

## ANEXO II. Criterios de medición para cada uno de los Niveles de Madurez de la Gestión de Datos por Componentes

Establecer un marco para la gestión de datos no es suficiente si no se mide el avance en la consecución de sus objetivos. Por esta razón el MGD-CDMX se acompaña de un Modelo de Madurez para poder monitorear el avance en la implementación de la gestión de datos y así diseñar estrategias de seguimiento y mejora.

El modelo de madurez del MGD-CDMX es una adaptación con cuatro niveles de madurez (1. Inicial, 2. Repetible, 3. Definido y, 4. Gestionado), de la propuesta de Integración del Modelo de Madurez de Capacidades (CMMI, por sus siglas en inglés<sup>1</sup>) desarrollado por la Universidad de Carnegie Mellon y administrado por el Instituto CMMI.

Estos niveles de madurez son medidos para cada uno de los Componentes de la gestión de datos; y el desempeño conjunto de todas ellas determina el nivel general de madurez en la gestión de datos de cada uno de los Entes. Mientras que el promedio de la madurez de todos los Entes indicará el nivel de madurez general para toda la Administración Pública de la Ciudad de México.

Los elementos y criterios de medición para cada uno de los Niveles de Madurez de la Gestión de datos, por Componente, son los siguientes:

### 1. Gobernanza y Cultura de Datos: niveles de madurez

Dado su carácter transversal, y su objetivo de incidir en la cultura y cambio organizacionales, la medición de los niveles de madurez para la Gobernanza y Cultura de Datos, se desglosa en cinco elementos distintos, a saber: **i)** cultura de datos; **ii)** estándares, normativa y procesos; **iii)** respuesta a problemáticas de datos; **iv)** definición de roles para la gestión de datos; y, **v)** documentación de datos. Cada uno de estos elementos tiene sus características particulares según el nivel de madurez en que se encuentra, como se puede ver a continuación:

Tabla 1. Niveles de Madurez para la Gobernanza y Cultura de Datos		
Elementos	Nivel	Descripción
<b>Cultura de datos:</b> Se sabe de la importancia de los datos para aumentar valor organizacional y público. Se impulsan acciones afirmativas para su promoción	<b>Inicial</b>	Los asuntos sobre datos son vistos como asuntos exclusivos del área de Sistemas y Tecnologías de la Información. No existe un impulso a la cultura de datos.
	<b>Repetible</b>	Se comienzan a impulsar proyectos con el uso de datos para mostrar su valor.
	<b>Definido</b>	Se impulsa y promueve el uso de datos para distintos proyectos con la finalidad de impactar en una mejor toma de decisiones.
	<b>Gestionado</b>	Los datos se reconocen como un activo clave que debe ser gestionado para cada una de las áreas. Se impulsan capacitaciones constantes en cultura de datos.
<b>Estándares, normativa y procesos:</b> Existencia y robustez de estándares, normativa y procesos sobre datos.	<b>Inicial</b>	No existen estándares, normativa o procesos definidos sobre cómo gestionar los datos.
	<b>Repetible</b>	Prácticas locales exitosas son compartidas con otras áreas y algunas buenas prácticas comienzan a emerger.
	<b>Definido</b>	Algunas mejores prácticas son elevadas a estándares de facto y hay una adopción de éstas en áreas específicas.
	<b>Gestionado</b>	Es definido un programa de normativas y estándares de gestión de datos a nivel de toda la organización.

<sup>1</sup> Carnegie Mellon Univ. Software Eng. Inst., Capability Maturity Model: Guidelines for Improving the Software Process, 1st ed., Addison-Wesley Professional, 1995.

