

PROYECTO DE TRANSFORMACIÓN DIGITAL DE LA DIRECCIÓN GENERAL DE PERSONAL



Olivia Suárez Martínez
José Antonio Contreras Benítez
Carlos Alberto Román Zamitiz

Abril 2020

Transformación Digital



Fuente: Webinar "Transformación Digital" por Lucas Adolfo Giraldo Ríos de la empresa RSM de Colombia

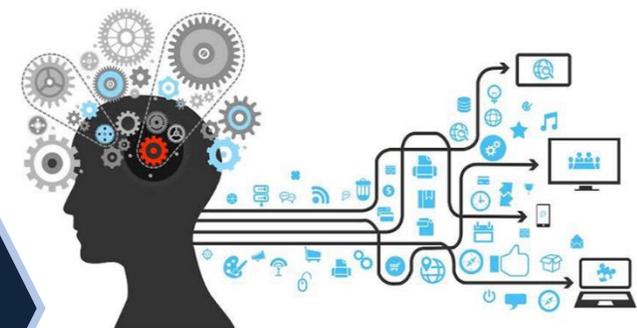


- *Micro-servicios
- *Escalable
- *Flexible

“Tendencia Mundial” Transformación Digital



En la transformación digital existen dos elementos que apuntalan este proceso, la innovación y la tecnología. Estos elementos permitirán combinar prácticas y nuevas formas de operar que darán como resultado nuevas técnicas y habilidades; pero para poner en marcha esta transformación se requiere de un cambio de actitud y de salir de la zona de confort para buscar nuevas oportunidades^{1/}.



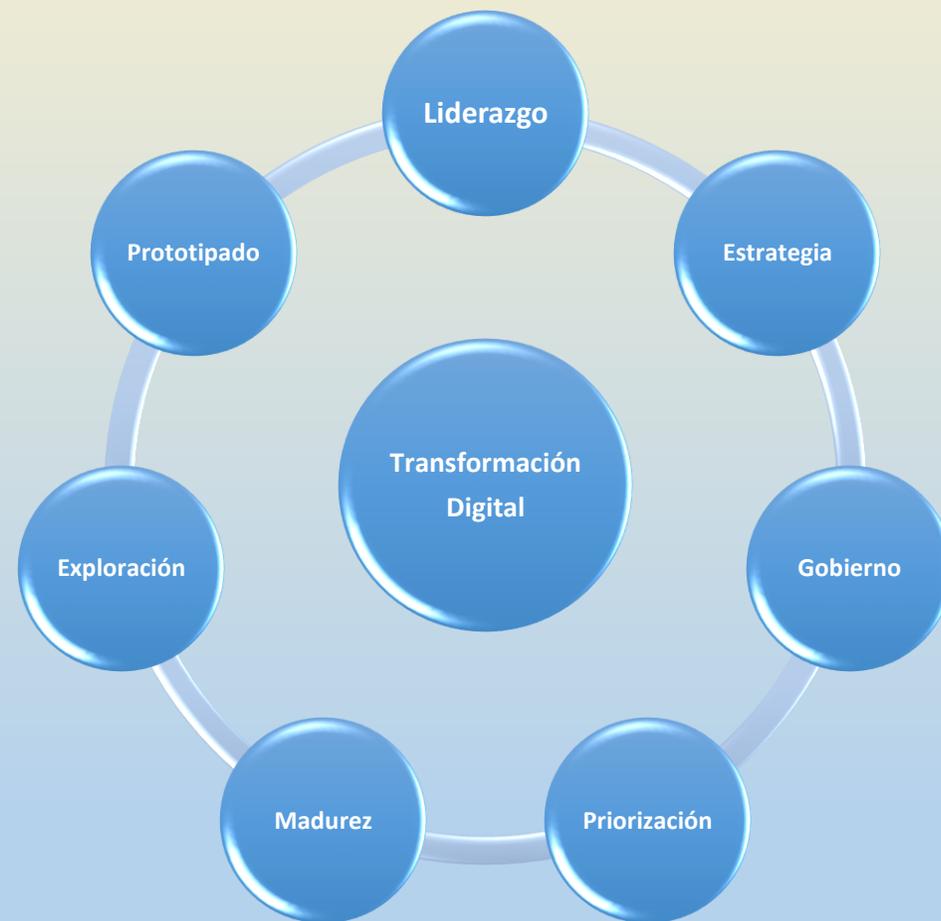
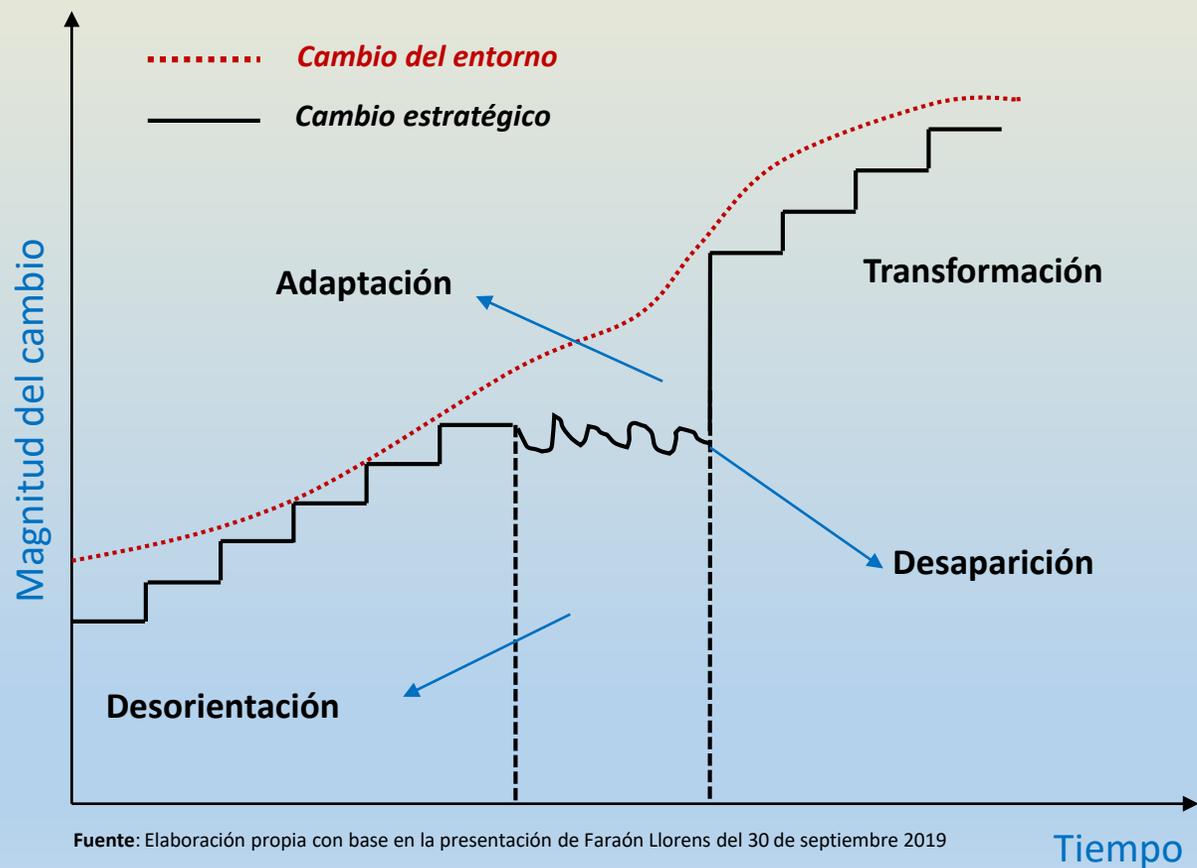
- *Minería de datos y algoritmos
- *Datos de valor
- *Patrones de comportamiento
- *Predicciones



- *Análisis de grandes datos
- *Información “No estructurada”
- *Procesamiento y análisis

^{1/}La transformación digital en las empresas. Rosana Rosas. Obtenido de <https://rosanarosas.com/transformacion-digital-empresas/>

Requerimientos de la Transformación Digital



Transformación Digital

“Hacer cosas nuevas de modos nuevos”¹

Transformación de fondo

Transformación de forma

Transformación digital

Digitalización
Automatización

Evolución digital hace referencia a la correlación continua de transformación e innovación digital con el fin de conseguir los objetivos.

