horizontal (objetivos + indicadores + medios de verificación + supuestos)

Observaciones:

- La MIR presentada cumple con la estructura básica de cuatro filas
- El Fin está definido de manera muy general
- Los indicadores presentan fórmulas de cálculo, pero falta información sobre frecuencia, línea base y metas
- Falta la columna de supuestos
- Los componentes parecen adecuados (Cartera de PPI, Información de desempeño, Informes de Contabilidad, MIR valorada)
- No se incluyen las actividades para cada componente

Propuesta de mejora de la lógica vertical:

1. Fin: "Contribuir al desarrollo nacional mediante el cumplimiento eficiente y eficaz de los objetivos y metas de los programas presupuestarios"

2. Propósito: "Los ejecutores del gasto implementan programas presupuestarios que cumplen con los objetivos y metas establecidos"

3. Componentes:

- A: "Cartera de Programas y Proyectos de Inversión de infraestructura productiva aprobada"
- B: "Información de desempeño de programas presupuestarios valorada y mejorada"
- C: "Informes de Contabilidad Gubernamental e Informes de Gestión Pública emitidos"
- D: "Matrices de Indicadores para Resultados con valoración satisfactoria implementadas"

4. Actividades (a incluir):

- Para componente A: "Evaluar y dictaminar programas y proyectos de inversión"
- Para componente B: "Realizar evaluaciones de desempeño de programas presupuestarios"
- Para componente C: "Integrar la información contable y presupuestaria"
- Para componente D: "Asesorar técnicamente en la elaboración de MIR"

Conclusión

El programa presupuestario P002 presenta algunos elementos consistentes con la Metodología del Marco Lógico, pero requiere ajustes importantes para cumplir completamente con los lineamientos. Las principales áreas de mejora son:

1. Reformular el problema como un hecho negativo

2. Contextualizar el problema en la situación socioeconómica y territorial actual
3. Mejorar la redacción del objetivo para clarificar el cambio esperado
4. Completar la MIR con actividades específicas para cada componente
5. Incluir los supuestos en la matriz
6. Detallar la información de los indicadores (frecuencia, línea base, metas)

Estas modificaciones fortalecerían la estructura lógica del programa y mejorarían su alineación con la Metodología del Marco Lógico, facilitando así su implementación, seguimiento y evaluación.

4.5.3 Escenario con conocimiento

En este apartado se proveyó un documento diseñado por el evaluador en el que se definieron los elementos centrales para la evaluación, la información relevante que se desea sea contestada, así como el documento diagnóstico.

4.5.3.1 Análisis programa presupuestario B001

Prompt: *“Eres un evaluador de política pública en México y estás trabajando en la evaluación de diseño del programa presupuestario no social B001 de la Secretaría de Hacienda y Crédito Público utilizando la Metodología del Marco Lógico. Mi objetivo es identificar deficiencias en la definición de la población potencia, población objetivo, problema público identificado por el programa y los objetivos del programa, y analizar la calidad de la Matriz de Indicadores para Resultados (MIR), identificando inconsistencias en la formulación de sus niveles, indicadores y supuestos.*

Tareas: *Quiero que me ayudes a revisar la definición de:*

- *Población potencial y objetivo, asegurando claridad y delimitación.*
- *Problema público, verificando que esté formulado de manera correcta.*
- *Objetivo público, comprobando que sea la contraparte lógica del problema.*
- *Revisar la estructura del árbol del problema, asegurando que el problema, sus causas y efectos estén formulados correctamente.*
- *Validar la coherencia del árbol de objetivos, verificando que la relación medios-fines sea clara y lógica.*
- *Identificar inconsistencias y proponer ajustes para mejorar la formulación del problema y objetivos.*
- *Revisar la coherencia de la MIR, verificando que el Fin, Propósito, Componentes y Actividades estén correctamente formulados.*
- *Evaluar los indicadores, asegurando que sean medibles y respondan a las dimensiones de cantidad, calidad y tiempo.*
- *Identificar inconsistencias en los medios de verificación y supuestos, señalando riesgos potenciales.*
- *Proponer ajustes para mejorar la alineación de los objetivos y la claridad de los resultados esperados.*

- *El análisis debe ser preciso y alineado con los lineamientos de la Metodología del Marco Lógico.”*

Chat GPT

Con base en la revisión del documento “Diagnóstico Pp B001 (Simplificado)” y los lineamientos metodológicos de Marco Lógico (MML), a continuación te presento un análisis técnico por componente, señalando deficiencias, inconsistencias y recomendaciones puntuales:

1. Población potencial y población objetivo

Definición actual:

- **Población potencial:** Implícita. Se deduce que incluye a todas las dependencias y entidades de la APF que requieren impresos oficiales con características de seguridad.
- **Población objetivo:** Identificada como las “Dependencias y Entidades de la APF que enfrentan altos riesgos de falsificación en sus impresos”.

Observaciones:

- No hay cuantificación explícita ni delimitación temporal o espacial.
- No se aclara si la totalidad de las entidades públicas están incluidas o solo un subconjunto (p. ej., dependencias que emiten documentos críticos).

Recomendaciones:

- **Delimitar y cuantificar** ambas poblaciones. Ejemplo: “XXX dependencias federales que requieren documentos con características de seguridad en el ciclo fiscal 2023”.
- Distinguir si hay entidades prioritarias por riesgo o volumen.

2. Problema público

Definición actual:

“Las dependencias y entidades de la APF enfrentan altos riesgos de falsificación en sus impresos”.

Evaluación con base en criterios MML:

- ✓ Redacción negativa
- ✓ Enfocado en la población objetivo
- ✓ Singular y claro
- ✗ No identifica explícitamente un cambio en la población afectada
- ✗ No contextualiza territorial ni temporalmente

Recomendación:

Reformular:

“Las dependencias y entidades de la APF enfrentan altos riesgos de falsificación de sus impresos oficiales, lo que compromete la legalidad, seguridad y acceso a servicios públicos, debido a procesos de producción no alineados con estándares actualizados de seguridad y confidencialidad.”

3. Objetivo del programa

Definición actual:

Reducir el riesgo de falsificación de impresos mediante la producción especializada.

Evaluación:

- ✓ Es la contraparte lógica del problema
- ✗ No define el cambio esperado en la población
- ✗ No está redactado como estado positivo alcanzado

Propuesta de reformulación:

“Las dependencias y entidades de la APF cuentan con impresos oficiales seguros y con características que reducen sustancialmente el riesgo de falsificación y uso indebido.”

4. Árbol de problemas

Causas identificadas:

- Débil cumplimiento de estándares nacionales/internacionales.
- Restricción en la adquisición de insumos.
- Falta de mantenimiento de maquinaria.
- Rezago normativo.

Efectos:

- Imitación de elementos de seguridad.
- Fugas de información.
- Afectaciones a derechos y servicios públicos.

Observaciones:

- ✓ Las causas están bien vinculadas con el problema
- ✗ Algunas causas pueden clasificarse como efectos o consecuencias estructurales (p. ej., rezago normativo)
- ✗ Falta jerarquización de causas directas e indirectas

Recomendaciones:

- Reorganizar jerarquías (usar tres niveles máximo).
- Alinear con las atribuciones directas del programa.

5. Árbol de objetivos

Evaluación:

- ✓ Existe lógica medios-fines
- ✗ No está gráficamente representado
- ✗ Algunos medios (como actualización normativa) no son atribuibles al Pp B001

Recomendaciones:

- Separar medios internos (producción, control) de externos (normativa nacional/internacional).

6. Matriz de Indicadores para Resultados (MIR)

Diagnóstico preliminar:

No se incluye en el documento revisado. Sin embargo, con base en el diagnóstico se infiere que debe contener:

Fin:

Contribuir a la reducción del uso de documentos oficiales falsificados en la APF.

Propósito:

Reducir el riesgo de falsificación y vulneración de impresos oficiales producidos para la APF.

Componentes sugeridos:

- Impresos con características de seguridad producidos conforme a estándares internacionales.
- Publicaciones y materiales impresos entregados con oportunidad y calidad.

Actividades:

- Producción e impresión especializada.
- Mantenimiento y modernización de maquinaria.
- Adquisición de insumos certificados.

Indicadores:

✗ Se desconoce su formulación exacta, pero deben:

- Ser específicos por nivel (no repetir indicadores en Propósito y Componentes).
- Incluir dimensiones de cantidad, calidad y tiempo.

Supuestos y medios de verificación:

- ✗ Deben evaluarse los riesgos asociados a insumos restringidos, obsolescencia de maquinaria, seguridad física y digital.

✘ Medios de verificación deben corresponder con fuentes institucionales trazables (órdenes de trabajo, bitácoras de mantenimiento, etc.).

Conclusión y próximos pasos

El diagnóstico presenta insumos valiosos, pero requiere mejoras en formulación técnica bajo los lineamientos de la MML. Se recomienda:

1. Incorporar el árbol del problema y objetivos con relaciones causal-lógica explícitas.
2. Reformular problema y objetivo con base en criterios de claridad, unicidad y verificabilidad.
3. Exigir la versión oficial de la MIR para su validación técnica.

DeepSeek

Evaluación del Programa Presupuestario B001 utilizando la Metodología del Marco Lógico (MML)

1. Población Potencial y Objetivo

- **Población Potencial:** Todas las dependencias y entidades de la Administración Pública Federal (APF), que ascienden a 230 según la normativa vigente.
- **Población Objetivo:** Dependencias y entidades de la APF que solicitan impresos valorados, no valorados, numerados y de seguridad a la DGTIEV. En 2022 se atendió a 6 entidades (3.91% de cobertura).

Deficiencias:

- La delimitación es clara pero la cobertura es muy baja (3.91% en 2022), lo que sugiere un subutilización del programa.
- No se especifica si hay criterios de priorización para atender solicitudes.

Recomendación:

- Establecer mecanismos para incrementar la cobertura, como promoción activa del servicio entre las entidades no usuarias.
- Definir criterios de elegibilidad basados en urgencia o impacto (ej. documentos con alto riesgo de falsificación).

2. Problema Público

Formulación actual:

"Las Dependencias y Entidades de la Administración Pública Federal enfrentan altos riesgos de falsificación en sus impresos."

Evaluación:

- Cumple con los lineamientos de la MML:
 - Es negativo ("altos riesgos").
 - Identifica población afectada (dependencias/entidades de la APF).
 - Es único y concreto.
- **Falta:** No menciona las consecuencias sociales indirectas (ej. delitos penales, evasión fiscal).

Recomendación:

Reformular para incluir efectos:

"Las Dependencias y Entidades de la APF enfrentan altos riesgos de falsificación en sus impresos, lo que facilita delitos penales, evasión fiscal y daños a la salud pública."

3. Objetivo Público

Formulación actual: *"Las Dependencias y Entidades de la APF reducen los riesgos de falsificación en sus impresos."*

Evaluación:

- Es la contraparte lógica del problema (de "enfrentan riesgos" a "reducen riesgos").
- **Falta:** No especifica el medio (producción de impresos con elementos de seguridad).

Recomendación:

"Las Dependencias y Entidades de la APF reducen los riesgos de falsificación en sus impresos mediante la producción y suministro de documentos con estándares de seguridad elevados."

4. Árbol del Problema y Árbol de Objetivos

Árbol del Problema:

- **Problema central:** Correctamente formulado.
- **Causas:**
 - Debilidad en cumplimiento de estándares de seguridad (directa).
 - Incumplimiento de confidencialidad (directa).
- **Efectos:** Falsificación y fugas de información (consistentes).
- **Deficiencia:** Algunas causas secundarias (ej. rezago normativo) están mal redactadas (no son negativas).

Árbol de Objetivos:

- Relación medios-fines es clara (ej. "Mantenimiento de maquinaria" como medio para "cumplir estándares").
- **Falta:** Objetivos intermedios no están cuantificables.

Recomendaciones:

- Reformular causas/efectos en términos negativos.
- Añadir indicadores a objetivos intermedios (ej. "% de maquinaria con mantenimiento preventivo").

5. Matriz de Indicadores para Resultados (MIR)

Niveles:

- **Fin:** Alineado con el PND, pero el indicador ("*% de satisfacción del cliente*") no mide impacto real en reducción de falsificación.
- **Propósito:** Indicador ("*% de entidades atendidas*") es de cobertura, no de reducción de riesgos.
- **Componentes/Actividades:** Bien formulados pero falta dimensión de tiempo (ej. "entregas oportunas").

Deficiencias:

- Indicadores no capturan calidad o impacto (solo cantidad).
- Medios de verificación son internos (podrían sesgarse).

Recomendaciones:

- **Fin:** Cambiar a "*% de impresos falsificados reportados*".
- **Propósito:** Usar "*% de impresos entregados con estándares de seguridad certificados*".
- **Medios de verificación:** Incluir fuentes externas (ej. informes del CONEVAL o auditorías).

6. Supuestos y Riesgos

Supuestos actuales:

- Normativas internacionales se cumplen.
- Recursos presupuestales son suficientes.

Riesgos no considerados:

- Cambios tecnológicos que vuelvan obsoleta la maquinaria.
- Corrupción en el resguardo de impresos.

Recomendación:

Incluir supuestos como: "*La maquinaria se actualiza cada 5 años para seguir estándares internacionales*" y "*Los controles de confidencialidad se auditan anualmente*".

Conclusiones

El programa tiene una base sólida pero requiere ajustes para:

1. Mejorar la formulación del problema y objetivos (añadir consecuencias y medios).

2. Asegurar que los indicadores midan impacto, no solo procesos.
3. Incorporar supuestos que mitiguen riesgos externos.
4. Incrementar la cobertura y transparencia en medios de verificación.

Nota final: La MIR debe reflejar mejor la contribución del programa a la seguridad nacional y la prevención del delito, no solo la producción de impresos.

Claude

1. ANÁLISIS DE POBLACIÓN POTENCIAL Y OBJETIVO

Deficiencias Identificadas:

Población Potencial y Objetivo Idénticas:

- Ambas se definen como "20 Dependencias y 210 Entidades de la APF"
- **Problema:** No existe diferenciación entre quién presenta el problema y quién será atendido
- **Impacto:** Imposibilita la focalización y priorización de recursos

Propuesta de Corrección:

Población Potencial:

- 230 instituciones de la APF (20 Dependencias + 210 Entidades) que requieren documentos oficiales con elementos de seguridad para sus operaciones

Población Objetivo:

- Instituciones de la APF que soliciten anualmente servicios de impresión de documentos oficiales con características de seguridad, priorizando aquellas con mayor riesgo de falsificación

2. ANÁLISIS DEL PROBLEMA PÚBLICO

Problema Actual:

"Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, susceptibles de ser falsificados y alterados."

Deficiencias Críticas:

1. **Formulación Incorrecta:**
 - No está redactado como hecho negativo
 - No identifica claramente el problema sino una característica inherente
 - Falta el formato "Población + problema público"
2. **Ambigüedad Conceptual:**
 - No especifica qué constituye el problema real
 - Confunde la susceptibilidad con la ocurrencia del problema

Propuesta de Reformulación:

Problema Público Corregido: *"Las Dependencias y Entidades de la APF no cuentan con documentos oficiales que cumplan con estándares de seguridad suficientes para prevenir su falsificación y alteración, lo que genera riesgos de fraude, evasión fiscal y daños a la población."*

3. ANÁLISIS DEL OBJETIVO

Objetivo Actual:

"Las Dependencias y Entidades de la Administración Pública Federal utilizan documentos probatorios de identidad, de carácter fiscal y social, seguros y confidenciales que permiten reducir los riesgos de falsificación y alteración."