<b>Respuesta a problemáticas de datos:</b> Gestión o respuesta a problemáticas de datos (data issues).	<b>Inicial</b>	Se responde ante problemáticas de datos cuando éstas aparecen (es decir, de manera reactiva) y no necesariamente se busca prevenir que vuelvan a suceder.
	<b>Repetible</b>	Las problemáticas de datos comienzan a ser respondidas usando procesos repetibles que comienzan a ser formalizados.
	<b>Definido</b>	Las problemáticas sobre datos comienzan a ser rastreadas de manera rigurosa. La organización incluye evaluaciones de riesgo para la calidad de datos en proyectos en proceso.
	<b>Gestionado</b>	Los problemas y las soluciones a la calidad de los datos son medidos, monitoreados y comunicados. Las evaluaciones de riesgos de datos en torno a proyectos son realizadas de manera oportuna.
<b>Definición de roles para la gestión de datos:</b> La gestión depende de la definición concreta de roles para la gestión de datos.	<b>Inicial</b>	Cada área general, y por iniciativas aisladas de algunos individuos, define sus propios roles de datos basados en necesidades específicas.
	<b>Repetible</b>	En algunas áreas generales, la responsabilidad sobre la gestión de los datos está más formalizada y descrita. Comienza una delineación de los roles generales y de Tecnologías de la Información sobre los datos.
	<b>Definido</b>	Un programa de Administración de Datos a nivel organización es definido y comunicado, con roles de responsabilidad explícitos.
	<b>Gestionado</b>	La administración de datos se considera una responsabilidad organizacional. Los responsables de datos llevan a cabo sus funciones respondiendo a estas responsabilidades previamente definidas, institucionalizadas y estandarizadas.
<b>Documentación de datos:</b> Existencia de documentación (diccionario de datos, síntesis metodológicas, etc.) para los conjuntos de datos.	<b>Inicial</b>	No existe documentación para los datos.
	<b>Repetible</b>	Existe de forma aislada solo para algunos casos; y ésta se implementó a iniciativa de algún/algunos individuos de la organización.
	<b>Definido</b>	Existe un plan de documentación, aunque no se cumple para todos los casos.
	<b>Gestionado</b>	Existe documentación para todos los datos y ésta responde a un plan de documentación estandarizado e institucionalizado.

## 2. Calidad de Datos: niveles de madurez

Los niveles de madurez de este Componente se miden a partir de la evaluación de la existencia o no, de procedimientos, estándares y/o métricas para garantizar un nivel mínimo de calidad de los datos. Así como si estos datos cuentan con algunos de los criterios relacionados con las dimensiones de accesibilidad, certeza, adecuados, oportunidad e integridad, como las dimensiones mínimas necesarias:

Tabla 2. Niveles de madurez para Calidad de Datos

Nivel de madurez	Descripción
Inicial	<p>No existen procedimientos para garantizar y evaluar la calidad de los datos. La mayoría de los datos no cuentan con calidad, lo que implica que:</p> <ul style="list-style-type: none"> <li>No son accesibles por presentarse en formatos y sistemas no adecuados.</li> <li>Tienen que pasar por varios y complicados procesos de limpieza para poder ser utilizados, y, además dichos procesos no son documentados para que otros usuarios los consulten.</li> <li>No se puede garantizar la integridad de los datos, al no existir procesos para asegurar que éstos no fueron modificados o si están desactualizados.</li> <li>Se presentan en formatos no legibles por máquinas.</li> <li>En caso de publicarse, estos siempre se presentan después de las fechas previamente acordadas.</li> </ul>
Repetible	<p>Solo para algunas áreas y de manera informal y esporádica se implementan procedimientos para la calidad de los datos. Solo algunos datos cuentan con calidad, lo que implica que:</p> <ul style="list-style-type: none"> <li>Los formatos y sistemas no son adecuados.</li> </ul>

	<ul style="list-style-type: none"> <li>• La mayoría tienen que pasar por varios procesos de limpieza para poder ser utilizados, otros no, y, además dichos procesos no son documentados para que otros usuarios los consulten.</li> <li>• Solo en algunas ocasiones se puede garantizar la integridad de los datos.</li> <li>• La mayoría se presentan en formatos no legibles por máquinas.</li> <li>• Solo algunos conjuntos de datos se publican en las fechas previamente acordadas.</li> </ul>
Definido	<p>Se han definido de manera formal procedimientos para la calidad de los datos de manera formal y para toda la organización.</p> <p>La mayoría de los datos son de calidad, lo que implica que:</p> <ul style="list-style-type: none"> <li>• Son accesibles a la mayoría de usuarios por presentarse en formatos y sistemas adecuados.</li> <li>• Los procesos de limpieza son menores y, son documentados para la consulta de otros usuarios.</li> <li>• Se puede garantizar la integridad de la mayoría de los datos, al existir procesos para asegurar que éstos no fueron modificados y que están actualizados.</li> <li>• Se presentan en formatos legibles por máquinas.</li> <li>• La mayoría de los datos se publican en las fechas previamente acordadas.</li> </ul>
Gestionado	<p>Los procedimientos para la calidad de los datos son aplicados de manera rutinaria en toda la organización y para todos los datos.</p> <p>Todos los datos son de calidad, lo que implica que:</p> <ul style="list-style-type: none"> <li>• Son accesibles a la mayoría de usuarios.</li> <li>• Ningún conjunto de datos necesita pasar por procesos de limpieza previo a su utilización.</li> <li>• En todas las ocasiones se puede garantizar la integridad de los datos, pues se sabe a ciencia cierta que éstos no fueron alterados y están actualizados.</li> <li>• Se presentan en formatos legibles por máquinas.</li> <li>• Se publican de manera oportuna en las fechas previamente acordadas.</li> </ul>