Para lograr un **cambio transformador**, es necesario la implantación de un **centro o equipo de innovación**. Lo que permitirá una mayor rapidez en el cambio, con menos problemas. Servirá de apoyo en la creación de ideas y prototipos, y se podrá experimentar con ellos con el fin de desbloquear los problemas en el modelo u obtener nuevos modelos^{2/}.

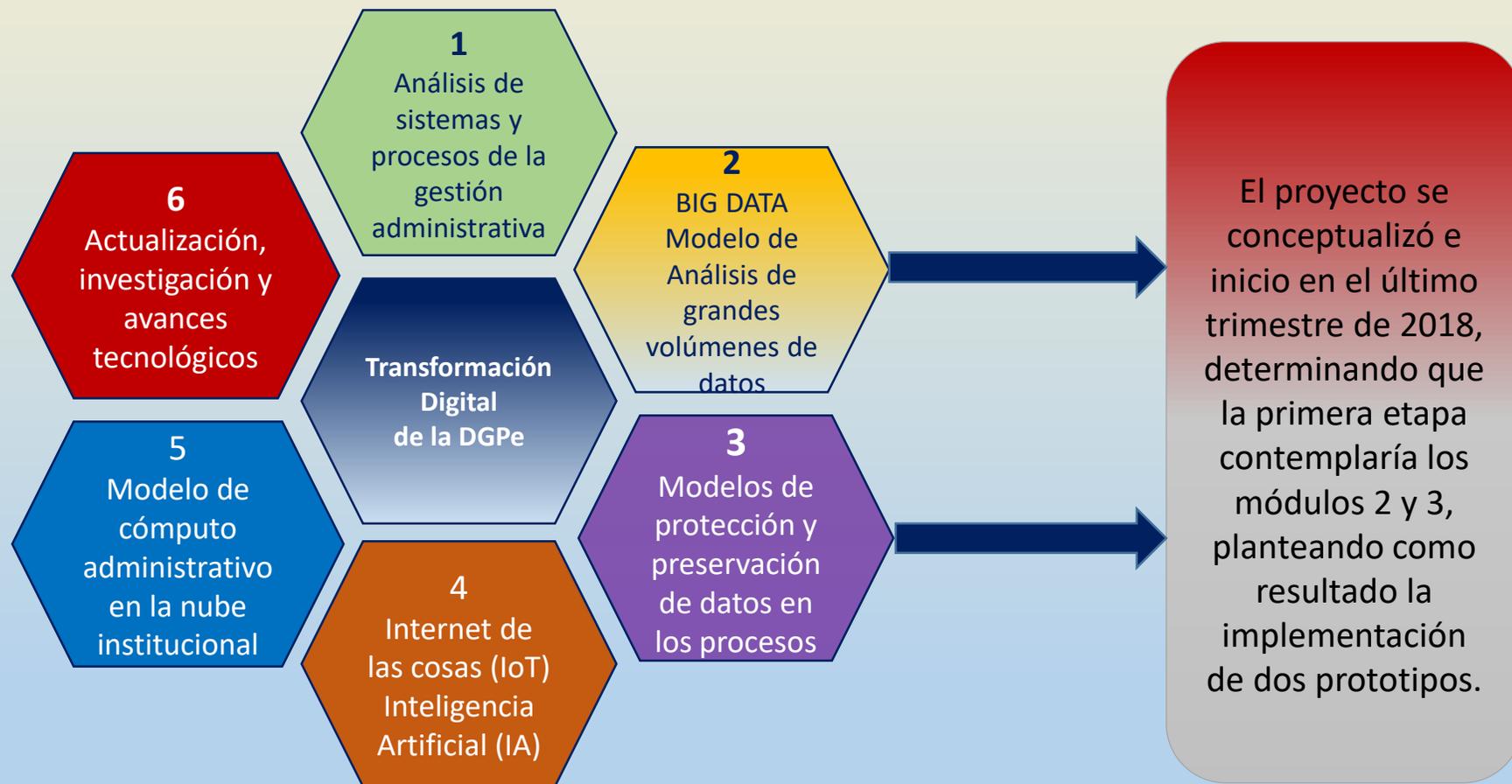
^{1/}Elaboración propia con base en la presentación de Faraón Llorens del 30 de septiembre 2019

^{2/} Correlación entre Transformación Digital e innovación. <https://www.synonize.com/transformation-digital-innovacion/>

Objetivo

Modernizar la administración de la DGPe con tecnologías inteligentes que hagan uso de la información y la conectividad para innovar la gestión de trámites y servicios con enfoque al usuario final, el ahorro y optimización de los recursos, la simplificación de procesos y el aumento en la calidad, seguridad y productividad de los servicios.

Modularidad





Alcance

- Establecer modelos predictivos para ejercicios presupuestales, mantenimiento de instalaciones, mitigación de riesgos, aprovechamiento de materiales, vigilancia y seguridad de las instalaciones.
- Contar con sistemas modernos, orientados en resolver necesidades actuales de los usuarios de trámites y servicios digitales (derechos y obligaciones con bases en los compromisos y tipos de contratación).
- Identificar patrones para un mejor desempeño del personal (habilidades, competencias, capacidad para relacionarse, evaluación, capacitación, actualización, prestaciones, etc.) en la gestión administrativa.
- Contar con un modelo de preservación digital del acervo histórico con políticas y lineamientos institucionales, conforme a normas y estándares internacionales.
- Ampliar lineamientos, procesos y sistemas de preservación digital hacia el subsistema administrativo.
- Modernizar los procesos de evaluación de la calidad para la mejora continua de la administración universitaria.
- Diseñar un modelo de cómputo en la nube para la DGPe para el análisis de grandes volúmenes de información que coadyuve en la toma de decisiones para mejorar y modernizar los procesos, trámites y servicios en la gestión administrativa.

Modelo de Análisis de grandes volúmenes de datos

PROTOTIPO “BIG DATA”