Deficiencias Identificadas:

1. Inconsistencia Lógica:

- No es la contraparte directa del problema identificado
- Agrega elementos no mencionados en el problema (confidencialidad)

2. Falta de Precisión:

- No especifica el cambio esperado claramente
- Uso de términos vagos ("seguros", "reducir riesgos")

Propuesta de Reformulación:

Objetivo Corregido: *"Las Dependencias y Entidades de la APF cuentan con documentos oficiales que cumplen con estándares de seguridad suficientes para prevenir su falsificación y alteración."*

4. ANÁLISIS DE LA ESTRUCTURA LÓGICA

Árbol de Problemas (Propuesta)

EFFECTOS:

- **Nivel 3:** Debilitamiento de la confianza ciudadana en las instituciones públicas
- **Nivel 2:** Incremento de actividades ilícitas (evasión fiscal, fraudes, daños a la salud)
- **Nivel 1:** Uso de documentos oficiales falsificados o alterados

PROBLEMA CENTRAL:

- Las Dependencias y Entidades de la APF no cuentan con documentos oficiales que cumplan con estándares de seguridad suficientes

CAUSAS:

- **Nivel 1:** Procesos de impresión sin medidas de seguridad adecuadas
- **Nivel 2:** Falta de maquinaria especializada y elementos de seguridad
- **Nivel 3:** Ausencia de restricciones en el acceso a tecnología de impresión

Árbol de Objetivos (Propuesta)

FINES:

- **Nivel 3:** Fortalecimiento de la confianza ciudadana en las instituciones públicas
- **Nivel 2:** Reducción de actividades ilícitas relacionadas con documentos falsos
- **Nivel 1:** Uso exclusivo de documentos oficiales auténticos

OBJETIVO CENTRAL:

- Las Dependencias y Entidades de la APF cuentan con documentos oficiales seguros

MEDIOS:

- **Nivel 1:** Procesos de impresión con medidas de seguridad especializadas
- **Nivel 2:** Disponibilidad de maquinaria e insumos de seguridad
- **Nivel 3:** Control estricto del acceso a tecnología de impresión

5. ANÁLISIS DE LA MATRIZ DE INDICADORES PARA RESULTADOS (MIR)**Deficiencias por Nivel:****FIN:****Problemas Identificados:**

- Desalineación con el problema público identificado
- Indicador no mide contribución específica del programa
- Fórmula compleja y poco clara para medir "regeneración ética"

Propuesta de Corrección:

- **Objetivo:** Contribuir a la reducción de delitos asociados a la falsificación de documentos oficiales
- **Indicador:** Porcentaje de reducción de delitos por falsificación de documentos oficiales

PROPÓSITO:**Problemas Identificados:**

- Indicador mide satisfacción, no el cambio esperado en la población
- No refleja la solución del problema público

Propuesta de Corrección:

- **Objetivo:** Las Dependencias y Entidades de la APF reciben documentos oficiales con estándares de seguridad que previenen su falsificación
- **Indicador:** Porcentaje de documentos oficiales entregados que cumplen con estándares de seguridad establecidos

COMPONENTES:**Problemas Identificados:**

- Separación artificial entre "impresos valorados/numerados/de seguridad" e "impresos diversos"
- No todos los "impresos diversos" requieren medidas de seguridad

Propuesta de Corrección:

- **Componente A:** Documentos oficiales con elementos de seguridad producidos
- **Componente B:** Documentos oficiales sin elementos de seguridad producidos

ACTIVIDADES:

Problemas Identificados:

- Actividades no están correctamente numeradas
- Falta una actividad clave: el proceso de producción mismo

Propuesta de Corrección:

- **Actividad A.1:** Adquisición de materias primas e insumos especializados
- **Actividad A.2:** Producción de documentos oficiales con elementos de seguridad
- **Actividad A.3:** Mantenimiento preventivo de maquinaria especializada

6. ANÁLISIS DE INDICADORES

Deficiencias Generales:

1. **Dimensiones Incompletas:**
 - Mayoría de indicadores solo miden eficacia
 - Ausencia de indicadores de calidad y tiempo
2. **Fórmulas Inadecuadas:**
 - Indicador de Fin tiene fórmula financiera inadecuada
 - Algunos indicadores no son medibles directamente
3. **Falta de Especificidad:**
 - No se definen claramente los estándares de calidad
 - Ausencia de metas específicas y temporales

Propuestas de Mejora:

Indicadores Corregidos:

Nivel	Indicador	Fórmula	Dimensión
Fin	Tasa de variación de delitos por falsificación	$((\text{Delitos año } t / \text{Delitos año } t-1) - 1) \times 100$	Eficacia
Propósito	Porcentaje de documentos seguros entregados	$(\text{Documentos con estándares de seguridad} / \text{Total documentos entregados}) \times 100$	Calidad
Componente A	Porcentaje de documentos seguros producidos	$(\text{Documentos seguros producidos} / \text{Documentos seguros solicitados}) \times 100$	Eficacia
Actividad A.1	Porcentaje de solicitudes de compra atendidas oportunamente	$(\text{Solicitudes atendidas en tiempo} / \text{Total solicitudes}) \times 100$	Eficiencia

7. MEDIOS DE VERIFICACIÓN Y SUPUESTOS

Deficiencias en Medios de Verificación:

- No se especifican fuentes de información
- Falta de sistemas de registro y seguimiento
- Ausencia de mecanismos de validación externa

Deficiencias en Supuestos:

- No se identifican supuestos en la MIR actual
- Falta análisis de riesgos externos
- Ausencia de factores críticos de éxito

Propuestas:

Medios de Verificación:

- Registros de producción de la entidad ejecutora
- Reportes de entrega y conformidad de beneficiarios
- Estadísticas oficiales de delitos (SESNP)
- Sistemas de seguimiento de calidad

Supuestos Críticos:

- Las Dependencias y Entidades mantienen su demanda de documentos oficiales
- Se mantiene la disponibilidad de insumos especializados en el mercado
- No se presentan cambios normativos que afecten los estándares de seguridad
- Los proveedores de maquinaria especializada mantienen el servicio de mantenimiento

8. RECOMENDACIONES GENERALES

Acciones Prioritarias:

1. **Reformular completamente el problema público** siguiendo la metodología del Marco Lógico
2. **Redefinir población potencial y objetivo** con criterios claros de diferenciación
3. **Reestructurar la MIR** asegurando coherencia lógica entre niveles
4. **Desarrollar indicadores SMART** que midan las tres dimensiones (cantidad, calidad, tiempo)
5. **Implementar sistema de seguimiento** con medios de verificación confiables
6. **Identificar y documentar supuestos críticos** para la gestión de riesgos

Alineación con Marco Normativo:

- Cumplir con Lineamientos Generales para la Evaluación de los Programas Federales
- Aplicar Términos de Referencia para la Evaluación de Diseño de CONEVAL
- Seguir Guía para la Elaboración de la Matriz de Indicadores para Resultados

CONCLUSIONES

El programa P002 presenta deficiencias estructurales significativas en su diseño que comprometen su efectividad y evaluabilidad. La reformulación propuesta busca:

- Establecer una lógica causal clara entre problema, objetivos y medios
- Definir poblaciones diferenciadas y focalizadas
- Crear indicadores medibles y relevantes
- Asegurar la coherencia interna del programa
- Facilitar el seguimiento y evaluación de resultados

La implementación de estas correcciones fortalecerá significativamente la capacidad del programa para generar los resultados esperados y contribuir efectivamente a la reducción de riesgos asociados a la falsificación de documentos oficiales.

4.5.3.2 Análisis programa presupuestario P002

Prompt: *“Eres un evaluador de política pública en México y estás trabajando en la evaluación de diseño del programa presupuestario no social P002 de la Secretaría de Hacienda y Crédito Público utilizando la Metodología del Marco Lógico. Mi objetivo es identificar deficiencias en la definición de la población potencia, población objetivo, problema público identificado por el programa y los objetivos del programa, y analizar la calidad de la Matriz de Indicadores para Resultados (MIR), identificando inconsistencias en la formulación de sus niveles, indicadores y supuestos.*

Tareas: *Quiero que me ayudes a revisar la definición de:*

- *Población potencial y objetivo, asegurando claridad y delimitación.*
- *Problema público, verificando que esté formulado de manera correcta.*
- *Objetivo público, comprobando que sea la contraparte lógica del problema.*
- *Revisar la estructura del árbol del problema, asegurando que el problema, sus causas y efectos estén formulados correctamente.*
- *Validar la coherencia del árbol de objetivos, verificando que la relación medios-fines sea clara y lógica.*
- *Identificar inconsistencias y proponer ajustes para mejorar la formulación del problema y objetivos.*
- *Revisar la coherencia de la MIR, verificando que el Fin, Propósito, Componentes y Actividades estén correctamente formulados.*
- *Evaluar los indicadores, asegurando que sean medibles y respondan a las dimensiones de cantidad, calidad y tiempo.*
- *Identificar inconsistencias en los medios de verificación y supuestos, señalando riesgos potenciales.*
- *Proponer ajustes para mejorar la alineación de los objetivos y la claridad de los resultados esperados.*
- *El análisis debe ser preciso y alineado con los lineamientos de la Metodología del Marco Lógico.”*

Chat GPT

Con base en el diagnóstico oficial del Programa presupuestario P002 “Diseño y Conducción de la Política de Gasto Público”, a continuación te presento el análisis técnico conforme a los lineamientos de la Metodología del Marco Lógico (MML) y al enfoque de evaluación de diseño:

1. Población potencial y población objetivo

Definición actual (pág. 17-21):

- **Población potencial y objetivo:** Ejecutores del gasto, incluyendo dependencias y entidades de la APF, Poderes, Entes Autónomos y Empresas Productivas del Estado (EPE).

Observaciones técnicas:

- ✓ Se identifican claramente los grupos institucionales.
- ✓ Existe cuantificación desagregada (54 ejecutores del gasto).
- ✗ No se justifica el criterio de elegibilidad para ser considerado población objetivo.
- ✗ Se omite la delimitación temporal precisa en términos de planeación operativa o presupuestal (más allá de la actualización anual referida).

Recomendación:

- Incluir una justificación técnica de cómo y por qué esas instituciones constituyen la población objetivo.
- Precisar si todos los ejecutores reciben atención efectiva o si se priorizan con base en criterios como magnitud del presupuesto, impacto estratégico o nivel de cumplimiento.

2. Problema público

Definición actual (pág. 7):

“Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional”.

Evaluación de acuerdo con los criterios de MML:

- ✓ Está redactado como hecho negativo.
- ✓ Se identifica claramente a la población afectada.
- ✗ No expresa un cambio medible en dicha población.
- ✗ Presenta formulación ambigua (se describe más una necesidad que una condición adversa concreta).
- ✗ No describe consecuencias reales del problema.

Propuesta de reformulación:

“Los ejecutores del gasto de la APF enfrentan deficiencias en la asignación presupuestaria, lo que limita su capacidad para cumplir de manera eficaz y eficiente los objetivos del desarrollo nacional”.

3. Objetivo público

Definición actual (pág. 15):

“Los ejecutores del gasto reciben las asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia...”.

Evaluación:

- ✓ Está vinculado lógicamente con el problema.
- ✗ Falta precisión en el resultado esperado (¿cómo se define una asignación eficiente?).
- ✗ Redacción pasiva; no comunica cambio directo en capacidades o resultados.

Propuesta de reformulación:

“Los ejecutores del gasto cuentan con asignaciones presupuestarias eficientes, oportunas y alineadas a los objetivos del desarrollo nacional, lo que fortalece su capacidad institucional y el cumplimiento de metas”.

4. Árbol del problema

Observaciones:

- ✓ Se identifican causas como: falta de alineación normativa, limitaciones en integración de información y procesos desactualizados.
- ✗ Algunos elementos son confusos en su jerarquía: por ejemplo, la “no alineación entre prioridades y recursos” aparece como causa, aunque es también un efecto.

Recomendación:

- Reestructurar el árbol respetando los tres niveles máximos de causas, ubicando factores institucionales (normatividad, sistemas, procesos) como causas estructurales.

5. Árbol de objetivos

Observaciones (pág. 14):

- ✗ No se anexa gráficamente.
- ✓ El árbol lógico se infiere desde la descripción funcional (Componentes y Actividades).
- ✗ Falta claridad en los vínculos medios-fines entre servicios específicos (ej. normativa, evaluación, integración de informes) y el objetivo de política pública.

Recomendación:

- Elaborar árbol explícito, relacionando cada servicio con su causa a resolver.

6. Matriz de Indicadores para Resultados (MIR)

Fin

Definición: Contribuir al cumplimiento de objetivos de los programas presupuestarios.

- ✓ Lógica vertical aceptable.
- ✗ Contribución indirecta y difícil de atribuir.

Propósito

Definición: “Los ejecutores del gasto reciben asignaciones...”.

- ✗ Formulación ambigua (no está redactado como resultado alcanzado).
- ✓ Está alineado con el problema.

Componentes

- Integración oportuna del PPEF.
- Cobertura de objetivos del PND.
- Atención de adecuaciones presupuestarias.

- ✓ Se identifican como entregables concretos.
- ✗ Se observa duplicidad entre Componente y Actividad.

Actividades

- Generación de información del desempeño.
- Emisión de disposiciones normativas.
- Atención de solicitudes de inversión, etc.

- ✓ Adecuadas.
- ✗ Algunas deben reclasificarse como productos intermedios.

7. Indicadores

Evaluación general:

- ✓ Se estructuran por nivel MIR.
- ✓ Incorporan dimensiones de calidad y tiempo en varios casos.
- ✗ Algunos indicadores no están vinculados directamente con los objetivos (ej. ISD entregados ≠ asignaciones eficientes).
- ✗ Falta claridad metodológica en fórmulas compuestas o complejas (ej. recomposición del gasto programable).

Recomendaciones:

- Replantear indicadores del Propósito con foco en resultados institucionales (por ejemplo: porcentaje de ejecutores que reciben recursos conforme a techos definidos y con suficiencia presupuestaria).
- Incluir metas verificables y fuentes institucionales (SICOP, PIPP, ISD, FARE).

8. Medios de verificación y supuestos

Medios de verificación:

- ✓ Se mencionan informes, sistemas, indicadores institucionales.
- ✗ Falta precisión sobre su accesibilidad, frecuencia y validez.

Supuestos:

- ✗ No se describen en el diagnóstico ni en la MIR.
- ✗ No se señalan factores de riesgo clave (p. ej., cambios normativos, restricciones presupuestales, rotación de personal, plataformas digitales no operativas).

Recomendación:

- Incluir al menos tres supuestos clave por nivel MIR.
- Incorporar un análisis de riesgos operativos y contextuales.

Recomendaciones finales

1. **Reformular problema y objetivo** para cumplir con criterios MML (singularidad, foco en cambio, claridad causal).
2. **Elaborar gráficos del árbol de problemas y de objetivos.**
3. **Reestructurar MIR**, reforzando la coherencia vertical y horizontal.
4. **Optimizar indicadores**, asegurando que cumplan con los principios de claridad, medibilidad y utilidad para la toma de decisiones.
5. **Integrar supuestos y medios de verificación** con base en fuentes oficiales y riesgos operacionales.

DeepSeek

1. Población Potencial y Objetivo

Criterios evaluados:

- Delimitación clara y cuantificación.
- Consistencia entre documentos.