### 3. Seguridad de Datos: niveles de madurez

Los niveles de madurez de la seguridad de los datos se evalúan según la existencia o no de estándares y procedimientos para garantizar la seguridad de los datos, tanto física como de software; así como si se cuentan con controles de acceso, de autorización y/o de autenticación:

Tabla 3. Niveles de madurez para Seguridad de Datos	
Nivel de madurez	Descripción
Inicial	No existen estándares ni procedimientos en materia de seguridad física, ni de software. No hay controles de acceso, ni autorización, ni se aseguran de la autenticación. No se realizan auditorías de seguridad.
Repetible	Ciertas prácticas de seguridad, tanto físicas como de software, se comienzan a generalizar y replicar por más miembros de la organización. Hay algunos controles de acceso, autorización y autenticación, pero no de forma integral. Se realizan auditorías del software de forma automática, sin seguimiento ni monitoreo.
Definido	Se siguen estándares y procedimientos homologados en materia de seguridad física y de software, asegurando la protección de datos contra accesos no autorizados evitando modificaciones y/o pérdidas de información parcial o totales. Controles de acceso, autorización y autenticación establecidos y seguidos. Las auditorías son monitoreadas con regularidad.
Gestionado	Los estándares y procesos para seguridad física, y de software no sólo se cumplen sino que están en constante revisión para innovar según las mejores prácticas internacionales. Los controles de acceso, autorización, autenticación y auditoría garantizan la seguridad de los datos de la organización.

#### 4. Gestión de Metadatos: niveles de madurez

La madurez de este componente se evalúa según si existen o no metadatos para todos los conjuntos de datos, para algunos de ellos o para ninguno; y, además si se implementa un catálogo interoperable de búsqueda y consulta de dichos metadatos:

Tabla 4. Niveles de madurez para Gestión de Metadatos	
Nivel de madurez	Descripción
Inicial	No se usan metadatos y no se cuenta con un estándar. No existe un catálogo de metadatos.
Repetible	Algunos conjuntos de datos tienen metadatos, pero no hay un estándar. No hay un catálogo de metadatos interoperable.
Definido	La mayoría de los conjuntos de datos cuentan con metadatos establecidos mediante un estándar. Se busca hacer el catálogo interoperable.
Gestionado	Todos los conjuntos de datos tienen sus respectivos metadatos. Existe un estándar y un catálogo interoperable.

#### 5. Datos maestros y de referencia: niveles de madurez

Los niveles de madurez de este componente se evalúan según la existencia o no de datos maestros y de referencia comunes, integrados e interoperables para toda la organización:

Tabla 5. Niveles de madurez para Datos maestros y de Referencia	
Nivel de madurez	Descripción
Inicial	No existen datos de referencia que faciliten la integración de los datos de la organización.
Repetible	Existen datos de referencia de uso compartido para algunos conjuntos de datos de la organización.
Definido	Existen datos de referencia para la mayoría de los conjuntos de datos de la organización. Se plantean los primeros datos maestros comunes.
Gestionado	Existen datos maestros y de referencia claros y consensuados para todos los conjuntos de datos a lo largo de la organización.