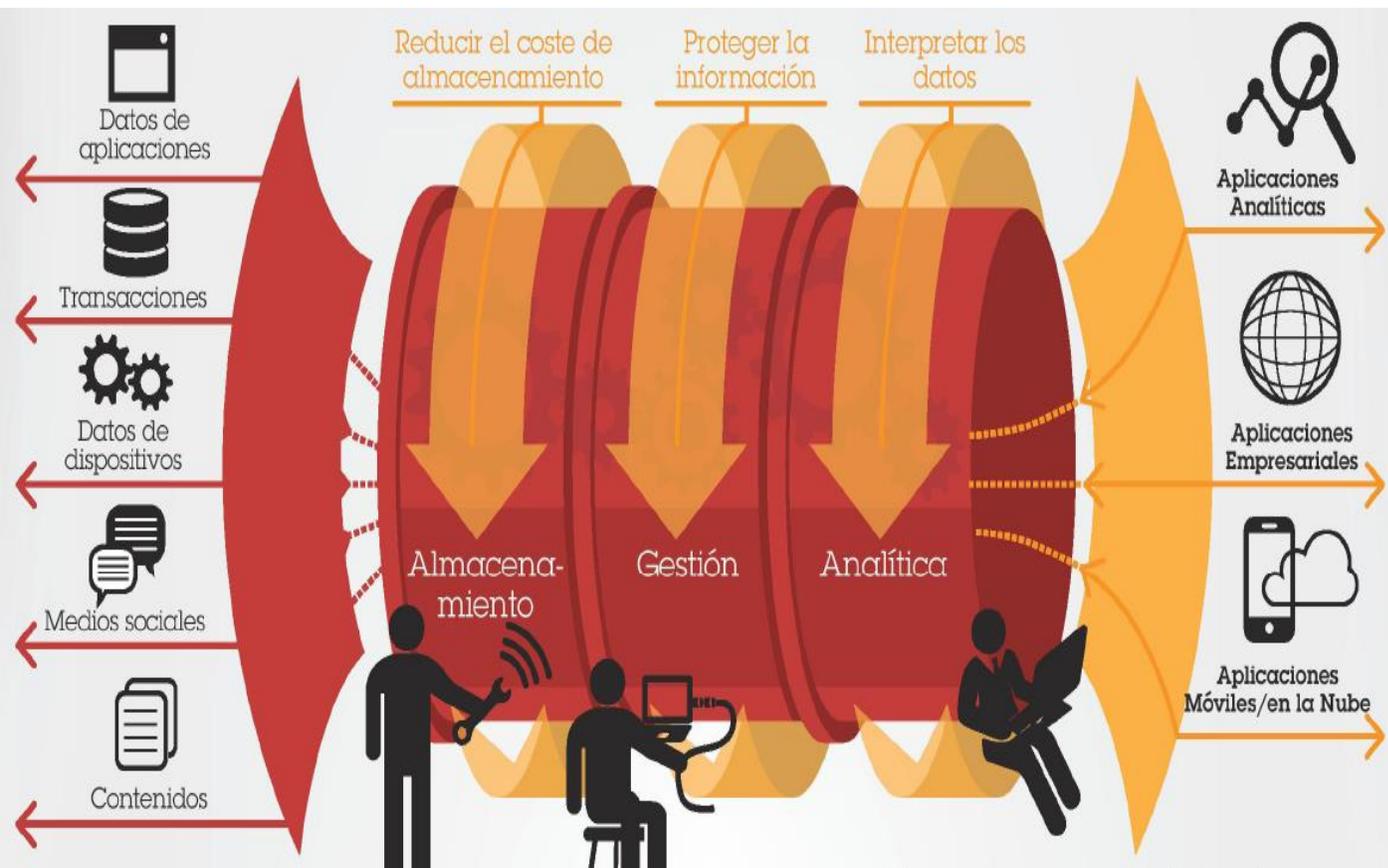


Fuente: <https://smartrural.net/wp-content/uploads/analitica-datos-smartrural.png>

Abril 2020

Modelo de Análisis de Datos

Cada día se generan en el mundo 2.5 trillones de bytes de información que se almacena de modo disperso. Para que sean de utilidad, las empresas necesitan recogerlos, identificar cuáles son importantes y extraer conocimiento para una mejor toma de decisiones.



Fuente: <https://www-03.ibm.com/press/es/es/photo/46325.wss>

Hoy no sólo hay que gestionar la información sino hacerlo de forma segura y rápida.



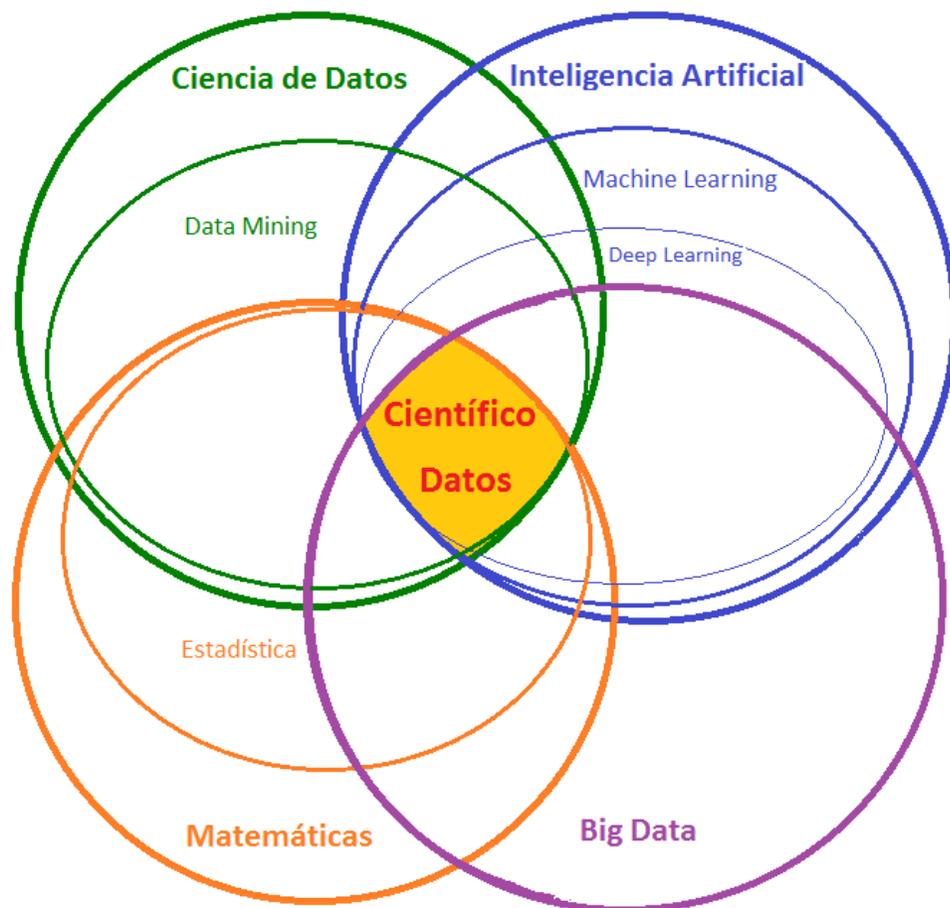
Objetivo

Implementar un modelo de análisis de grandes volúmenes de datos que permita la explotación de la información generada en la Dirección General de Personal para encontrar datos de valor y patrones de comportamiento que coadyuven a la toma de decisiones y mejoren los trámites y servicios de la gestión administrativa.

Metas

- ❖ Diseñar e implementar un ecosistema tecnológico de Big Data para la Dirección General de Personal y sus dependencias.
- ❖ Analizar grandes volúmenes de datos para obtener información de valor que aporte los elementos que permitan mejorar procesos, trámites y servicios de la Administración de la DGPe.
- ❖ Diseñar y desarrollar modelos predictivos y descriptivos aplicando técnicas inteligentes que permitan coadyuvar en la toma de decisiones para mejorar la calidad y oportunidad de los servicios.
- ❖ Capacitar al personal de las áreas de la DGPe en la aplicación de tecnologías inteligentes para generar información de valor en la toma de decisiones.