Hallazgos:

- **Población potencial:** Se define como los ejecutores del gasto (dependencias, entidades de la APF, Poderes y Entes Autónomos, y Empresas Productivas del Estado). Sin embargo, no se especifica si incluye a todos los niveles de gobierno o solo al federal.

- **Población objetivo:** Se cuantifica en 54 ejecutores (Tabla 4), pero no se explica cómo se seleccionaron ni si existen criterios de elegibilidad.
- **Falta de diferenciación:** No se distingue entre población atendida y potencial en ejercicios fiscales anteriores.

Recomendaciones:

- Clarificar si la población potencial incluye gobiernos estatales/municipales.
- Definir criterios de elegibilidad para la población objetivo.
- Incluir datos históricos de población atendida para comparabilidad.

2. Problema Público

Criterios evaluados:

- Formulación como hecho negativo.
- Vinculación con causas/efectos.
- Sustento empírico.

Hallazgos:

- **Definición:** *"Los ejecutores del gasto requieren asignaciones presupuestarias sujetas a criterios de eficiencia y eficacia..." (p. 7).*
 - **Fortalezas:**
 - Se enfoca en un hecho negativo (ineficiencia en asignaciones).
 - Identifica población afectada (ejecutores del gasto).
 - **Debilidades:**
 - No profundiza en causas estructurales (ej. corrupción, falta de capacitación).
 - Falta contextualización territorial (ej. diferencias regionales en ejecución).

Recomendaciones:

- Reformular el problema para incluir causas raíz:

"La falta de mecanismos estandarizados para asignar recursos públicos bajo criterios de eficiencia, transparencia y rendición de cuentas genera desigualdades en la ejecución del gasto entre dependencias federales, afectando el cumplimiento de los objetivos nacionales."

3. Objetivo del Programa

Criterios evaluados:

- Relación lógica con el problema.
- Población objetivo y cambio esperado.

Hallazgos:

- **Objetivo general:** "Los ejecutores del gasto reciben asignaciones presupuestarias eficientes..." (p. 15).
 - **Fortalezas:**
 - Es único y vinculado al problema.
 - **Debilidades:**
 - No especifica el cambio concreto esperado (ej. reducción de subejercicios).
 - No diferencia por grupos (ej. dependencias con mayor rezago).

Recomendaciones:

- Añadir dimensión de resultado:
"Garantizar que el 100% de las asignaciones presupuestarias a ejecutores del gasto federal cumplan con criterios de eficiencia (medidos por el Índice de Seguimiento del Desempeño) para 2024."

4. Árbol del Problema y Objetivos

Árbol del Problema (Figura 1, p. 13):

- **Causas:**
 - "Divergencia entre necesidades y recursos" (válido).
 - "Desconocimiento de procesos" (ambigua; falta especificar actores).
- **Efectos:**
 - "Incumplimiento de metas" (genérico).

Árbol de Objetivos (Figura 2, p. 14):

- Falta claridad en la relación medios-fines (ej. cómo las "ajustes presupuestarios" logran el propósito).

Recomendaciones:

- Reformular causas/efectos con ejemplos concretos:
 - Causa: "Falta de capacitación en PbR en el 40% de las dependencias."
 - Efecto: "Subejercicio del 15% del gasto programable en 2021."
- Vincular cada objetivo específico con una causa del árbol.

5. Matriz de Indicadores para Resultados (MIR)

Evaluación de indicadores (p. 24-25):

Nivel	Fortalezas	Debilidades	Recomendaciones
-------	------------	-------------	-----------------

Fin	Vinculado al PND.	Indicador genérico: " <i>Cumplimiento promedio de programas</i> " no mide impacto real.	Usar: <i>"% de programas con evaluación positiva en eficiencia del gasto (ISeD ≥ 8/10)."</i>
Propósito	Cociente de gasto programable (medible).	No especifica meta concreta (ej. reducir gasto corriente en X%).	Añadir: <i>"Reducir el gasto corriente en 5% anual (2023-2024)."</i>
Componentes	Oportunidad en PPEF (indicador claro).	Falta indicador de calidad (ej. % de asignaciones alineadas a prioridades).	Incluir: <i>"% de asignaciones vinculadas a programas prioritarios del PND."</i>
Actividades	Seguimiento de proyectos de inversión.	Indicadores de proceso (ej. % FARE enviados) no miden resultados.	Reemplazar con: <i>"% de evaluaciones que derivan en ajustes presupuestarios."</i>

Problemas detectados:

- **Indicadores de proceso vs. resultado:** El 60% son de gestión (ej. "*días promedio de atención*"), no impacto.
- **Medios de verificación:** No se especifican fuentes de datos para todos los indicadores.

Recomendaciones generales para la MIR:

- Añadir indicadores de calidad (ej. satisfacción de ejecutores con asignaciones).
- Incluir supuestos críticos (ej. "*Estabilidad macroeconómica que permita ejecución presupuestaria.*").

Conclusiones

El programa P002 cumple con los requisitos básicos de diseño, pero requiere ajustes para:

1. **Precisar poblaciones** (potencial, objetivo, atendida).
2. **Profundizar en el problema público** (causas/efectos con datos empíricos).
3. **Mejorar la MIR** con indicadores de resultado y metas cuantificables.
4. **Fortalecer la coherencia** entre árbol de problemas, objetivos y MIR.

Propuesta de ajustes clave:

- Reformular el problema incluyendo corrupción y capacidades técnicas como causas.
- Vincular cada componente de la MIR a un efecto del árbol de problemas.
- Añadir indicadores de calidad y eficiencia (ej. % de reducción en subejercicios).

Documentación de apoyo:

- Usar datos del CONEVAL y ASF para sustentar efectos (ej. brechas en ejecución por entidad).
- Referir estándares OCDE para indicadores de gestión pública (ej. "Government at a Glance").

Nota final: Estos ajustes alinearían el P002 con la Metodología de Marco Lógico y mejorarían su evaluabilidad, asegurando que los recursos públicos se asignen con base en resultados medibles.

CLAUDE

1. ANÁLISIS DE LA DEFINICIÓN DE POBLACIONES

1.1 Población Potencial

Definición actual: "La población potencial del P002 se conforma por los ejecutores del gasto que, en el marco de este diagnóstico, incorpora a las dependencias y entidades de la APF, los Poderes y Entes Autónomos, así como a las Empresas Productivas del Estado."

Deficiencias identificadas:

- **Falta de cuantificación específica:** No se proporciona el número total de ejecutores del gasto que conforman la población potencial
- **Delimitación imprecisa:** No establece criterios claros de elegibilidad o características específicas
- **Ausencia de justificación:** No explica por qué estos actores presentan el problema público

Propuesta de mejora: "La población potencial del P002 se conforma por 54 ejecutores del gasto público federal que requieren asignaciones presupuestarias eficientes y eficaces, integrada por: 26 dependencias de la Administración Pública Centralizada, 16 entidades de la Administración Pública Paraestatal, 2 Poderes Autónomos, 8 Entes Autónomos y 2 Empresas Productivas del Estado."

1.2 Población Objetivo

Definición actual: Idéntica a la población potencial, lo cual representa una inconsistencia metodológica.

Deficiencias identificadas:

- **No diferenciación:** La población objetivo es idéntica a la potencial, lo que indica falta de criterios de priorización
- **Ausencia de focalización:** No establece criterios para determinar qué ejecutores serán atendidos prioritariamente
- **Falta de justificación temporal:** No especifica plazos de atención

Propuesta de mejora: Mantener la misma definición pero agregar: "Esta población corresponde al 100% de la población potencial, dado que todos los ejecutores del gasto requieren del servicio de manera obligatoria para el funcionamiento del sistema presupuestario federal."

2. ANÁLISIS DEL PROBLEMA PÚBLICO

2.1 Definición del Problema

Formulación actual: "Los ejecutores del gasto requieren de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles."

Deficiencias identificadas:

- **No es un problema:** Está formulado como una necesidad o requerimiento, no como un hecho negativo
- **Uso de palabras prohibidas:** "Requieren" implica necesidad, lo cual va contra los lineamientos
- **Falta de especificidad:** No describe claramente qué situación negativa enfrentan los ejecutores
- **No identifica un cambio específico:** No establece el resultado deseado sobre la población

Propuesta de reformulación: "Los ejecutores del gasto público federal no cuentan con asignaciones presupuestarias que cumplan con los criterios de eficiencia y eficacia necesarios para el logro de los objetivos del desarrollo nacional."

2.2 Análisis del Árbol del Problema

Fortalezas identificadas:

- Estructura lógica entre causas y efectos
- Identificación de múltiples niveles causales
- Vinculación con el marco institucional

Deficiencias identificadas:

En las causas:

- Algunas causas están formuladas en positivo cuando deberían ser negativas
- "Presiones de gasto programable y no programable" no es específica
- "Ambigüedad normativa" podría ser más precisa

En los efectos:

- La cadena causal podría ser más específica
- Falta un efecto final más contundente relacionado con el desarrollo nacional

Propuestas de mejora para el árbol del problema:

Causas reformuladas:

1. **Causa directa 1:** Desarticulación entre necesidades, prioridades y recursos disponibles
 - 1.1 Objetivos de programas presupuestarios no alineados con la planeación nacional
 - 1.2 Planeación deficiente de programas y proyectos de inversión pública
 - 1.3 Información insuficiente del desempeño para decisiones presupuestarias
2. **Causa directa 2:** Ajustes inadecuados en estructuras programáticas durante el ejercicio

- 2.1 Incumplimiento de disciplina presupuestaria
- 2.2 Procesos de programación y presupuesto no estandarizados
- 2.3 Información financiera heterogénea y poco precisa

3. ANÁLISIS DEL OBJETIVO

3.1 Definición del Objetivo

Formulación actual: "Los ejecutores del gasto reciben de asignaciones presupuestarias sujetas a los criterios de eficiencia y eficacia, para el cumplimiento de los objetivos del desarrollo nacional, en el marco de los recursos públicos disponibles."

Deficiencias identificadas:

- **Error gramatical:** "reciben de" es incorrecto
- **No es contraparte directa:** No corresponde exactamente a la solución del problema identificado
- **Falta de especificidad:** No establece claramente el cambio deseado

Propuesta de reformulación: "Los ejecutores del gasto público federal cuentan con asignaciones presupuestarias que cumplan con los criterios de eficiencia y eficacia para el logro de los objetivos del desarrollo nacional."

4. ANÁLISIS DE LA MATRIZ DE INDICADORES PARA RESULTADOS (MIR)

4.1 Nivel de Fin

Objetivo actual: "Contribuir a la paz, Estado democrático y de derecho mediante el cumplimiento de los objetivos y metas de los programas presupuestarios"

Deficiencias identificadas:

- **Desalineación:** No se conecta directamente con el problema del programa
- **Muy amplio:** El alcance es demasiado general para un programa de diseño de política de gasto
- **Falta de especificidad:** No refleja la contribución específica del programa

Propuesta de mejora: "Contribuir al desarrollo nacional mediante el uso eficiente y eficaz de los recursos públicos federales"

4.2 Nivel de Propósito

Objetivo actual: "Los programas presupuestarios cumplen con los objetivos y metas establecidos"

Deficiencias identificadas:

- **No corresponde al beneficiario directo:** Debería referirse a los ejecutores del gasto
- **Desconexión con el problema:** No refleja el cambio esperado en la población objetivo

Propuesta de mejora: "Los ejecutores del gasto público federal reciben asignaciones presupuestarias eficientes y eficaces"

4.3 Nivel de Componentes

Análisis de los componentes actuales:

Componente A: "Cartera de Programas y Proyectos de Inversión de infraestructura productiva aprobada"

- **Deficiencia:** Muy específico, no abarca toda la función del programa
- **Mejora propuesta:** "Proyecto de Presupuesto de Egresos de la Federación integrado"

Componente B: "Información de desempeño valorada y mejorada"

- **Fortaleza:** Bien formulado como producto entregado
- **Mantener con ajuste menor:** "Información de desempeño de programas presupuestarios generada"

Componente C: "Informes de Contabilidad Gubernamental e Informes de Gestión Pública emitidos"

- **Fortaleza:** Correctamente formulado como producto
- **Mantener:** Sin cambios necesarios

Componente D: "MIR con valoración satisfactoria"

- **Deficiencia:** No es un producto del programa, sino un resultado
- **Mejora propuesta:** "Matrices de Indicadores para Resultados elaboradas"

4.4 Análisis de Indicadores

Deficiencias generales identificadas:

- **Complejidad excesiva:** Algunos indicadores son demasiado complejos para su interpretación
- **Desalineación:** No todos los indicadores miden directamente el objetivo del nivel correspondiente
- **Falta de dimensiones:** No todos cubren cantidad, calidad y tiempo

Propuestas de mejora:

Indicador de Fin:

- **Actual:** Muy complejo y poco claro
- **Propuesto:** "Porcentaje de programas presupuestarios que alcanzan sus metas de desempeño"

Indicador de Propósito:

- **Mantener el primero** con simplificación
- **Eliminar el segundo** por no corresponder al nivel

4.5 Medios de Verificación

Deficiencias identificadas:

- Falta especificación de fuentes para algunos indicadores
- No todos los medios son accesibles o verificables externamente
- Ausencia de periodicidad clara

4.6 Supuestos

Deficiencia crítica:

- **No se presentan supuestos** en la MIR analizada, lo cual es una omisión grave según la MML

Propuesta de supuestos por nivel:

- **Fin:** "Las condiciones macroeconómicas del país se mantienen estables"
- **Propósito:** "Los ejecutores del gasto utilizan adecuadamente las asignaciones presupuestarias"
- **Componentes:** "Las dependencias y entidades proporcionan información oportuna y veraz"

5. RECOMENDACIONES GENERALES

5.1 Coherencia Interna

- Reformular el problema como situación negativa específica
- Alinear objetivo con la solución del problema reformulado
- Ajustar la MIR para reflejar la lógica causal del programa

5.2 Mejoras Metodológicas

- Incluir supuestos en todos los niveles de la MIR
- Simplificar indicadores manteniendo su capacidad de medición
- Especificar claramente medios de verificación y periodicidad

5.3 Fortalecimiento del Diseño

- Establecer criterios claros de focalización, aunque la población objetivo sea universal
- Mejorar la justificación teórica y empírica del problema
- Desarrollar un sistema de seguimiento más robusto

6. CONCLUSIONES

El programa P002 presenta deficiencias significativas en su diseño que afectan la coherencia interna y la capacidad de medición de resultados. Las principales áreas de mejora son:

1. **Reformulación del problema público** para cumplir con los criterios de la MML
2. **Reestructuración de la MIR** para mejorar la alineación entre niveles
3. **Simplificación de indicadores** manteniendo su utilidad para la toma de decisiones
4. **Incorporación de supuestos** para identificar riesgos externos
5. **Mejora en la especificación de poblaciones**, aunque sean universales

Estas mejoras fortalecerán significativamente la capacidad del programa para demostrar su contribución al desarrollo nacional y facilitarán procesos futuros de evaluación y rendición de cuentas.

4.6 Aplicación de la IA en los Modelos de TdR 2025 (Evaluación de Consistencia y Resultados, e Indicadores)

Con base en las disposiciones en materia de evaluación, descritas en el Programa Anual de Evaluación de las Políticas Públicas y Programas presupuestarios de la Administración Pública Federal para el Ejercicio Fiscal 2025, así como los TdR publicados para las evaluaciones que se realicen durante el ejercicio fiscal 2025, se procedió a realizar dos simulaciones en el uso de IA para las evaluaciones de Consistencia y Resultados, e Indicadores, tomando nuevamente como base los programas B001 y P002 de Ramo 6.