#### 6. Arquitectura de datos: niveles de madurez

La madurez de arquitectura de datos es medida en virtud de si los datos de la organización están dotados de una lógica y estructura que permita integrarlos fácilmente, mediante la planeación, el modelo y diseño adecuadas que cumplan los requisitos organizacionales y según los propósitos de dichos datos:

Tabla 6. Niveles de madurez Arquitectura de Datos	
Nivel de madurez	Descripción
Inicial	Cada conjunto de datos tiene su propia lógica y estructura, su arquitectura responden a necesidades específicas, por falta de modelación y diseño previo que no garantizan integridad ni accesibilidad. Abundan silos de información.
Repetible	Existen elementos en común entre las bases de datos, pero no existe una arquitectura definida. Fragmentación parcial y alta redundancia de datos. Algunos datos de la organización están integrados, pero siguen existiendo silos.

Definido	Se siguen procesos definidos para tener arquitecturas estandarizadas en la mayoría de las organizaciones. Se definen y siguen procedimientos de diseño y modelado de bases de datos, lo que permite que se integren datos bajo procesos definidos en tiempos dentro del margen. Pocos silos de información.
Gestionado	Existen arquitecturas estandarizadas, diseñadas para explotar los datos al máximo. Se evita la fragmentación y la redundancia de los datos, por diseñar y modelar los requerimientos de las bases de datos con anticipación. No existen silos de información. Todos los datos están integrados.

### 7. Gestión de Documentos y contenido: niveles de madurez

El nivel de madurez de documentos y contenido se mide según la existencia o no de lineamientos para su gestión; y, además, de si éstos pueden encontrarse fácilmente y con cierta calidad:

Tabla 7. Niveles de madurez para Gestión de Documentos y Contenido	
Nivel de madurez	Descripción
Inicial	No existen lineamientos generales para la gestión de datos dentro de documentos y contenido, por ello dichos datos no están organizados ni catalogados. No se puede hacer uso de ellos a tiempo y de encontrarlos no sería con buena calidad.
Repetible	Existen algunos lineamientos y estándares para la gestión de datos dentro de documentos y contenidos, pero éstos no se aplican de manera generalizada. Algunos documentos se pueden usar, pero depende de la habilidad y cuidado de las personas. La calidad de los documentos no está asegurada.
Definido	Existen lineamientos y procesos estandarizados para todos los datos que se encuentran en documentos y contenido que ayudan a que dichos datos estén almacenados de forma ordenada y puedan ser accesados con facilidad. Los documentos y contenido se hallan fácilmente y se encuentran con buena calidad.
Gestionado	Se aplican e innovan regularmente los lineamientos y estándares para datos dentro de documentos y contenido, por lo que éstos están almacenados de forma ordenada y pueden ser accesados con calidad y facilidad.

### 8. Almacén de datos e inteligencia gubernamental: niveles de madurez

Los niveles de madurez de este componente se evalúan según la existencia o no de este repositorio central; así como la existencia de procesos de extracción, limpieza, transformación y carga de datos indispensables para alimentar y mantener dicho repositorio:

Tabla 8. Niveles de madurez para Almacén de datos e inteligencia gubernamental	
Nivel de madurez	Descripción
Inicial	No existe un repositorio central donde los datos estén albergados de forma ordenada y sistemática; por lo tanto, es difícil utilizar los datos para la toma inteligente de decisiones. No existen procesos de extracción, limpieza, transformación y carga de datos, desde una variedad de fuentes externas.
Repetible	Solo algunos datos están albergados de forma ordenada en un repositorio central. La toma de decisiones basadas en datos sigue siendo difícil. Solo para algunos datos se realizan procesos de extracción, limpieza, transformación y carga de datos desde una variedad de fuentes externas.
Definido	La mayor parte de los datos de la organización están en un repositorio centralizado y tienen estructuras definidas. Aunque los datos tienen estructuras definidas, aún no se puede hacer análisis de de inteligencia de decisiones de manera fácil. Para la mayoría de datos se llevan a cabo procesos de extracción, limpieza, transformación y carga de datos, desde una variedad de fuentes externas.

<p>Gestionado</p>	<p>Todos los datos de la organización están en un repositorio centralizado, tienen estructuras definidas y se puede acceder a ellos fácilmente para su explotación. La toma de decisiones basadas en datos se puede hacer de manera ágil e inteligente.</p> <p>Existen procesos definidos y gestionados de extracción, limpieza, transformación y carga de datos, desde una variedad de fuentes externas.</p>
-------------------	---