Integración del modelo de análisis

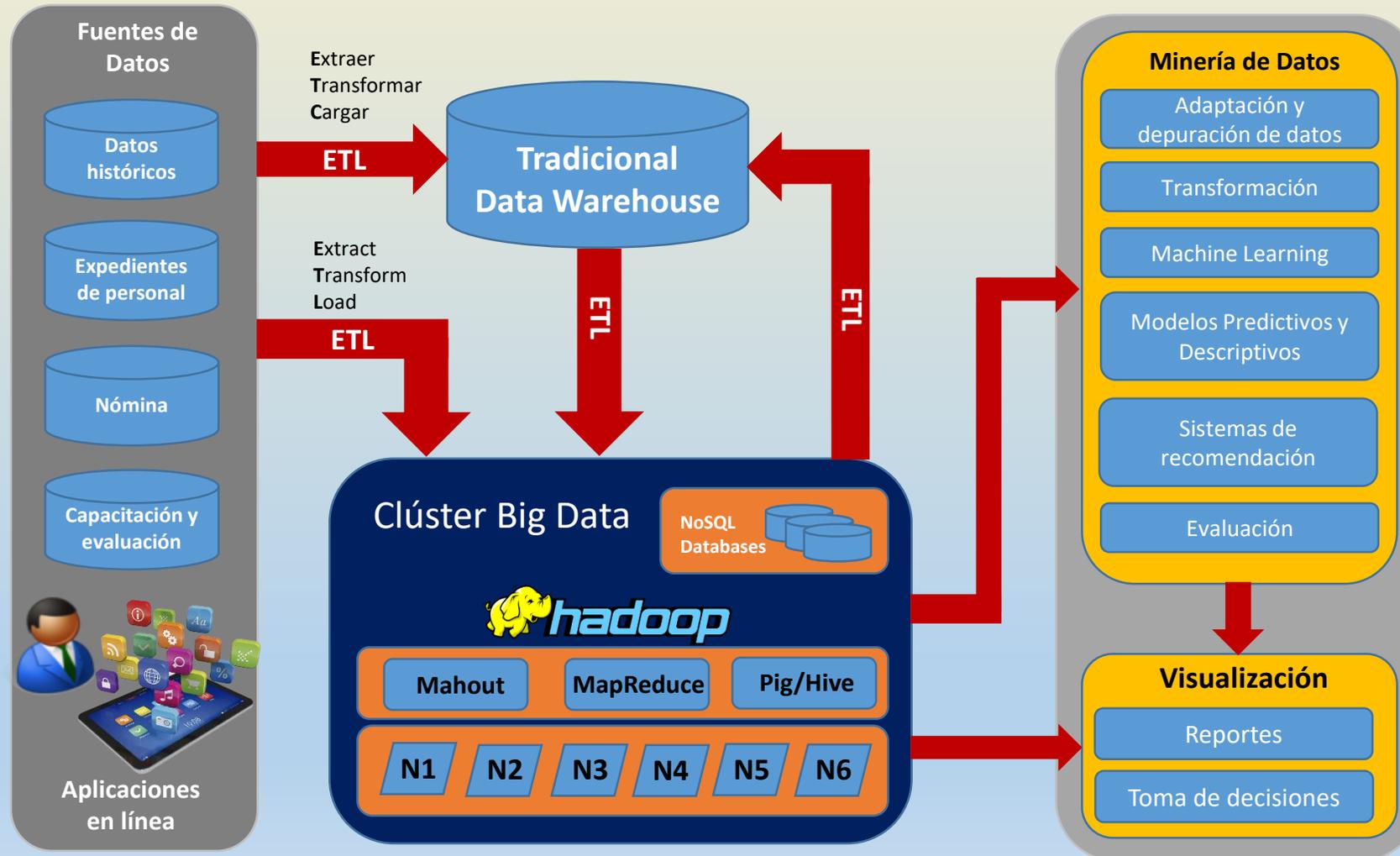


El modelo de análisis contempla la integración de diferentes ramas del conocimiento y tecnologías:

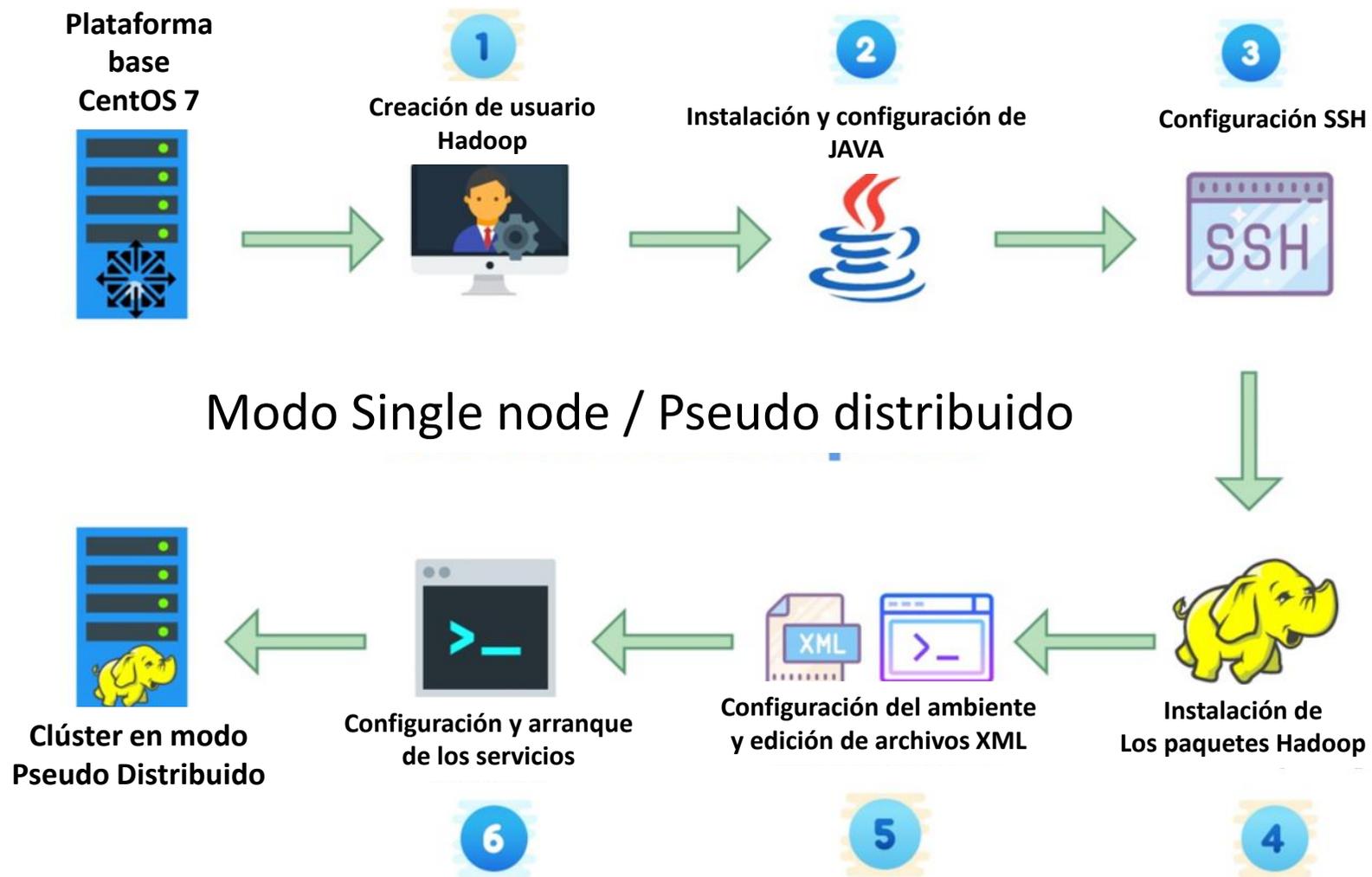
- ❖ Big Data
- ❖ Minería de Datos
- ❖ Matemáticas (estadística)
- ❖ Machine Learning
- ❖ Deep Learning

Fuente: Open Mind BBVA- Análisis de datos, Inteligencia Artificial y Big Data en la banca, 16/04/2018.
<https://www.bbvaopenmind.com/economia/finanzas/analitica-de-datos-inteligencia-artificial-y-big-data-en-la-banca/>