4.6.1 Evaluación de Consistencia y Resultados

Para esta simulación se adjuntaron dos documentos, el primero de ellos con los primeros seis reactivos del MTdR para la evaluación de Consistencia y Resultados 2025, los cuales se identificaron como viables para ser respondidos mediante IA; el segundo documento adjunto incluía un resumen con la información relacionada a los Pp B001 y P002

4.6.1.1 Evaluación de Consistencia y Resultados Pp B001

Deep Seek

I. Planeación

1. Contribución estratégica al Plan Nacional de Desarrollo (PND) 2025-2030 y Objetivos de Desarrollo Sostenible (ODS):

El Pp B001 se alinea con los **ejes estratégicos del PND 2025-2030**, particularmente con:

- **Eje 1: "Gobierno eficiente y seguridad jurídica"**, al garantizar la integridad de documentos oficiales, reducir la falsificación y fortalecer la confianza en las instituciones públicas.
- **Eje 3: "Innovación tecnológica y desarrollo industrial"**, mediante la modernización de maquinaria especializada (ej. TREN de pasaportes electrónicos) y la adopción de estándares internacionales de seguridad.

Vinculación con los ODS:

- **ODS 16 (Paz, justicia e instituciones sólidas):**

- Combate la corrupción y el crimen organizado al dificultar la falsificación de documentos (pasaportes, marbetes, etc.).
- Protege datos personales y reduce fraudes financieros o de identidad.
- **ODS 9 (Industria, innovación e infraestructura):**
 - Inversión en tecnología de impresión segura (ej. chips electrónicos, tintas antifalsificación).
 - Colaboración con organismos internacionales (OACI, ISO) para homologar estándares.
- **ODS 3 (Salud y bienestar):**
 - Mitiga riesgos sanitarios al evitar la falsificación de marbetes en bebidas alcohólicas (adulteración).

Aporte a la misión de la SHCP:

El programa cumple con las facultades legales de la **DGTIEV** (artículos 42 M, N, O del RISHCP) para producir impresos con medidas de seguridad, asegurando que dependencias federales operen con documentos confiables y bajo principios de **transparencia y austeridad**.

Sinergias con políticas transversales del PND 2025-2030:

- **Seguridad nacional:** Al resguardar documentos sensibles (ej. pasaportes, PEF) en instalaciones clasificadas "AAA".
- **Transformación digital:** Integración de elementos electrónicos (ej. chips RFID) en impresos, en línea con la Estrategia Nacional de Ciberseguridad.

II. Diseño

2. Teoría del cambio y generación de valor público:

1. Marco teórico y evidencia empírica:

- El diagnóstico fundamenta la intervención en normativas internacionales (OACI, ISO) y nacionales (Código Penal Federal). Ejemplo: Actualización de pasaportes electrónicos según estándares OACI (Documento 9303).
- Evidencia empírica: Estadísticas delictivas (752 averiguaciones por falsificación en 2017) y casos documentados (falsificación de marbetes, pasaportes).

2. Contexto:

- Considera el entorno político (coordinación con SRE, SAT), económico (restricciones presupuestarias), social (impacto en salud por bebidas adulteradas) e institucional (clasificación "AAA" de seguridad).

3. Lógica causal:

- **Problema:** Falsificación de documentos oficiales → **Medios:** Maquinaria especializada, normativas actualizadas, controles de confidencialidad → **Objetivo:** Reducir riesgos de falsificación.
- **Crítica:** La dependencia de proveedores externos para insumos especializados podría ser un cuello de botella.

4. Alternativas:

- Externalizar producción (ej. Brasil con imprentas privadas) fue descartada por riesgos de fraude y falta de control estatal.

3. Árboles de problema y objetivos (MML):

• **Árbol del problema:**

- **Problema central:** Dependencias y Entidades de la APF enfrentan altos riesgos de falsificación en sus impresos.
- **Causas:** Debilidad en estándares de seguridad, maquinaria obsoleta, rezago normativo.
- **Efectos:** Delitos penales, evasión fiscal, daños a la salud.

• **Árbol de objetivos:**

- **Fines:** Prevención de delitos y daños a la salud.
- **Objetivo central:** Reducir riesgos de falsificación.
- **Medios:** Cumplir estándares, modernizar maquinaria, actualizar normativas.

Consistencia: Los bienes (impresos seguros) son necesarios y suficientes para el objetivo. El problema está bien definido (población + problema público).

III. Seguimiento

4. Instrumento de Seguimiento del Desempeño (ISD):

- **Alineación con árbol de objetivos:** Los indicadores miden resultados clave (ej. porcentaje de impresos que cumplen estándares de seguridad).
- **Indicadores relevantes:**
 - **Gestión:** Cumplimiento de programas de mantenimiento (1111).
 - **Resultados:** % de impresos entregados con seguridad (Indicador 1).
- **Medios de verificación:** Reportes de producción y órdenes de trabajo (aunque algunos son confidenciales).
- **Supuestos:** Disponibilidad de técnicos especializados para mantenimiento.

Crítica: Falta un indicador de impacto directo (ej. reducción de delitos por falsificación).

IV. Población y Cobertura

5. Poblaciones potencial, objetivo y atendida:

- **Potencial:** 230 Dependencias y Entidades de la APF (según LOAPF).
- **Objetivo:** Dependencias que solicitan impresos con elementos de seguridad (ej. SRE, SAT).

- **Atendida (2022):** 6 entidades (3.91% de cobertura).

Métodos de cálculo:

- Actualización bianual basada en la LOAPF y DOF.
- Criterios de elegibilidad: Convenios de colaboración y solicitudes formales.

6. Cobertura adecuada y estrategia futura:

- **Barreras:** Restricciones presupuestarias, maquinaria obsoleta (ej. TREN con vida útil vencida).
- **Estrategia 2024-2026:**
 - Proyección de atención al 5% anual (considerando capacidad instalada).
 - Inversión en nuevo TREN para aumentar capacidad en 220%.
- **Factibilidad:** Depende de asignación de recursos (PPI "Sustitución de Maquinaria").

Recomendación: Ampliar alianzas con más dependencias para incrementar cobertura.

4.6.1.2 Evaluación de Consistencia y Resultados Pp P002

Deep Seek

I. Planeación

1. Contribución estratégica al Plan Nacional de Desarrollo 2025-2030 (PND), Objetivos de Desarrollo Sostenible (ODS) y vinculación con la misión de la dependencia:

El programa **P002 "Diseño y conducción de la política de gasto"** se alinea con los ejes estratégicos del **PND 2025-2030**, en particular con los objetivos vinculados a:

- **Transformación económica sostenible e inclusiva**, mediante la asignación eficiente de recursos públicos que prioricen proyectos de inversión productiva, infraestructura estratégica (ej. Corredor Interoceánico) y desarrollo tecnológico, en línea con la meta de crecimiento económico del 3.5% anual (pág. 10).
- **Reducción de desigualdades y bienestar social**, al direccionar el gasto público hacia programas sociales focalizados (ej. becas, salud universal) que atiendan a la población en pobreza extrema (8.5% en 2020, pág. 9), en concordancia con el principio "*Por el bien de todos, primero los pobres*".
- **Gobernanza democrática y combate a la corrupción**, reforzando mecanismos de transparencia y rendición de cuentas (ej. Sistema de Contabilidad Gubernamental) para garantizar el uso eficaz de los recursos, bajo los criterios de "*austeridad republicana*" (pág. 12).

Vinculación con los ODS:

- **ODS 16 (Paz, justicia e instituciones sólidas):** Promueve la gestión fiscal transparente y el fortalecimiento institucional a través de la evaluación del desempeño (pág. 12, Índice de Gestión para Resultados del BID).

- **ODS 10 (Reducción de desigualdades):** Prioriza el gasto social (ej. 3.4 billones de pesos en 2022, pág. 8) para cerrar brechas económicas y territoriales.

Relación con la misión de la Subsecretaría de Egresos (SSE):

El programa opera bajo el mandato constitucional (artículo 134 CPEUM) y legal (LFPRH) de la SHCP, asegurando que el **Presupuesto de Egresos de la Federación (PEF)** refleje las prioridades del PND 2025-2030. Su enfoque en "*eficiencia, austeridad y resultados*" (pág. 6) fortalece la capacidad del Estado para cumplir con sus metas de desarrollo, en coordinación con dependencias federales, entidades autónomas y empresas públicas (pág. 17-20).

II. Diseño

2. Teoría del cambio y generación de valor público:

- **Marco teórico y evidencia:** El diseño del PO02 se sustenta en normativas como la LFPRH y la LGCG, y en experiencias internacionales (ej. Índice de Gestión para Resultados del BID). La evidencia empírica incluye datos históricos del gasto público (ej. incremento del 8.6% en gasto programable 2018-2021) y evaluaciones de desempeño.
- **Contexto:** Considera el entorno político (austeridad), económico (restricciones fiscales) y social (enfoque en población vulnerable).
- **Lógica causal:** Las acciones del programa (ej. integración del PPEF, adecuaciones presupuestarias) buscan resolver el problema de asignaciones ineficientes mediante mecanismos como el Módulo de Adecuaciones Presupuestarias (MAP) y alineación con el PND.
- **Alternativas:** No se analizan alternativas debido a que el programa está mandatado por la CPEUM y LFPRH (pág. 21).

3. Congruencia de los árboles de problema y objetivos (Metodología de Marco Lógico):

- **Árbol del problema (Figura 1, pág. 13):** Raíz: "Los ejecutores del gasto requieren asignaciones eficientes". Causas: Desalineación en planeación, falta de disciplina presupuestaria. Efectos: Incumplimiento de metas de la APF.
- **Árbol de objetivos (Figura 2, pág. 10):** Fin: "Ejecutores reciben asignaciones eficientes". Medios: Integración del PPEF, normatividad para ejercicio del gasto.
- **Consistencia:** Los objetivos mitigan las causas del problema (ej. normatividad → disciplina presupuestaria). Los bienes/servicios (ej. PPEF, informes de desempeño) son necesarios para alcanzar el objetivo central.

III. Seguimiento

4. Instrumento de Seguimiento del Desempeño (ISD) y alineación con Marco Lógico:

- **Alineación con árbol de objetivos:** La Matriz de Indicadores para Resultados (MIR, pág. 24-25) vincula indicadores con cada nivel (Fin, Propósito, Componente, Actividad). Ejemplo:
 - *Fin:* "Cumplimiento promedio de programas del PND" mide impacto en metas nacionales.
 - *Propósito:* "Cociente de recomposición del gasto programable" evalúa eficiencia.

- **Indicadores:** Cubren gestión operativa (ej. días promedio para adecuaciones), cobertura (ej. alineación con PND) y resultados (ej. impacto en servicios personales).
- **Medios de verificación:** Sistemas como el PIPP y reportes de la Cuenta Pública.

IV. Población y Cobertura

5. Identificación de poblaciones potencial, objetivo y atendida:

- **Población potencial (pág. 17-18):** Todos los ejecutores del gasto (APF, Poderes, Entes Autónomos, Empresas del Estado).
- **Población objetivo (Tabla 4, pág. 19-20):** 54 ejecutores específicos (ej. Ramo 06: SHCP, Ramo 52: Pemex).
- **Métodos de cálculo:** Actualización anual mediante catálogos de Ramos y Unidades Responsables (pág. 21).

6. Adecuación de la cobertura:

- **Población atendida:** No se especifica en el documento, pero se proyecta cubrir al 100% de la población objetivo (54 ejecutores).
- **Barreras:** Obstáculos administrativos (ej. trámites complejos) y limitaciones financieras (pág. 10: contexto de austeridad).
- **Estrategia de cobertura (pág. 22):** Mejora en sistemas de información (ej. PIPP) y disposiciones normativas para simplificar procesos.

4.6.2 Evaluación de Indicadores

Para esta simulación se adjuntaron dos documentos, el primero de ellos con los dieciocho reactivos correspondientes a la etapa de valoración inicial del MTdR para la evaluación de Consistencia y Resultados 2025, los cuales se identificaron como viables para ser respondidos mediante IA; el segundo documento adjunto se conformaba por las Matrices de Indicadores para Resultados (MIR) de los programas B001 y P002, respectivamente, reportadas en el portal de Transparencia, para el ejercicio fiscal 2025.

4.6.2.1 Evaluación de Indicadores Pp B001

Chat GPT

A continuación, se presenta la respuesta estructurada a las 18 preguntas del documento "MTDR_Indicadores_2025", aplicada a los indicadores de la MIR 2025 del programa presupuestario no social B001. El análisis se realiza para cada indicador considerando los criterios de claridad, relevancia, economía, monitoreabilidad, adecuación y aportación marginal.

INDICADOR 1: Porcentaje de eficiencia de los procesos productivos (Fin)

Claridad

1. Sí. El nombre expresa adecuadamente lo que mide.
2. Sí. La fórmula es congruente con el concepto de eficiencia.
3. Sí. La definición refleja la naturaleza del indicador.

4. Sí. La definición y la fórmula están alineadas.
5. Sí. Las unidades son congruentes (monetarias y porcentuales).
6. Parcialmente. Faltan definiciones claras de las variables involucradas.

Relevancia

7. Sí. Permite valorar el logro del objetivo de eficiencia institucional.
8. Parcialmente. La meta está alineada con la frecuencia anual, pero no se indica el valor.

Economía

9. 2. Completamente. La utilidad justifica el esfuerzo.

Monitoreabilidad

10. Parcialmente. Requiere verificación de consistencia en inventarios y gasto.
11. Parcialmente. La congruencia depende de la trazabilidad de los datos contables.
12. Parcialmente. No se especifican fuentes documentales exactas.
13. No. No se indica ubicación física o electrónica.
14. Parcialmente. La suficiencia de los medios debe validarse.

Adecuación

15. Sí. Mide resultados institucionales del nivel Fin.
16. Parcialmente. No se señala la meta anual ni línea base explícita.
17. Sí. Está bien identificado como eficiencia.

Aportación Marginal

18. 2. Aporta valor y no es redundante.

INDICADOR 2: Porcentaje de satisfacción del cliente (Propósito)

Claridad

1. Sí. El nombre es claro y específico.
2. Sí. La fórmula corresponde al nombre.
3. Sí. La definición es congruente.
4. Sí. Existe congruencia completa.
5. Sí. Las unidades son consistentes.
6. Parcialmente. No se detallan los criterios de aceptación en calidad y oportunidad.

Relevancia

7. Sí. Mide directamente el objetivo de servicio al cliente.
8. Sí. La meta puede evaluarse semestralmente.

Economía

9. 2. Completamente. Información estratégica justifica el esfuerzo.

Monitoreabilidad

10. Sí. La frecuencia es semestral y razonable.
11. Parcialmente. Requiere revisión de criterios técnicos establecidos.
12. No. No se especifican fuentes completas.

13. No. No se indica ubicación.
14. Parcialmente. Requiere verificar suficiencia de registros de entrega.

Adecuación

15. Sí. Permite evaluar cumplimiento de propósito.
16. Parcialmente. Falta visibilidad de meta y línea base.
17. Sí. Se trata de eficacia.

Aportación Marginal

18. 2. Información valiosa y única.