Esquema del Modelo de Análisis

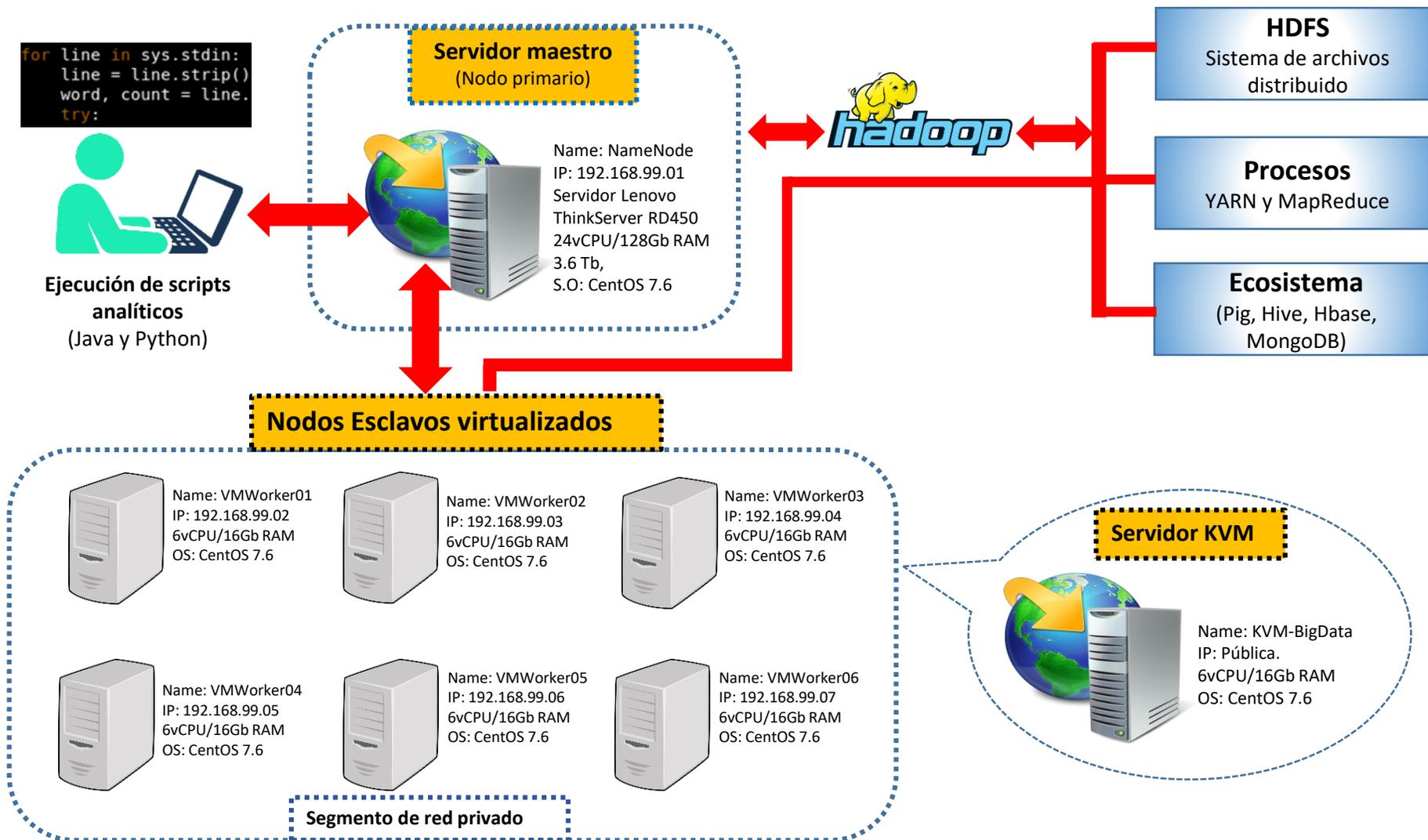


Configuración del prototipo con ambiente Hadoop

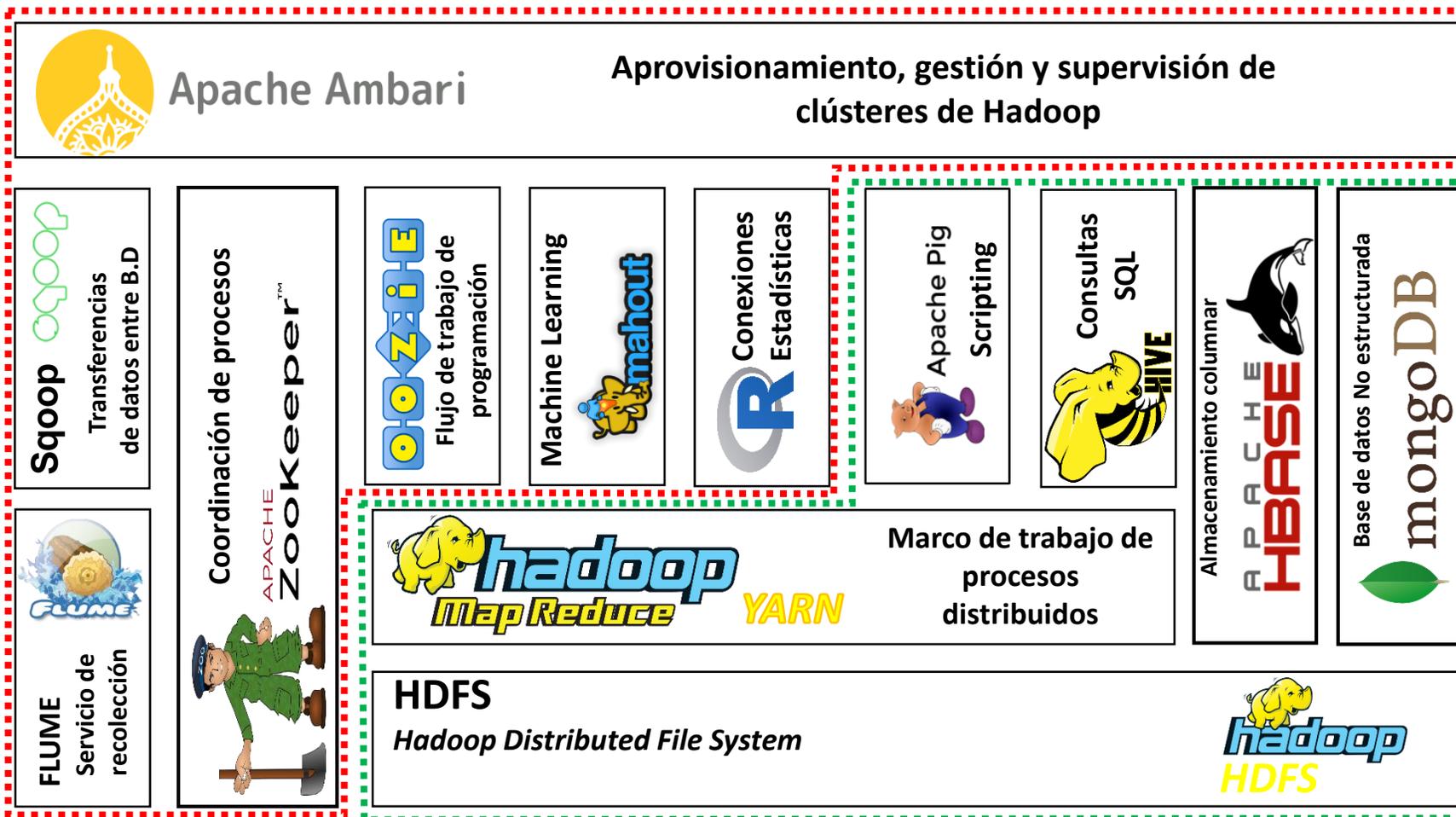




Clúster del prototipo



Ecosistema **hadoop**

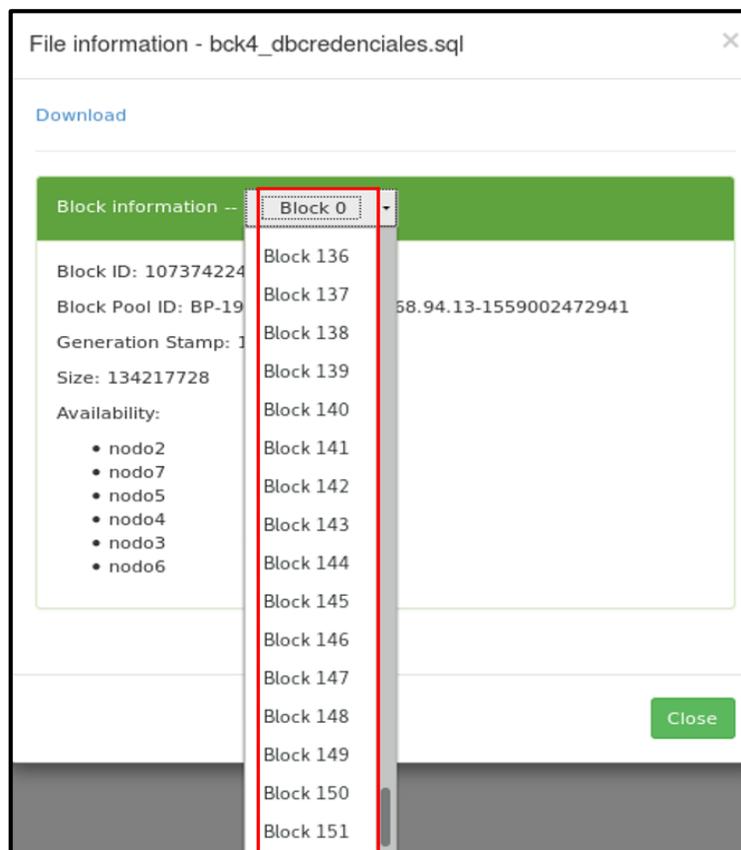
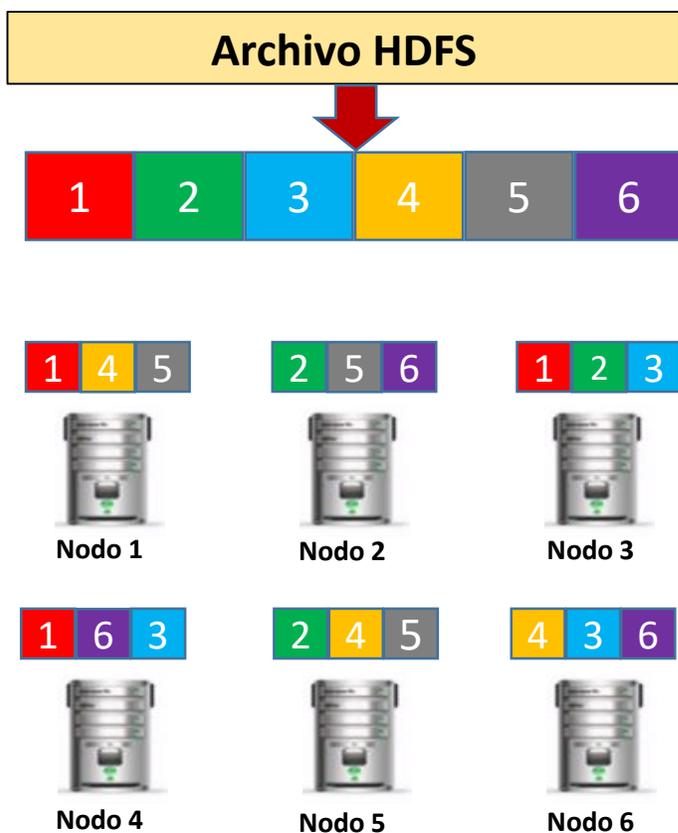


..... Pendiente
..... Implementado

Fuente: Elaboración propia con base en "Componentes Principales del Ecosistema