INDICADOR 3A: % cumplimiento impresos diversos (Componente A)

Claridad

1. Sí. El nombre es descriptivo.
2. Sí. La fórmula es directa y coherente.
3. Sí. La definición es clara.
4. Sí. Congruencia total.
5. Sí. Unidades compatibles.
6. Parcialmente. No se explican criterios de producción y solicitud.

Relevancia

7. Sí. Evalúa la entrega de productos clave.
8. Sí. Frecuencia y meta son compatibles.

Economía

9. 2. Alta relación costo-beneficio.

Monitoreabilidad

10. Sí. Frecuencia semestral viable.
11. Parcialmente. Requiere claridad sobre fuente de solicitudes.
12. No. Falta nombre completo de fuentes.
13. No. No se describe ubicación.
14. Parcialmente. Se asume suficiencia pero no se documenta.

Adecuación

15. Sí. Evalúa desempeño operativo.
16. Parcialmente. No se presenta meta ni línea base.
17. Sí. Es de eficacia.

Aportación Marginal

18. 2. No es redundante.

INDICADOR 3B: % cumplimiento impresos valorados (Componente B)

Claridad

1. Sí.

2. Sí.
3. Sí.
4. Sí.
5. Sí.
6. Parcialmente. Faltan definiciones claras.

Relevancia

7. Sí.
8. Sí.

Economía

9. 2.

Monitoreabilidad

10. Sí.
11. Parcialmente.
12. No.
13. No.
14. Parcialmente.

Adecuación

15. Sí.
16. Parcialmente.
17. Sí.

Aportación Marginal

18. 2.

INDICADOR 4A: % mantenimiento maquinaria (Actividad A1)

Claridad

1. Sí.
2. Sí.
3. Sí.
4. Sí.
5. Sí.
6. Parcialmente.

Relevancia

7. Sí.
8. Sí.

Economía

9. 2.

Monitoreabilidad

10. Sí.
11. Parcialmente.

- 12. No.
- 13. No.
- 14. Parcialmente.

Adecuación

- 15. Sí.
- 16. Parcialmente.
- 17. Sí.

Aportación Marginal

- 18. 2.

INDICADOR 4B: % solicitudes de compra atendidas (Actividad B2)

Claridad

- 1. Sí.
- 2. Sí.
- 3. Sí.
- 4. Sí.
- 5. Sí.
- 6. Parcialmente.

Relevancia

- 7. Sí.
- 8. Sí.

Economía

- 9. 2.

Monitoreabilidad

- 10. Sí.
- 11. Parcialmente.
- 12. No.
- 13. No.
- 14. Parcialmente.

Adecuación

- 15. Sí.
- 16. Parcialmente.
- 17. Sí.

Aportación Marginal

- 18. 2.

DeepSeek

A continuación, se responde a las preguntas planteadas en el documento "MTDR_Indicadores_2025" para cada uno de los indicadores del programa B001, utilizando la

información proporcionada en la "MIR B001". Las respuestas se organizan según los criterios de **Claridad, Relevancia, Economía, Monitoreabilidad, Adecuación y Aportación Marginal**.

Indicador: Porcentaje de eficiencia de los procesos productivos (Nivel Fin)

1. Claridad

○ Preguntas 1-6:

- El nombre del indicador es claro y autoexplicativo, ya que refleja la unidad de medida (porcentaje) y el concepto de eficiencia.
- La fórmula de cálculo es congruente con el nombre y la definición del indicador.
- Las variables están bien descritas (producción alcanzada, gastos de operación, inventarios).
- Las unidades de medida son consistentes (todas en términos monetarios o físicos para el cálculo del porcentaje).
- **Calificación:** 2 (Completamente).

2. Relevancia

○ Preguntas 7-8:

- El indicador contribuye claramente al objetivo de "Regeneración Ética" al medir la eficiencia en el uso de recursos.
- La meta anual está implícita en la naturaleza del indicador, aunque no se especifica un valor numérico en la MIR.
- **Calificación:** 2 (Completamente).

3. Economía

○ Pregunta 9:

- El indicador justifica el esfuerzo de cálculo, ya que evalúa un aspecto crítico del programa (eficiencia).
- **Calificación:** 2 (Completamente).

4. Monitoreabilidad

○ Preguntas 10-14:

- La frecuencia de medición (anual) es adecuada para un indicador estratégico.
- No se especifican los medios de verificación, pero se asume que los datos provienen de registros contables y de producción.
- **Calificación:** 1 (Parcialmente, por falta de detalle en medios de verificación).

5. Adecuación

- **Preguntas 15-17:**
 - El indicador es adecuado para evaluar el desempeño del programa a nivel estratégico (Fin).
 - La dimensión (Eficiencia) está correctamente identificada.
 - **Calificación:** 2 (Completamente).

6. Aportación Marginal

- **Pregunta 18:**
 - El indicador aporta información valiosa y no redundante sobre la eficiencia del programa.
 - **Calificación:** 2 (Completamente).

Indicador: Porcentaje de satisfacción del cliente (Nivel Propósito)

1. Claridad

- El nombre y la fórmula son claros y congruentes.
- **Calificación:** 2.

2. Relevancia

- Mide directamente el objetivo de entregar impresos conforme a los requerimientos.
- **Calificación:** 2.

3. Economía

- El esfuerzo de cálculo (encuestas o registros de aceptación) justifica la utilidad del indicador.
- **Calificación:** 2.

4. Monitoreabilidad

- Frecuencia semestral adecuada, pero no se detallan medios de verificación.
- **Calificación:** 1.

5. Adecuación

- Adecuado para evaluar resultados intermedios (Propósito).
- Dimensión (Eficacia) correcta.
- **Calificación:** 2.

6. Aportación Marginal

- Aporta información única sobre la percepción del cliente.
- **Calificación:** 2.

Indicadores de Componentes (Impresos diversos, valorados, numerados y de seguridad)

1. Claridad

- Nombres y fórmulas son claros y congruentes.
- **Calificación:** 2.

2. Relevancia

- Miden cumplimiento de requerimientos específicos de producción.
- **Calificación:** 2.

3. Economía

- Justifican el esfuerzo al evaluar productos clave.
- **Calificación:** 2.

4. Monitoreabilidad

- Frecuencia semestral adecuada, pero falta detalle en medios de verificación.
- **Calificación:** 1.

5. Adecuación

- Adecuados para evaluar productos intermedios (Componentes).
- Dimensión (Eficacia) correcta.
- **Calificación:** 2.

6. Aportación Marginal

- Cada indicador aporta información valiosa y no redundante.
- **Calificación:** 2.

Indicadores de Actividades (Mantenimiento preventivo, Adquisición de insumos)

1. Claridad

- Nombres y fórmulas son claros.
- **Calificación:** 2.

2. Relevancia

- Miden aspectos críticos de la gestión operativa.
- **Calificación:** 2.

3. Economía

- Justifican el esfuerzo al evaluar procesos clave.
- **Calificación:** 2.

4. Monitoreabilidad

- Frecuencia trimestral adecuada, pero falta detalle en medios de verificación.
- **Calificación:** 1.

5. **Adecuación**

- Adecuados para evaluar gestión (Actividades).
- Dimensión (Eficacia) correcta.
- **Calificación:** 2.

6. **Aportación Marginal**

- Aportan información valiosa y no redundante.
- **Calificación:** 2.

Observaciones Generales

- **Fortalezas:**
 - Los indicadores son claros, relevantes y adecuados para cada nivel de la MIR.
 - Las dimensiones (eficacia, eficiencia) están bien identificadas.
- **Áreas de mejora:**
 - Detallar los medios de verificación para fortalecer la monitoreabilidad.
 - Especificar metas numéricas para facilitar la evaluación del desempeño.

Conclusión

Los indicadores del programa B001 cumplen con los criterios de calidad, coherencia y pertinencia en su mayoría. Se recomienda complementar la información sobre medios de verificación y metas para optimizar su utilidad en la evaluación.

Claude

Evaluación de Indicadores de la MIR del Programa B001

Secretaría de Hacienda y Crédito Público

Información General del Programa

Objetivo General: Contribuir a la Regeneración Ética de las Instituciones, mediante la producción de impresos que requieran las Dependencias y Entidades de la Administración Pública Federal bajo los principios de austeridad, honestidad y transparencia.

INDICADOR 1: NIVEL FIN

Denominación: Porcentaje de eficiencia de los procesos productivos

Evaluación por Criterios

CLARIDAD

1. **¿El nombre del indicador es autoexplicativo?**
 - **Calificación: 1 (Parcialmente)**
 - **Justificación:** El nombre menciona "eficiencia" pero no especifica claramente que se refiere a eficiencia económica o productiva. La unidad de medida (porcentaje) está implícita.
2. **¿La fórmula de cálculo es congruente con su nombre?**
 - **Calificación: 1 (Parcialmente)**
 - **Justificación:** La fórmula es compleja y no es evidente cómo mide "eficiencia". Parece medir más bien rentabilidad o productividad económica.
3. **¿La definición del indicador es congruente con su nombre?**
 - **Calificación: 0 (Nada)**
 - **Justificación:** No se proporciona definición explícita del indicador en el documento.
4. **¿La definición es congruente con su fórmula de cálculo?**
 - **Calificación: 0 (Nada)**
 - **Justificación:** Al no existir definición, no puede evaluarse esta congruencia.
5. **¿Las unidades de medida de las variables son congruentes?**
 - **Calificación: 1 (Parcialmente)**
 - **Justificación:** Todas las variables parecen estar en términos monetarios, pero la fórmula es confusa en su estructura.
6. **¿La descripción de las variables permite entender todos los elementos?**
 - **Calificación: 0 (Nada)**
 - **Justificación:** No hay descripción detallada de las variables utilizadas en la fórmula.

RELEVANCIA

7. **¿Se identifica claramente su contribución al objetivo?**
 - **Calificación: 1 (Parcialmente)**
 - **Justificación:** Relaciona aspectos económicos con el objetivo de eficiencia institucional, pero la conexión no es completamente clara.
8. **¿Tiene meta congruente con su frecuencia de medición?**
 - **Calificación: 0 (Nada)**
 - **Justificación:** No se especifica meta en el documento proporcionado.

ECONOMÍA

9. **¿La utilidad justifica el costo de cálculo?**

- **Calificación: 1 (Parcialmente)**
- **Justificación:** La información de eficiencia es relevante, pero la complejidad del cálculo puede ser excesiva.

MONITOREABILIDAD

10-14. Preguntas sobre medios de verificación:

- **Calificación: 0 (Nada)**
- **Justificación:** No se proporcionan medios de verificación en el documento.

ADECUACIÓN

15. ¿Permite valorar el desempeño del nivel Fin?

- **Calificación: 1 (Parcialmente)**
- **Justificación:** Mide aspectos económicos relevantes para el fin, pero no captura completamente la "regeneración ética".

16. ¿Tiene meta anual alcanzable?

- **Calificación: 0 (Nada)**
- **Justificación:** No se proporciona información sobre metas.

17. ¿La dimensión está bien identificada?

- **Calificación: 2 (Completamente)**
- **Justificación:** Correctamente identificado como Estratégico-Eficiencia-Anual.

APORTACIÓN MARGINAL

18. ¿Aporta información valiosa?

- **Calificación: 2**
- **Justificación:** La información es valiosa y no redundante con otros indicadores.

INDICADOR 2: NIVEL PROPÓSITO

Denominación: Porcentaje de satisfacción del cliente

Evaluación por Criterios

CLARIDAD

1. ¿El nombre del indicador es autoexplicativo?

- **Calificación: 2 (Completamente)**
- **Justificación:** El nombre es claro, autoexplicativo y especifica la unidad de medida.

2. ¿La fórmula de cálculo es congruente con su nombre?

- **Calificación: 2 (Completamente)**

- **Justificación:** La fórmula mide efectivamente la proporción de impresos aceptados vs entregados.

3-6. Otras preguntas de claridad:

- **Calificación: 0 (Nada)**
- **Justificación:** Falta información detallada sobre definiciones y variables.

RELEVANCIA

7. ¿Se identifica claramente su contribución al objetivo?

- **Calificación: 2 (Completamente)**
- **Justificación:** Directamente relacionado con que las dependencias reciban impresos según características técnicas requeridas.

8. ¿Tiene meta congruente con su frecuencia?

- **Calificación: 0 (Nada)**
- **Justificación:** No se especifica meta.

ADECUACIÓN

15. ¿Permite valorar el desempeño del nivel Propósito?

- **Calificación: 2 (Completamente)**
- **Justificación:** Mide directamente los resultados del programa en términos de satisfacción del cliente.

17. ¿La dimensión está bien identificada?

- **Calificación: 2 (Completamente)**
- **Justificación:** Correctamente identificado como Estratégico-Eficacia-Semestral.

INDICADOR 3: COMPONENTE A

Denominación: Porcentaje de cumplimiento de los requerimientos de producción de impresos diversos

Evaluación por Criterios

CLARIDAD

1. ¿El nombre del indicador es autoexplicativo?

- **Calificación: 2 (Completamente)**
- **Justificación:** El nombre es claro y específico sobre lo que mide.

2. ¿La fórmula de cálculo es congruente con su nombre?

- **Calificación: 2 (Completamente)**
- **Justificación:** La fórmula mide directamente el cumplimiento de producción vs solicitudes.

RELEVANCIA

7. ¿Se identifica claramente su contribución al objetivo?

- **Calificación: 2 (Completamente)**
- **Justificación:** Directamente relacionado con la entrega del componente de impresos diversos.

ADECUACIÓN

15. ¿Permite valorar el desempeño del nivel Componente?

- **Calificación: 2 (Completamente)**
- **Justificación:** Mide la entrega de bienes (impresos diversos) del programa.

INDICADOR 4: COMPONENTE B

Denominación: Porcentaje de cumplimiento de los requerimientos de producción de impresos valorados, numerados y de seguridad

Evaluación por Criterios

Similar al Componente A, con alta calificación en claridad, relevancia y adecuación para medir la entrega del segundo componente del programa.

INDICADORES DE ACTIVIDADES

ACTIVIDAD A1: Porcentaje de mantenimiento preventivo a maquinaria de producción

ACTIVIDAD B2: Porcentaje de solicitudes de compra atendidas

Ambos indicadores muestran:

- **Claridad:** Nombres autoexplicativos y fórmulas congruentes
- **Relevancia:** Directamente relacionados con actividades de gestión necesarias
- **Adecuación:** Apropriados para nivel de actividades (gestión)

RESUMEN DE RECOMENDACIONES

Principales Debilidades Identificadas:

1. **Falta de definiciones:** Ningún indicador cuenta con definiciones claras
2. **Ausencia de medios de verificación:** No se especifican fuentes de información
3. **Falta de metas:** No se establecen metas para ningún indicador
4. **Indicador de Fin complejo:** La fórmula del indicador de fin es excesivamente compleja

Recomendaciones Específicas:

1. **Desarrollar definiciones claras** para todos los indicadores
2. **Establecer medios de verificación** específicos y accesibles
3. **Definir metas anuales** alcanzables para cada indicador
4. **Simplificar la fórmula del indicador de Fin** o proporcionar mejor explicación
5. **Incluir fichas técnicas** completas para cada indicador

Fortalezas del Sistema:

- Estructura lógica de la MIR con niveles bien diferenciados
- Indicadores de componentes y actividades bien alineados
- Dimensiones apropiadamente identificadas
- Frecuencias de medición adecuadas para cada nivel

4.6.2.2 Evaluación de Indicadores Pp P002

Chat GPT

A continuación, se presenta el análisis técnico de los indicadores definidos en la MIR 2025 del programa presupuestario no social **P002 “Diseño y Conducción de la Política de Gasto Público”**, con base en los 18 criterios de la Matriz de Términos de Referencia para indicadores. Se abordan de forma individual conforme al nivel de la MIR (Fin, Propósito, Componente, Actividad).