Hadoop" <http://spaceanalytics.blogspot.com/2017/05/comp-apache-hadoop.html>

Funcionamiento del prototipo

El prototipo divide en "n" número de bloques de 128Mb, dependiendo del tamaño del archivo y los reparte en los diferentes nodos esclavos, permitiendo la ejecución simultánea de los procesos, trabajando bajo la filosofía de "divide y vencerás".

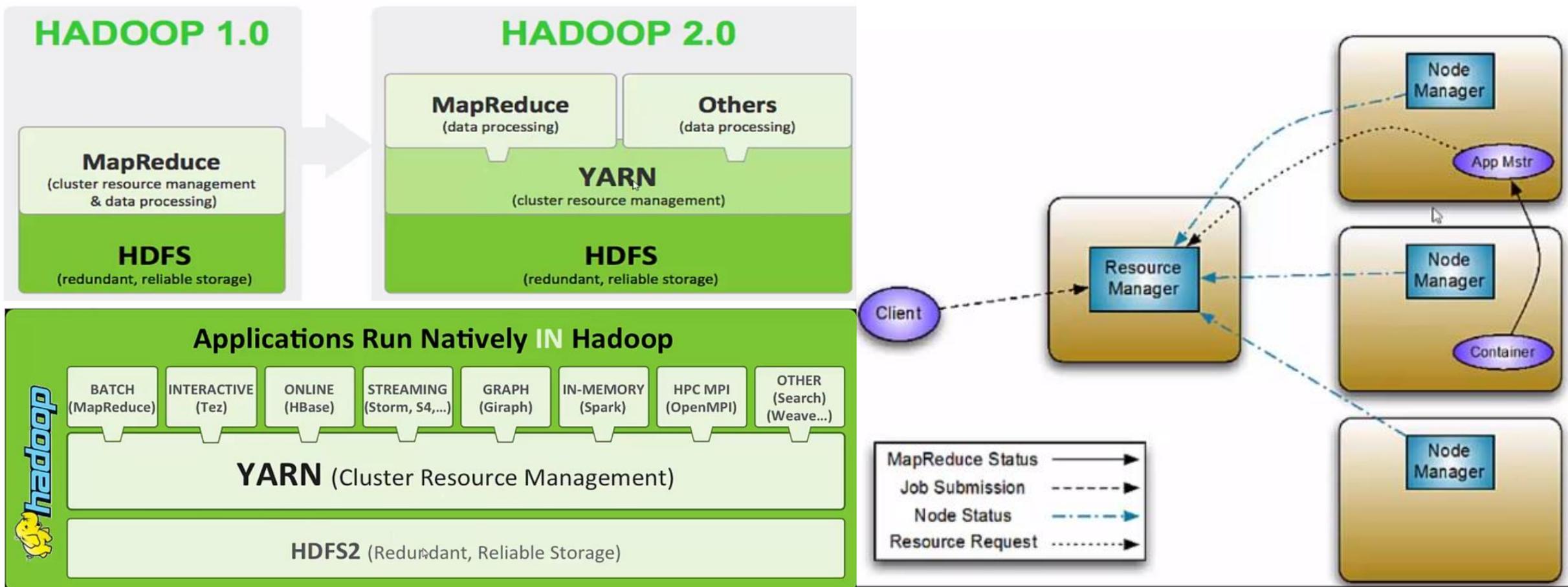




PROTOTIPO "BIG DATA"