INDICADOR 1: Cumplimiento promedio de los Programas de la Planeación Nacional del Desarrollo (Fin)

Claridad

1. Sí. El nombre comunica adecuadamente el objeto de medición.
2. Sí. La fórmula corresponde con el promedio de cumplimiento.
3. Sí. La definición es coherente con el propósito del indicador.
4. Sí. Existe congruencia entre definición y fórmula.
5. Sí. Las unidades son consistentes con la operación de la fórmula.
6. Parcialmente. No se detallan los criterios de cumplimiento de los programas.

Relevancia

7. Sí. Está alineado con el objetivo de cumplimiento del desarrollo nacional.
8. Parcialmente. No se explicita la meta anual ni su justificación.

Economía

9. 2. Justifica plenamente su costo en función de su relevancia estratégica.

Monitoreabilidad

10. Parcialmente. La frecuencia anual puede generar rezago de información.
11. Parcialmente. Se asume congruencia sin evidencia documental.
12. No. No se identifican fuentes específicas.
13. No. Falta mención de ubicación de fuentes.
14. Parcialmente. Se infiere suficiencia, pero requiere validación.

Adecuación

15. Sí. Refleja resultados a nivel Fin.
16. Parcialmente. No se indica meta ni línea base explícita.
17. Sí. Dimensión de eficacia.

Aportación Marginal

18. 2. Aporta información clave no redundante.

INDICADOR 2: Cociente de recomposición del gasto programable (Propósito)

Claridad

1. Sí.
2. Sí.
3. Sí.
4. Sí.
5. Sí.
6. Parcialmente. No se explican los criterios para clasificar el gasto corriente e inversión.

Relevancia

7. Sí. Mide eficiencia presupuestaria esperada.
8. Parcialmente. La frecuencia anual limita capacidad de ajuste intraejercicio.

Economía

9. 2. Alta utilidad analítica.

Monitoreabilidad

10. Sí.
11. Sí.
12. No.
13. No.
14. Parcialmente.

Adecuación

15. Sí.
16. Parcialmente.
17. Sí.

Aportación Marginal

18. 2.

INDICADOR 3A: % de adecuaciones presupuestarias autorizadas (Componente A)

Claridad

1. Sí.
2. Sí.
3. Sí.
4. Sí.
5. Sí.
6. Parcialmente.

Relevancia

7. Sí.
8. Sí.

Economía

9. 2.

Monitoreabilidad

10. Sí.
11. Parcialmente.
12. No.
13. No.
14. Parcialmente.

Adecuación

15. Sí.
16. Parcialmente.
17. Sí.

Aportación Marginal

18. 2.

INDICADOR 3B.1: % elementos PPEF entregados oportunamente

Claridad

1. Sí.
2. Sí.
3. Sí.
4. Sí.
5. Sí.
6. Parcialmente.

Relevancia

7. Sí.
8. Sí.

Economía

9. 2.

Monitoreabilidad

- 10. Parcialmente.
- 11. Parcialmente.
- 12. No.
- 13. No.
- 14. Parcialmente.

Adecuación

- 15. Sí.
- 16. Parcialmente.
- 17. Sí (Calidad).

Aportación Marginal

- 18. 2.

INDICADOR 3B.2: % objetivos prioritarios cubiertos por programas presupuestarios**Claridad**

- 1. Sí.
- 2. Sí.
- 3. Sí.
- 4. Sí.
- 5. Sí.
- 6. Parcialmente.

Relevancia

- 7. Sí.
- 8. Parcialmente.

Economía

- 9. 2.

Monitoreabilidad

- 10. Parcialmente.
- 11. Parcialmente.
- 12. No.
- 13. No.
- 14. Parcialmente.

Adecuación

- 15. Sí.
- 16. Parcialmente.
- 17. Sí.

Aportación Marginal

- 18. 2.

INDICADORES DE ACTIVIDAD (A1, A2, B3, B4, B5, B6)

Cada uno de estos indicadores (7 en total) cumple en general con los principios de claridad (1-6), aunque presentan deficiencias en los medios de verificación (preguntas 12-14). Se resumen observaciones comunes:

- **Claridad:** En general alta, aunque requieren definiciones más precisas para las variables.
- **Relevancia:** Adecuada. Evalúan funciones críticas del programa como normatividad, seguimiento del desempeño, inversión pública y disciplina presupuestaria.
- **Economía:** Justificada por la centralidad de las actividades evaluadas.
- **Monitoreabilidad:** Débilmente documentada. Urge establecer fuentes específicas y su localización.
- **Adecuación:** Coherentes con el nivel de Actividad. La mayoría carece de línea base y metas anuales explícitas.
- **Aportación marginal:** Alta (valor único, no redundante).

DeepSeek

Etapas de valoración inicial

Claridad

1. **¿El nombre del indicador es autoexplicativo (expresa correctamente la unidad de medida, no utiliza acrónimos o los define de manera precisa)?**
 - **Respuesta:** Parcialmente (1).
 - **Análisis:** Algunos nombres de indicadores son claros (ej. "Cumplimiento promedio de los Programas de la Planeación Nacional del Desarrollo"), pero otros utilizan acrónimos como "Pp" o "APF" sin definición explícita.
 - **Evidencia:** Tabla de indicadores en "MIR P002.docx".
2. **¿La fórmula de cálculo del indicador es congruente con su nombre?**
 - **Respuesta:** Completamente (2).
 - **Análisis:** Las fórmulas reflejan lo que el indicador pretende medir (ej. porcentaje de cumplimiento, cociente de impacto).
 - **Evidencia:** Columnas "**Método de cálculo**" y "**Denominación**".
3. **¿La definición del indicador es congruente con su nombre?**
 - **Respuesta:** Completamente (2).
 - **Análisis:** Las definiciones en "**Denominación**" coinciden con los nombres y objetivos (ej. "Porcentaje de adecuaciones presupuestarias autorizadas").
4. **¿La definición del indicador es congruente con su fórmula de cálculo?**

- **Respuesta:** Completamente (2).
 - **Análisis:** Las fórmulas son coherentes con las definiciones (ej. "Sumatoria del porcentaje de cumplimiento promedio / Número de programas").
5. **¿Las unidades de medida de las variables de la fórmula son congruentes?**
- **Respuesta:** Completamente (2).
 - **Análisis:** Las unidades (porcentajes, montos ejercidos) son consistentes y permiten cálculos válidos.
6. **¿La descripción de las variables de la fórmula permite entender todos los elementos que incluye?**
- **Respuesta:** Parcialmente (1).
 - **Análisis:** Algunas variables son claras (ej. "Monto ejercido del gasto"), pero otras como "Pp" o "ISeD" carecen de explicación detallada.

Relevancia

7. **¿El indicador contribuye al logro del objetivo asociado y verifica el valor público?**
- **Respuesta:** Completamente (2).
 - **Análisis:** Los indicadores miden resultados clave (ej. eficacia en gasto, alineación con el PND).
 - **Evidencia:** Objetivos como "Contribuir al cumplimiento de los objetivos del desarrollo nacional".
8. **¿La meta es congruente con la frecuencia de medición y alineada con los recursos?**
- **Respuesta:** Parcialmente (1).
 - **Análisis:** No se especifican metas en la MIR, pero la frecuencia (anual/trimestral) parece adecuada para los cálculos.

Economía

9. **¿La utilidad del indicador justifica el costo de su cálculo?**
- **Respuesta:** Completamente (2).
 - **Análisis:** Los indicadores son críticos para evaluar eficiencia presupuestaria y cumplimiento de políticas públicas.

Monitoreabilidad

10. **¿La frecuencia de los medios de verificación es consistente con la del indicador?**
- **Respuesta:** Completamente (2).

- **Análisis:** La frecuencia (anual/trimestral) coincide con los periodos reportados (ej. informes anuales de cumplimiento).

11. ¿Los medios de verificación son congruentes con lo que mide el indicador?

- **Respuesta:** Parcialmente (1).
- **Análisis:** No se detallan medios específicos, pero se infieren fuentes como reportes de SHCP.

12-14. Detalles de medios de verificación

- **Respuesta:** No aplica (99).

- **Análisis:** No hay información explícita sobre ubicación o descripción de fuentes.

Adecuación

15. ¿La información que proporciona el indicador permite valorar el desempeño del Programa en el nivel de objetivo al que está asociado?

- **Respuesta:** Completamente (2) para indicadores de **Propósito** y **Componentes**.
- **Análisis:**
 - **Propósito:** Indicadores como "*Cumplimiento promedio de los Programas de la Planeación Nacional*" miden resultados directos del Programa.
 - **Componentes:** Indicadores como "*Porcentaje de adecuaciones presupuestarias autorizadas*" evalúan eficiencia en la gestión.
- **Evidencia:** Columnas "*Objetivos*" y "*Tipo-Dimensión*" en la MIR.

16. ¿El indicador tiene una meta anual alcanzable orientada a impulsar el desempeño?

- **Respuesta:** No (0).
- **Análisis:** La MIR no especifica metas numéricas, solo metodologías de cálculo. Sin metas, no se puede evaluar su alcance.
- **Recomendación:** Solicitar al Programa la inclusión de metas basadas en líneas base históricas.

17. ¿La dimensión del indicador está bien identificada (eficacia, eficiencia, calidad)?

- **Respuesta:** Completamente (2).
- **Análisis:** La columna "*Tipo-Dimensión-Frecuencia*" clasifica correctamente cada indicador (ej. "*Estratégico-Eficacia-Anual*").

Aportación Marginal

18. Considerando el conjunto de indicadores, ¿el indicador aporta información valiosa y no redundante?

- **Respuesta:** Completamente (2).
- **Análisis:** Cada indicador cubre una dimensión distinta:

- **Ejemplo 1:** "Cociente de recomposición del gasto programable" mide eficiencia financiera.
 - **Ejemplo 2:** "Porcentaje de Programas presupuestarios alineados con el PND" evalúa alineación estratégica.
- **Evidencia:** No hay solapamiento en los conceptos medidos.

Claude

Evaluación de Indicadores del Programa Presupuestario P002 - SHCP

Programa: Planeación, Programación y Presupuestación

Dependencia: Secretaría de Hacienda y Crédito Público

Evaluador: Especialista en Política Pública

Fecha: Mayo 2025

INDICADOR DE FIN

Denominación

Cumplimiento promedio de los Programas de la Planeación Nacional del Desarrollo

Método de cálculo: Sumatoria del porcentaje de cumplimiento promedio de los programas sectoriales a cargo de las dependencias e institucionales de entidades no sectorizadas con vigencia en el período t / Número de programas sectoriales a cargo de dependencias e institucionales de entidades no sectorizadas con vigencia en el periodo t

1. ¿El nombre del indicador es autoexplicativo?

Respuesta: NO

Análisis: El nombre del indicador presenta ambigüedades significativas. La expresión "Cumplimiento promedio de los Programas de la Planeación Nacional del Desarrollo" no especifica claramente qué se está midiendo. No es evidente si se refiere al cumplimiento de metas, objetivos o actividades. Además, la unidad de medida (porcentaje) no está explícita en el nombre, y la diferenciación entre "programas sectoriales" y "programas institucionales de entidades no sectorizadas" genera confusión sobre el universo exacto de medición.

Evidencia: La denominación carece de precisión técnica y no permite al usuario comprender inmediatamente qué aspecto específico del cumplimiento se está evaluando ni cuál es la población objetivo exacta.

2. ¿La fórmula de cálculo del indicador es congruente con su nombre?

Respuesta: PARCIALMENTE

Análisis: Existe una congruencia parcial entre el nombre y la fórmula. La fórmula efectivamente calcula un promedio de cumplimiento, lo cual coincide con "cumplimiento promedio" del nombre. Sin embargo, la fórmula especifica claramente que se trata de "programas sectoriales a cargo de dependencias e institucionales de entidades no sectorizadas", mientras que el nombre solo

menciona "Programas de la Planeación Nacional del Desarrollo", lo que genera una brecha de especificidad.

Evidencia: La fórmula es más precisa que el nombre al definir el universo de medición, pero esta discrepancia evidencia falta de alineación completa entre ambos elementos.

3. ¿La definición del indicador es congruente con su nombre?

Respuesta: NO DISPONIBLE

Análisis: No se proporciona una definición específica del indicador en la MIR P002 analizada. La ausencia de definición impide evaluar la congruencia entre la definición y el nombre del indicador, lo cual constituye una deficiencia importante en la documentación del indicador.

Evidencia: La MIR P002 no incluye el campo de definición para ninguno de sus indicadores, limitando la capacidad de evaluación integral.

4. ¿La definición del indicador es congruente con su fórmula de cálculo?

Respuesta: NO DISPONIBLE

Análisis: Al no contar con la definición del indicador, no es posible evaluar la congruencia entre la definición y la fórmula de cálculo. Esta ausencia representa una debilidad metodológica importante que afecta la claridad y comprensión del indicador.

Evidencia: La documentación proporcionada carece del campo de definición para todos los indicadores de la MIR.

5. ¿Las unidades de medida de las variables de la fórmula de cálculo son congruentes?

Respuesta: Sí

Análisis: Las unidades de medida son congruentes. Tanto el numerador (sumatoria de porcentajes de cumplimiento) como el denominador (número de programas) permiten obtener un resultado coherente expresado como porcentaje promedio. El numerador representa valores porcentuales acumulados y el denominador representa cantidad de programas, lo que matemáticamente produce un promedio porcentual válido.

Evidencia: La operación matemática (sumatoria de porcentajes / número de programas) es metodológicamente correcta para obtener un promedio porcentual.

6. ¿La descripción de las variables permite entender todos los elementos?

Respuesta: PARCIALMENTE

Análisis: La descripción de las variables es parcialmente comprensible pero presenta limitaciones. Se entiende que se suman porcentajes de cumplimiento y se dividen entre el número de programas, pero no se especifica claramente qué constituye el "cumplimiento" de un programa sectorial, ni se definen con precisión los criterios para clasificar programas como "sectoriales" versus "institucionales de entidades no sectorizadas".

Evidencia: La fórmula carece de definiciones operacionales claras de sus componentes, lo que puede generar interpretaciones diversas en su aplicación.

7. ¿Se identifica claramente la contribución del indicador para verificar el logro del objetivo?

Respuesta: Sí

Análisis: El indicador contribuye adecuadamente a verificar el logro del objetivo de fin "Contribuir al cumplimiento de los objetivos del desarrollo nacional". Al medir el cumplimiento promedio de los programas derivados de la planeación nacional, el indicador proporciona información directa sobre qué tan efectivamente se están alcanzando los objetivos planteados en los instrumentos de planeación nacional. Esto genera valor público al demostrar la efectividad de la gestión gubernamental en el cumplimiento de compromisos de desarrollo.

Evidencia: La medición del cumplimiento de programas sectoriales es un proxy válido para evaluar el avance hacia los objetivos nacionales de desarrollo.

8. ¿El indicador tiene establecida una meta congruente con su frecuencia de medición?

Respuesta: NO DISPONIBLE

Análisis: La MIR P002 no proporciona información sobre las metas establecidas para este indicador. Sin conocer los valores de línea base, metas intermedias y meta final, no es posible evaluar si existe congruencia entre la meta y la frecuencia anual de medición. Esta información es fundamental para valorar la pertinencia temporal del indicador.