Procesos Hadoop



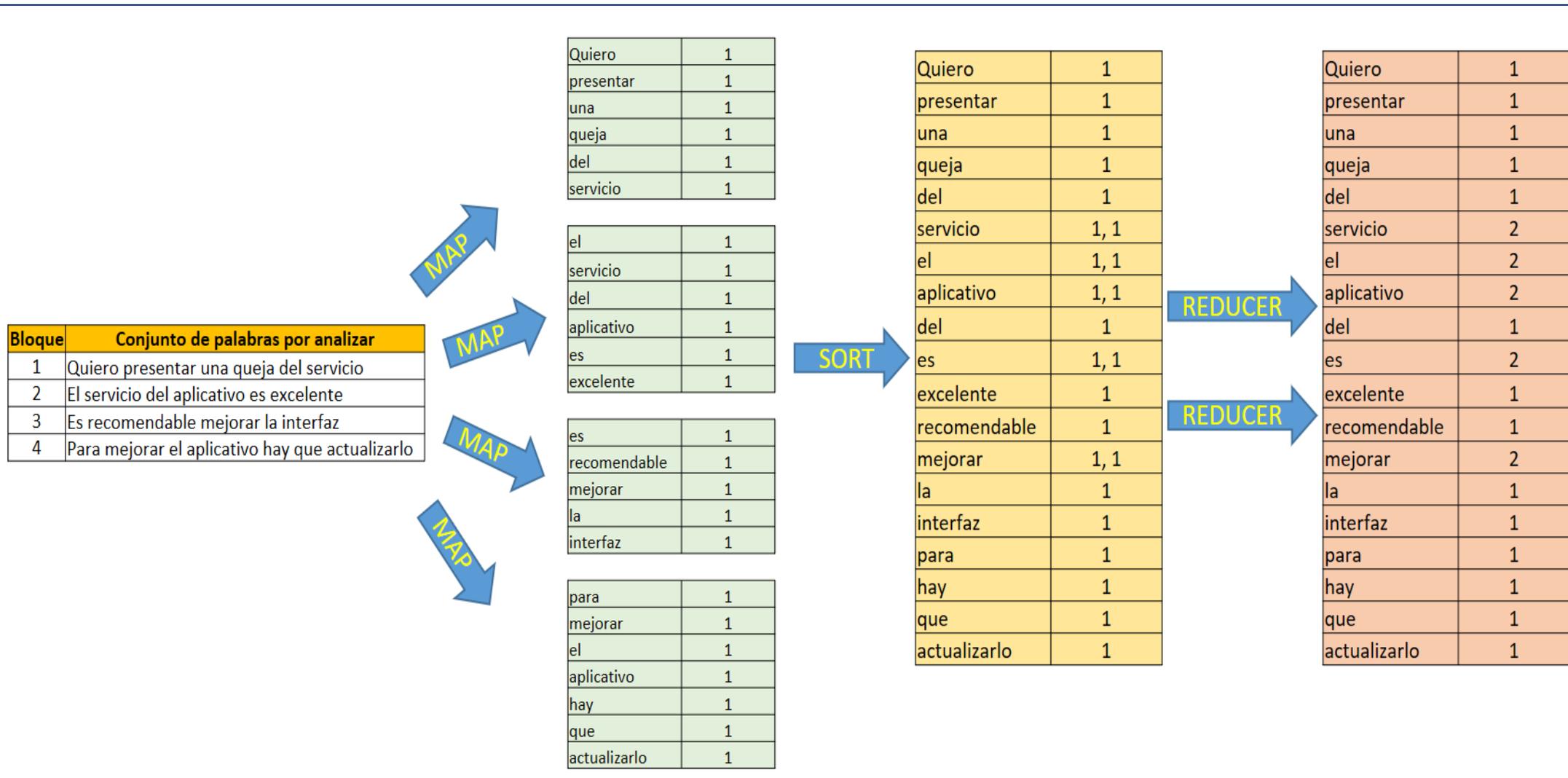
Fuente: Apasoft Formación 2018



PROTOTIPO "BIG DATA"

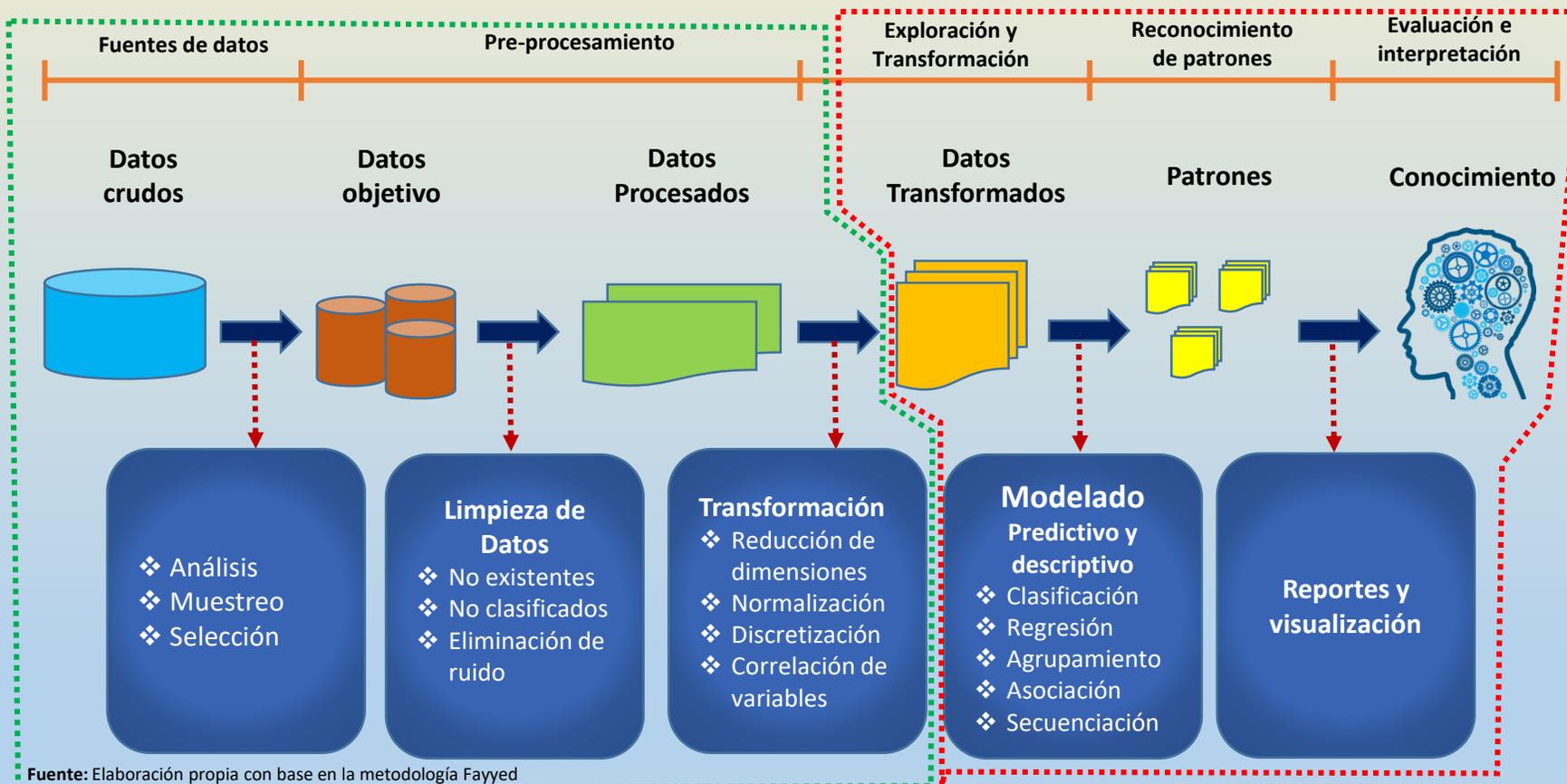


FUNCIONAMIENTO DE MAPREDUCE



Minería de Datos

Se inició el proceso de minería de datos con la información de las bases de datos de bitácoras y credenciales de la DGPe.