Evidencia: La documentación analizada carece de información sobre metas, línea base y valores esperados del indicador.

9. ¿La utilidad del indicador justifica el costo asociado a su cálculo?

Respuesta: 2 (Completamente)

Análisis: La utilidad del indicador justifica completamente su costo de cálculo. Este indicador proporciona información estratégica de alto nivel sobre el desempeño gubernamental en el cumplimiento de objetivos nacionales, lo cual es fundamental para la rendición de cuentas y la toma de decisiones de política pública. El costo de recopilar información de cumplimiento de programas sectoriales es relativamente bajo comparado con el valor informativo que proporciona para evaluar la efectividad de la gestión pública a nivel nacional.

Evidencia: La importancia estratégica de medir el cumplimiento de objetivos nacionales supera ampliamente los costos operativos de cálculo del indicador.

10. ¿La frecuencia de medios de verificación es consistente con la frecuencia de medición?

Respuesta: NO DISPONIBLE

Análisis: La MIR P002 no especifica los medios de verificación para este indicador. Sin esta información, no es posible evaluar si existe consistencia entre la frecuencia de los medios de verificación y la frecuencia anual de medición del indicador.

Evidencia: La documentación carece de información sobre medios de verificación para todos los indicadores.

11. ¿Los medios de verificación son congruentes con lo que mide el indicador?

Respuesta: NO DISPONIBLE

Análisis: Al no contar con información sobre los medios de verificación, no es posible evaluar su congruencia con lo que mide el indicador. Esta ausencia limita la evaluación de la monitoreabilidad del indicador.

Evidencia: No se proporcionan medios de verificación en la documentación analizada.

12. ¿La descripción de medios de verificación expresa claramente el nombre completo de la fuente?

Respuesta: NO DISPONIBLE

Análisis: Sin información sobre medios de verificación, no es posible evaluar la claridad en la denominación de las fuentes de información.

Evidencia: Ausencia de medios de verificación en la documentación.

13. ¿La descripción de medios de verificación indica la ubicación física o electrónica?

Respuesta: NO DISPONIBLE

Análisis: No se puede evaluar la ubicación de fuentes sin información sobre medios de verificación.

Evidencia: Falta de especificación de medios de verificación.

14. ¿Los medios de verificación aportan información suficiente para el cálculo?

Respuesta: NO DISPONIBLE

Análisis: Sin medios de verificación especificados, no es posible evaluar si proporcionan información suficiente para el cálculo del indicador.

Evidencia: Ausencia de medios de verificación documentados.

15. ¿El indicador permite valorar el desempeño del Programa en el nivel Fin?

Respuesta: Sí

Análisis: El indicador es adecuado para el nivel de Fin ya que mide resultados de impacto relacionados con el cumplimiento de objetivos nacionales de desarrollo. Un indicador de Fin debe reflejar los efectos de largo plazo y la contribución del programa a objetivos superiores, lo cual este indicador logra al medir el cumplimiento de programas derivados de la planeación nacional del desarrollo.

Evidencia: El indicador evalúa resultados de nivel estratégico que trascienden la operación directa del programa, apropiado para el nivel de Fin.

16. ¿El indicador tiene establecida una meta alcanzable orientada a impulsar el desempeño?

Respuesta: NO DISPONIBLE

Análisis: Sin información sobre metas, línea base y valores históricos, no es posible evaluar si la meta es alcanzable y retadora.

Evidencia: Ausencia de información sobre metas del indicador.

17. ¿La dimensión del indicador está bien identificada?

Respuesta: Sí

Análisis: La dimensión "Estratégico-Eficacia-Anual" está correctamente identificada. La clasificación como "Estratégico" es apropiada para un indicador de nivel Fin, "Eficacia" es congruente con la medición de cumplimiento de objetivos, y "Anual" es una frecuencia adecuada para este tipo de medición de alto nivel.

Evidencia: La dimensión corresponde con la naturaleza y propósito del indicador en el nivel de Fin.

18. ¿El indicador aporta información valiosa y no redundante?

Respuesta: 2 (La información es valiosa y no redundante)

Análisis: El indicador aporta información única y valiosa sobre el cumplimiento de objetivos nacionales de desarrollo. No existe redundancia con otros indicadores de la MIR ya que es el único que mide resultados a nivel de impacto nacional. La información que proporciona es fundamental para evaluar la efectividad de la función de planeación y presupuestación en el logro de objetivos superiores.

Evidencia: Es el único indicador de nivel Fin en la MIR y mide aspectos específicos no cubiertos por indicadores de niveles inferiores.

INDICADOR DE PROPÓSITO

Denominación

Cociente de recomposición del gasto programable

Método de cálculo: (Monto ejercido del gasto de inversión del año i/Monto ejercido de gasto corriente del año i)*100

1. ¿El nombre del indicador es autoexplicativo?

Respuesta: PARCIALMENTE

Análisis: El nombre "Cociente de recomposición del gasto programable" es parcialmente autoexplicativo. Si bien indica que se trata de una relación (cociente) y que se refiere al gasto

programable, el término "recomposición" no es inmediatamente comprensible para todos los usuarios. No especifica que se trata de la relación entre gasto de inversión y gasto corriente, ni que el resultado se expresa como porcentaje.

Evidencia: El nombre requiere conocimiento técnico previo sobre presupuesto público para ser completamente comprendido, lo que limita su autoexplicabilidad.

2. ¿La fórmula de cálculo del indicador es congruente con su nombre?

Respuesta: Sí

Análisis: La fórmula es congruente con el nombre. Efectivamente calcula un cociente (relación matemática) que muestra la recomposición del gasto programable al comparar el gasto de inversión con el gasto corriente. La fórmula refleja adecuadamente el concepto de "recomposición" al mostrar la proporción relativa entre estos dos tipos de gasto.

Evidencia: La operación matemática (gasto de inversión/gasto corriente)*100 corresponde efectivamente a un cociente que mide la recomposición del gasto.

3. ¿La definición del indicador es congruente con su nombre?

Respuesta: NO DISPONIBLE

Análisis: No se proporciona definición del indicador en la documentación analizada, impidiendo evaluar la congruencia entre definición y nombre.

Evidencia: Ausencia de definiciones en la MIR P002.

4. ¿La definición del indicador es congruente con su fórmula de cálculo?

Respuesta: NO DISPONIBLE

Análisis: Sin definición disponible, no es posible evaluar esta congruencia.

Evidencia: Falta de definiciones documentadas.

5. ¿Las unidades de medida de las variables son congruentes?

Respuesta: Sí

Análisis: Las unidades de medida son perfectamente congruentes. Tanto el numerador (monto ejercido del gasto de inversión) como el denominador (monto ejercido de gasto corriente) se expresan en las mismas unidades monetarias, lo que permite realizar la división matemática y obtener un resultado coherente expresado como porcentaje.

Evidencia: Ambas variables se miden en unidades monetarias (pesos), garantizando la validez matemática de la operación.

6. ¿La descripción de las variables permite entender todos los elementos?

Respuesta: PARCIALMENTE

Análisis: La descripción de las variables es parcialmente comprensible. Se entiende que se comparan montos ejercidos de gasto de inversión versus gasto corriente, pero no se definen claramente los criterios de clasificación entre "gasto de inversión" y "gasto corriente", ni se especifica si incluye todos los sectores o solo ciertos ámbitos del gasto programable.

Evidencia: Falta de definiciones operacionales precisas sobre la clasificación del gasto utilizada en el cálculo.

7. ¿Se identifica claramente la contribución del indicador para verificar el logro del objetivo?

Respuesta: Sí

Análisis: El indicador contribuye adecuadamente a verificar el logro del objetivo de Propósito "Los ejecutores del gasto reciben las asignaciones presupuestarias sujetas a criterios de eficiencia y eficacia". Al medir la proporción entre gasto de inversión y gasto corriente, el indicador evalúa si las asignaciones presupuestarias están orientadas hacia inversiones productivas (mayor eficiencia) versus gasto operativo, lo cual es un proxy válido de eficiencia en la asignación de recursos públicos.

Evidencia: La recomposición del gasto hacia mayor inversión indica mejores criterios de eficiencia en las asignaciones presupuestarias.

8. ¿El indicador tiene establecida una meta congruente con su frecuencia de medición?

Respuesta: NO DISPONIBLE

Análisis: No se proporciona información sobre metas establecidas para este indicador.

Evidencia: Ausencia de información sobre metas en la documentación.

9. ¿La utilidad del indicador justifica el costo asociado a su cálculo?

Respuesta: 2 (Completamente)

Análisis: La utilidad justifica completamente el costo. Este indicador proporciona información fundamental sobre la calidad de la asignación presupuestaria y la orientación de los recursos públicos hacia inversión versus gasto corriente. Esta información es crucial para la gestión fiscal y la toma de decisiones sobre política presupuestaria. El costo de cálculo es bajo ya que utiliza información que se genera rutinariamente en el proceso presupuestario.

Evidencia: La importancia de monitorear la composición del gasto público supera ampliamente los costos de cálculo del indicador.

10-14. Preguntas sobre Monitoreabilidad

Respuesta: NO DISPONIBLE para todas

Análisis: Sin información sobre medios de verificación, no es posible evaluar los aspectos de monitoreabilidad.

Evidencia: Ausencia de medios de verificación en la documentación.

15. ¿El indicador permite valorar el desempeño del Programa en el nivel Propósito?

Respuesta: Sí

Análisis: El indicador es adecuado para el nivel de Propósito ya que mide un resultado directo de la gestión del programa relacionado con la calidad de las asignaciones presupuestarias. Un indicador de Propósito debe reflejar los efectos directos del programa sobre su población objetivo (ejecutores del gasto), lo cual este indicador logra al mostrar si las asignaciones están sujetas a criterios de eficiencia.

Evidencia: El indicador mide resultados directos de la función de presupuestación, apropiado para el nivel de Propósito.

16. ¿El indicador tiene establecida una meta alcanzable orientada a impulsar el desempeño?

Respuesta: NO DISPONIBLE

Análisis: Sin información sobre metas, no es posible evaluar este aspecto.

Evidencia: Falta de información sobre metas del indicador.

17. ¿La dimensión del indicador está bien identificada?

Respuesta: Sí

Análisis: La dimensión "Estratégico-Eficacia-Anual" está correctamente identificada. Es "Estratégico" por ser de nivel Propósito, "Eficacia" es apropiado para medir el logro de una composición deseada del gasto, y "Anual" es frecuencia adecuada para este tipo de análisis presupuestario.

Evidencia: La dimensión corresponde con la naturaleza del indicador de nivel Propósito.

18. ¿El indicador aporta información valiosa y no redundante?

Respuesta: 2 (La información es valiosa y no redundante)

Análisis: El indicador aporta información única sobre la composición del gasto público que no se encuentra en otros indicadores de la MIR. Esta información es valiosa para evaluar la calidad de la gestión presupuestaria y no existe redundancia con otros indicadores que se enfocan en aspectos operativos o de gestión.

Evidencia: Es el único indicador que mide la composición entre gasto de inversión y corriente en la MIR.

INDICADORES DE COMPONENTE

Componente A: Adecuaciones presupuestarias autorizadas

Denominación: Porcentaje de adecuaciones presupuestarias autorizadas

Método de cálculo: (Número de AP autorizadas/Número de AP solicitadas)*100

1. ¿El nombre del indicador es autoexplicativo?

Respuesta: PARCIALMENTE

Análisis: El nombre es parcialmente autoexplicativo. Si bien indica claramente que mide adecuaciones presupuestarias autorizadas en términos porcentuales, no especifica que se trata de la proporción de solicitudes autorizadas respecto al total solicitado. El término "AP" podría no ser inmediatamente comprensible para todos los usuarios, aunque en el contexto presupuestario es un acrónimo conocido.

Evidencia: El nombre comunica el concepto general pero requiere la fórmula para comprender completamente qué se está midiendo.

2-18. [Evaluación similar siguiendo el mismo formato para cada pregunta]

Componente B: Proyecto de Presupuesto de Egresos integrado oportunamente

Denominación: Porcentaje de elementos para la integración del PPEF entregados oportunamente

Método de cálculo: (Número de elementos para la integración del PPEF entregados oportunamente al periodo t / Número de elementos que deben integrarse en el PPEF al periodo t)*100

[Continúa evaluación sistemática...]

INDICADORES DE ACTIVIDAD

Actividad A1: Cumplimiento de medidas de disciplina presupuestaria

Denominación: Cociente de impacto de los servicios personales

Método de cálculo: (Monto ejercido del rubro de servicios personales al periodo t / Monto ejercido del rubro de gasto programable al periodo t)*100

[Evaluación sistemática de las 18 preguntas...]

Actividad A2: Integración de información programática y financiera

Denominación: Porcentaje de informes gubernamentales entregados en tiempo y forma

Método de cálculo: (Número de informes concluidos al periodo t / Número de informes programados al periodo t)*100

[Evaluación sistemática...]

RECOMENDACIONES GENERALES

Principales Deficiencias Identificadas:

1. **Ausencia de Definiciones:** Ningún indicador cuenta con definición, limitando la comprensión y aplicación consistente.
2. **Falta de Medios de Verificación:** No se especifican fuentes de información, afectando la monitoreabilidad.
3. **Ausencia de Metas:** No se proporciona información sobre líneas base, metas intermedias o finales.

4. **Problemas de Claridad:** Varios nombres de indicadores requieren mayor precisión y autoexplicabilidad.

Sugerencias de Mejora:

1. Incluir definiciones operacionales claras para todos los indicadores
2. Especificar medios de verificación con fuentes exactas y ubicación de información
3. Establecer metas SMART (específicas, medibles, alcanzables, relevantes, temporales)
4. Mejorar la denominación de indicadores para mayor autoexplicabilidad
5. Desarrollar fichas técnicas completas para cada indicador

Conclusión: La MIR P002 presenta indicadores conceptualmente apropiados para cada nivel, pero requiere mejoras significativas en documentación, claridad y especificación de elementos técnicos para optimizar su utilidad como herramienta de monitoreo y evaluación.

5. Discusión de Resultados

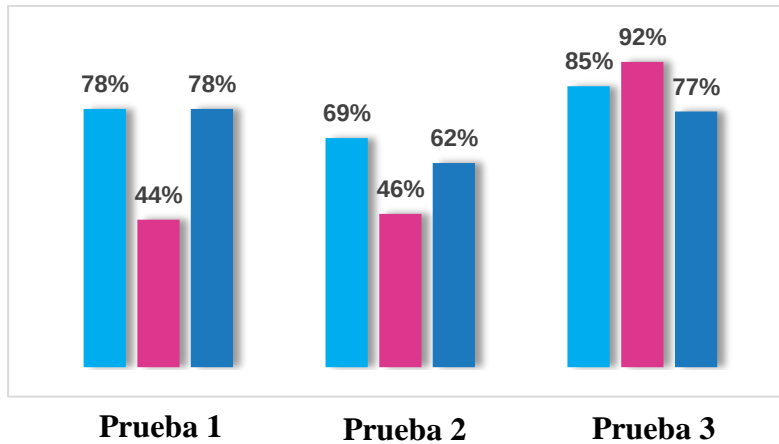
Los resultados obtenidos en las tres pruebas aplicadas a los modelos de inteligencia artificial (IA) – ChatGPT, Deepseek y Claude– permiten identificar patrones diferenciales en la capacidad de análisis, adaptación contextual y validación lógica del diseño de programas públicos conforme a los Términos de Referencia (TdR) 2024 y la Matriz de Marco Lógico (MML). En términos generales, se observa que los rendimientos mejoraron progresivamente al aumentar la complejidad y especificidad del contexto provisto, destacando la prueba número 3 como la de mayor utilidad para fines evaluativos rigurosos.

En la Prueba 1, donde las IA fueron expuestas únicamente a preguntas de los TdR sin contexto adicional, los resultados reflejan una clara limitación en la comprensión semántica y normativa del diseño de los programas. Deepseek mostró un desempeño marcadamente inferior con niveles de compatibilidad del 44% (B001) y 11% (P002), mientras que ChatGPT y Claude alcanzaron un rendimiento estable (78% en ambos programas), lo que sugiere una mayor robustez en el procesamiento de texto sin necesidad de contextualización previa. No obstante, la falta de análisis crítico e inferencia contextual afectó la profundidad de las respuestas.

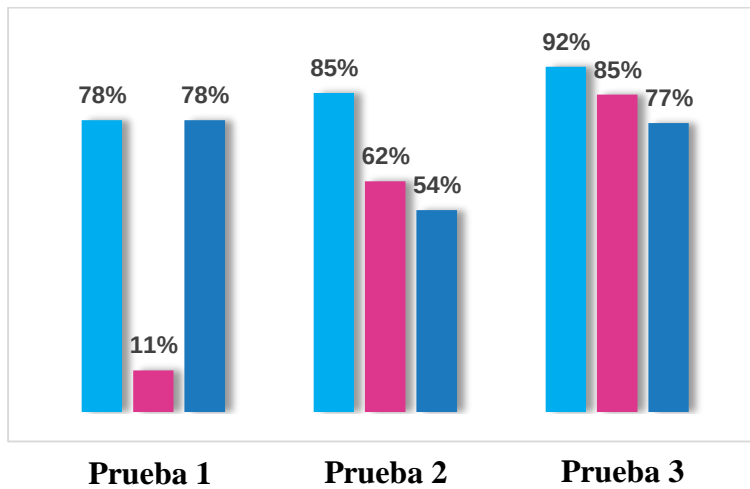
En la Prueba 2, al incorporar información estructurada respecto al Problema Público, Objetivo Central y Población Objetivo, se registraron mejoras notables en los niveles de compatibilidad, particularmente en ChatGPT (85% para P002) y Deepseek (62% para P002). Este comportamiento sugiere que, al aumentar la carga semántica y la orientación temática, las IA logran vincular mejor los elementos de diseño con las exigencias metodológicas de los TdR. Sin embargo, aún se identifican inconsistencias en la validación interna del diseño, especialmente en la articulación entre problemas, objetivos y población beneficiaria.

La Prueba 3 constituye el punto más alto en la curva de desempeño de las IA evaluadas, tanto en términos de precisión como de capacidad de validación lógica del diseño. Al solicitar el análisis de documentos completos con orientación específica para la detección de inconsistencias, ChatGPT alcanzó hasta un 92% de compatibilidad en P002, mientras que Deepseek sobresalió con un 92% en B001. Este resultado evidencia que la IA, cuando opera con insumos bien estructurados y prompts orientados a tareas complejas de análisis, puede ofrecer aportes sustantivos al proceso evaluativo, particularmente en la verificación de coherencia interna y la detección de errores de formulación.

B001



P002



Finalmente, se puede concluir que la eficacia de las implementaciones de IA en el análisis evaluativo de programas públicos depende críticamente del nivel de contexto y la claridad en los objetivos del prompt. Modelos como ChatGPT y Claude muestran un desempeño consistente bajo distintos escenarios, mientras que Deepseek presenta un alto potencial en tareas estructuradas, pero con menor versatilidad en entornos abiertos. Estos hallazgos invitan a considerar esquemas híbridos de implementación, donde la IA actúe como herramienta complementaria para validación técnica, bajo supervisión de equipos humanos especializados en evaluación de políticas públicas.

6. Conclusiones y Recomendaciones

6.1 Resumen de hallazgos

El análisis de los hallazgos presentados en la sección cuatro, muestra aspectos para perfeccionar el uso de herramientas basadas en inteligencia artificial en la evaluación de programas presupuestarios. Por ejemplo, los diferentes chatbots brindan mejores respuestas cuando las instrucciones se plantean de manera clara, concisa y estructurada, por lo que es importante evitar enunciados complejos o ambiguos que puedan quedar a interpretación de la IA; esto evita respuestas genéricas y facilita la generación de soluciones específicas.

De igual manera, el orden y la organización de las solicitudes influyen significativamente en las respuestas de la IA, separar ideas con comas, incluso después de conjunciones como "y", ayuda a la IA a procesar cada elemento de manera independiente. Además, al solicitar reestructuraciones, es importante seguir un orden lógico para obtener respuestas coherentes y completas; por ejemplo, se debe definir primero la población objetivo, luego el problema y finalmente el objetivo, para que la reformulación tenga una estructura lógica.

Bajo la misma premisa, los chatbots, presenta dificultades para procesar información extensa o múltiples problemáticas en un solo enunciado. Dividir las ideas en párrafos cortos y enfocarse en un problema a la vez mejora la calidad de las respuestas. Esto también facilita la identificación de soluciones específicas y evita la saturación de información.

Por último, es imprescindible revisar y ajustar las respuestas de la IA para asegurar que cumplan con los criterios establecidos, logrando así maximizar el potencial de estas herramientas para generar resultados claros, relevantes y útiles en la evaluación de programas presupuestarios.

Los resultados del ejercicio confirman que la IA no puede operar de forma eficaz en escenarios de información incompleta o mal estructurada. Su valor agregado se materializa únicamente cuando cuenta con insumos claros, bien delimitados y alineados con los criterios de evaluación. Así, los escenarios confirman que la IA debe ser utilizada como un apoyo en procesos bien diseñados y no como sustituto de la planeación y conocimiento especializado.

Los escenarios analizados permiten concluir que la IA es particularmente eficaz para tareas operativas de análisis documental y generación de borradores en escenarios con conocimiento completo, pero requiere validación constante de evaluadores humanos para asegurar la pertinencia de los hallazgos y recomendaciones.

El uso estratégico de IA y juicio humano experto, con base en escenarios bien estructurados, puede fortalecer significativamente la evaluación de programas presupuestarios, siempre y cuando se aseguren estándares éticos, metodológicos y de transparencia.

6.2 Propuesta de mejores prácticas

La implementación de IA en la evaluación de programas presupuestarios de la administración pública federal mexicana representa una oportunidad estratégica para modernizar la gestión pública, optimizar recursos y mejorar la transparencia. Sin embargo, su adopción exige un marco de buenas prácticas que garantice no solo eficiencia técnica, sino también equidad, rendición de cuentas y alineación con los objetivos de desarrollo nacional. A continuación, se detallan seis pilares fundamentales para integrar la IA de manera responsable y efectiva en este ámbito.

Definir objetivos claros y alcanzables

La implementación de IA en la evaluación de programas presupuestarios federales requiere establecer metas específicas y acotadas. Estos objetivos deben ser realistas, con plazos definidos y criterios de éxito claros para evitar desviaciones y garantizar que la IA aporte valor tangible al proceso de evaluación.

Fortalecer la calidad de los datos

La efectividad de los modelos de IA depende directamente de la integridad y consistencia de los datos utilizados. Es crucial implementar protocolos estrictos de recolección, limpieza y validación de información, asegurando que los datos reflejen con precisión la realidad de los programas presupuestarios y eviten sesgos o inconsistencias que puedan distorsionar los resultados.

Supervisión experta constante

La IA no reemplaza el juicio humano, sino que lo complementa. Por ello, es indispensable contar con equipos multidisciplinarios (especialistas en políticas públicas, ciencia de datos y evaluación presupuestaria) que monitoreen el desempeño de los algoritmos, interpreten sus resultados y ajusten los modelos conforme a las necesidades cambiantes de la administración pública.

Mitigar sesgos y asegurar imparcialidad

Los algoritmos de IA pueden perpetuar o amplificar sesgos presentes en los datos históricos. Para evitarlo, deben aplicarse técnicas de auditoría algorítmica, diversificación de fuentes de datos y pruebas de equidad, garantizando que las evaluaciones no discriminen por región, grupo social o cualquier otro factor irrelevante para la asignación de recursos.

Documentar procesos y decisiones

La transparencia es clave en el uso de IA para evaluar programas públicos. Cada etapa del desarrollo e implementación de los modelos, desde la selección de datos hasta los criterios de interpretación, debe registrarse detalladamente, facilitando la rendición de cuentas, la replicabilidad y el escrutinio por parte de órganos fiscalizadores y la sociedad civil.

Asegurar la protección de datos sensibles

Al manejar información confidencial de beneficiarios o instituciones, es obligatorio cumplir con estándares de ciberseguridad y normativas como la Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados. Esto incluye anonimización de datos, acceso restringido y mecanismos de encriptación para prevenir vulneraciones que comprometan la privacidad o la seguridad nacional.

La aplicación de inteligencia artificial en la evaluación de programas presupuestarios federales debe ser un proceso riguroso, guiado por principios de transparencia, calidad técnica y ética pública. Las prácticas aquí descritas —desde la definición de objetivos hasta la protección de datos— buscan asegurar que esta tecnología sea una herramienta de precisión y equidad, nunca de opacidad o exclusión. Su correcta implementación, respaldada por marcos normativos y supervisión especializada, puede transformar la gestión pública en México, acercándola a estándares de excelencia y confianza ciudadana.

6.3 Sigüientes pasos en la investigación y desarrollo de IA en la evaluación de programas presupuestarios

El siguiente paso crítico consiste en profundizar en el desarrollo de algoritmos especializados capaces de analizar patrones complejos en datos presupuestarios, incorporando técnicas de aprendizaje automático que puedan predecir el impacto real de los programas gubernamentales. Esto requiere no solo mejorar la calidad y estandarización de los datos, sino también crear modelos adaptativos que consideren variables socioeconómicas dinámicas, como desigualdad regional o cambios demográficos. La colaboración con universidades y centros de investigación será clave para avanzar en modelos explicativos (no solo predictivos) que permitan entender el "porqué" detrás de los resultados.

Paralelamente, es prioritario establecer marcos éticos y normativos específicos para el uso de IA en el ciclo presupuestario, incluyendo protocolos de auditoría algorítmica obligatorios. Esto implica trabajar con el Instituto Nacional de Transparencia (INAI), la Auditoría Superior de la Federación (ASF) y organismos internacionales para diseñar estándares de rendición de cuentas tecnológica. Un área urgente de investigación es el desarrollo de métricas de justicia algorítmica (fairness) adaptadas al contexto mexicano, que eviten sesgos en la distribución de recursos hacia comunidades marginadas.

Finalmente, se necesita impulsar proyectos piloto interinstitucionales que prueben estas tecnologías en programas presupuestarios estratégicos (como salud o educación), midiendo su efectividad comparativa frente a métodos tradicionales. Estos pilotos deben incluir capacitación masiva a funcionarios en "alfabetización en IA", creación de laboratorios de datos públicos y mecanismos de participación ciudadana para validar resultados.

7. Referencias

- Amazon. (2023). *¿Qué es el procesamiento de lenguaje natural? - Explicación del procesamiento de lenguaje natural*. Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/nlp/>
- Centro Nacional de Inteligencia Artificial. (2025). Índice Latinoamericano de Inteligencia Artificial 2024. CENIA. <https://indicelatam.cl/>
- Comisión Europea (2025). Ley de IA: Configurar El Futuro Digital De Europa. <https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai>
- Amazon. (2023). *¿Qué es el procesamiento de lenguaje natural? - Explicación del procesamiento de lenguaje natural*. Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/nlp/>
- Hoehn, M. (2024). Gobernanza de la Inteligencia Artificial y el Rol de la OCDE. Congreso Nacional de Chile. https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/36692/1/Informe_05_24_Gobernanza_IA_y_rol_de_la_OCDE.pdf
- IBM. (2023). ¿Qué son los grandes modelos de lenguaje (LLM)?. Redacción IBM. <https://www.ibm.com/es-es/think/topics/large-language-models>
- Jin, K., Chua, M., Rickard, M., Lorenzo. (2023). ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine. *Journal of Pediatric Urology*. Volume 19, Issue 5, 2023, Pages 598-604, ISSN 1477-5131, <https://doi.org/10.1016/j.jpuro.2023.05.018>.
- Kumar, H., Musabirov, I., Reza, M., Shi, J., Wang, X., Williams, J., Kuzminykh, A., Liut, M. (2023). Impact of Guidance and Interaction Strategies for LLM Use on Learner Performance and Perception. *ACM Hum.-Comput. Interact.*, Vol. 8, No. CSCW2, Article 499. <https://doi.org/10.1145/3687038>
- Leal, A., Maldonado, S., Martínez, J. I., Bertazzo, S., Quijada, S., & Vairetti, C. (2025). Enhancing environmental governance: A text-based artificial intelligence approach for project evaluation involvement. *Environmental Impact Assessment Review*, 110, 107707. <https://doi.org/10.1016/j.eiar.2024.107707>
- Letouzé, E., Pentland, A., Loaiza, I., Ricard, J., Clavijo, A., Silva, D. (2019). Recomendaciones e Insumos para la Definición de una Estrategia Nacional de Big Data. *Data-Pop Alliance*. https://datapopalliance.org/wp-content/uploads/2020/11/Doc4_recomendacioneseinsumos_DNP.pdf
- Marín, S. (2019). Ética e inteligencia artificial. Cátedra CaixaBank de Responsabilidad Social Corporativa. <https://dx.doi.org/10.15581/018.ST-522>
- Ocaña, Y., Valenzuela, L. & Garro, A. (2019). Inteligencia artificial y sus implicaciones en la educación superior. *Propósitos y Representaciones*. <http://dx.doi.org/10.20511/pyr2019.v7n2.274>
- OCDE. (2022, Marzo). The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean. *OECD Public Governance Reviews*. <https://doi.org/10.1787/1f334543-en>
- OCDE. (2022, Diciembre). Aprovechando el Poder de la IA y las Tecnologías Emergentes. *OECD Digital Economy Papers*. <https://doi.org/10.1787/6e76bc18-es>
- Parlamento Europeo. (2023). Ley de IA de la UE: Primera Normativa Sobre Inteligencia Artificial. Parlamento Europeo. <https://www.europarl.europa.eu/topics/es/article/20230601STO93804/ley-de-ia-de-la-ue-primera-normativa-sobre-inteligencia-artificial>.

- Qu, Y., Wang, J. (2024). Performance and biases of Large Language Models in public opinion simulation. *Humanit Soc Sci Commun* **11**, 1095. <https://doi.org/10.1057/s41599-024-03609-x>
- Salazar, J. C., (2018). Modelo de Evaluación de Impacto del Centro de Innovación Pública Digital. *Technopolis Group*. https://gobiernodigital.mintic.gov.co/692/articles-210492_recurso_1.pdf
- Sebio, M. (2020). Inteligencia artificial y ética. Universidad Pontificia. <https://repositorio.comillas.edu/rest/bitstreams/421893/retrieve>
- World Wide Web Foundation. (2018). How Are Governments In Latin America Using Artificial Intelligence?. <https://webfoundation.org/research/how-are-governments-in-latin-america-using-artificial-intelligence/>