Video Demo del prototipo



Avances y/o actividades

- ❖ Investigación de la arquitectura Big Data para la definición del modelo.
- ❖ Implementación del clúster Hadoop de manera pseudo-distribuida y distribuida.
- ❖ Análisis y pruebas con herramientas del ecosistema Big Data.
- ❖ Ejecución de scripts básicos para manipulación de archivos
- ❖ Instalación y prueba de herramientas de minería de datos (Weka y RapidMiner).
- ❖ Investigación y estudio de los principales algoritmos de Machine Learning.

Proyección

- ❖ Complementar las pruebas del prototipo
- ❖ Complementar el ecosistema Hadoop - Big Data
- ❖ Implementar uno o más modelos predictivos para procesos específicos de la DGPe.
- ❖ Implementación de un KDD (Knowledge Discovery in Databases) para la DGPe.
- ❖ Implementación de un sistema de recomendación



Proyección

Requerimientos

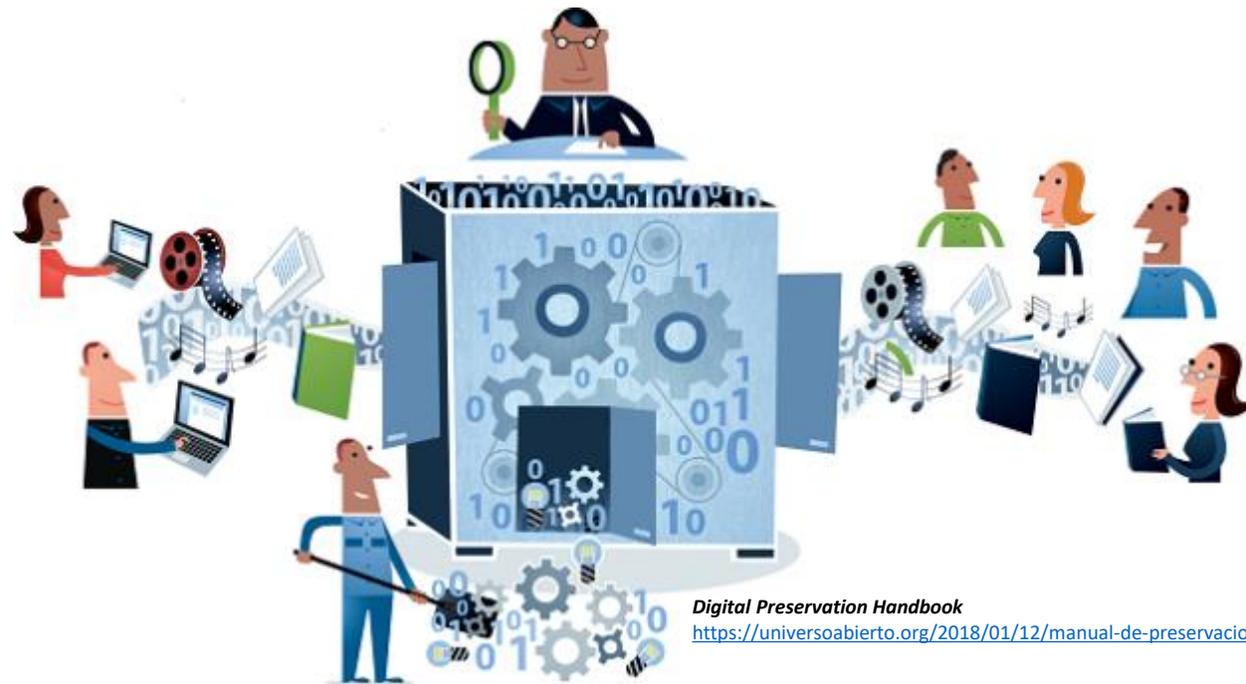
❖ Normativos

- ❖ Adaptar el modelo de análisis referente a Big Data y ciencia de datos.
- ❖ Apegarse a los aspectos éticos y legales para el análisis de información conforme a la Ley de Protección de Datos Personales, los Lineamientos para la Protección de Datos Personales en Posesión de la Universidad, y las Normas Complementarias para la Seguridad Técnica, Física y Administrativa para la PDPPU.
- ❖ Directrices de datos útiles que ayuden a la toma de decisiones y poder plantear los modelos de acuerdo a la información que se tiene.

❖ Técnicos

- ❖ Se requieren bases de datos de mucho mayor volumen para probar el prototipo Hadoop.
- ❖ Escalar la arquitectura del clúster cuando se aperture el análisis al resto de las direcciones de la DGPe.
- ❖ Es indispensable la participación de las áreas que cuentan con grandes volúmenes de información para identificar datos de valor y posibles modelos predictivos.
- ❖ Promover la apertura de un laboratorio en la DGPe con la temática de analítica de datos.
- ❖ Capacitación y actualización en tecnologías emergentes (Ciencia de Datos, I.A., etc.)

PLAN DE PRESERVACIÓN DIGITAL PROYECTO “BÓVEDA DIGITAL”



Digital Preservation Handbook
<https://universoabierto.org/2018/01/12/manual-de-preservacion-digital/>

Abril 2020

Vinton Cerf: "Vendrá una Edad Oscura si no logramos preservar los datos digitales"

"... las próximas generaciones no van a conocer nuestra historia, porque vamos a perder un montón de datos y de programas que nos explican. La evolución de los formatos ligada a sistemas operativos va a hacer que nadie recuerde cómo éramos. Nos espera una Edad Oscura Digital si no somos capaces de preservar nuestros textos, fotos y vídeos."



Vinton Cerf uno de los padre del Internet <https://www.elperiodico.com/es/sociedad/20160621/vinton-cerf-entrevista-5219022>



Fuente: <https://es.slideshare.net/DataReportal/digital-2019-mexico-es-january-2019-v03>



Información y documentos de la DGPe que son tema para un Plan de Preservación Digital

Ejemplo de series de tramites por año que utilizan información de los Sistemas Institucionales y son respaldan con documentos digitales

Tramites o servicios que utilizan o generan documentos	2018	2017	2016	2015	2014	2013
Versiones de Sistema Operativo y/o PDF		PDF Versión 2.0		Windows 10		Windows 8.0, PDF versión 1.7
Nombramientos de alta	114,110	106,536	104,076	94,214	91,341	86,779
Bajas de personal	19,508	17,227	17,831	16,494	16,134	16,067
Licencias de personal		25,490	20,223	17,460	13,430	10,758
Pagos por cheque y deposito	1,791,491	1,769,786	883,355	843,544	829,531	844,106
Aclaraciones a nómina	150,409	156,158	140,667			
Movimientos de personal	210,617	203,176	196,705	173,247	163,968	154,349
Constancias de sueldos del SAT		53,495	53,660	53,099	48,786	40,941
Total	2,286,135	2,331,868	1,416,517	1,198,058	1,163,190	1,153,000

Fuente: UNAM - Dirección General de Planeación- Memoria UNAM



Organismos o iniciativas comprometidos con la Preservación Digital

UNESCO

“La preservación del patrimonio documental abarca un conjunto de técnicas, tratamientos, procedimientos y tecnologías de carácter muy diverso, de tipo preventivo y correctivo, destinado a preservar los documentos y la información que en ellos figura^{1/}”.

InterPARES

“La **preservación digital** es el conjunto de principios, políticas, normas y estrategias diseñadas para asegurar que un objeto digital permanezca accesible, inteligible y usable a través del tiempo y de los cambios tecnológicos, su fiabilidad y exactitud están protegidas y su autenticidad es verificable^{2/}”.

Grupo de Preservación Digital de la UNAM

Preservación Digital: Combinación de estrategias y actividades encaminadas a garantizar el acceso a largo plazo, la integridad y autenticidad de los contenidos digitales^{3/}.

^{1/} <https://unescopersist.files.wordpress.com/2018/02/unesco-persist-guidelines-traducida-al-espac3b1ol.pdf>

^{2/} International Research on Permanent Authentic Records in Electronic Systems (InterPARES)

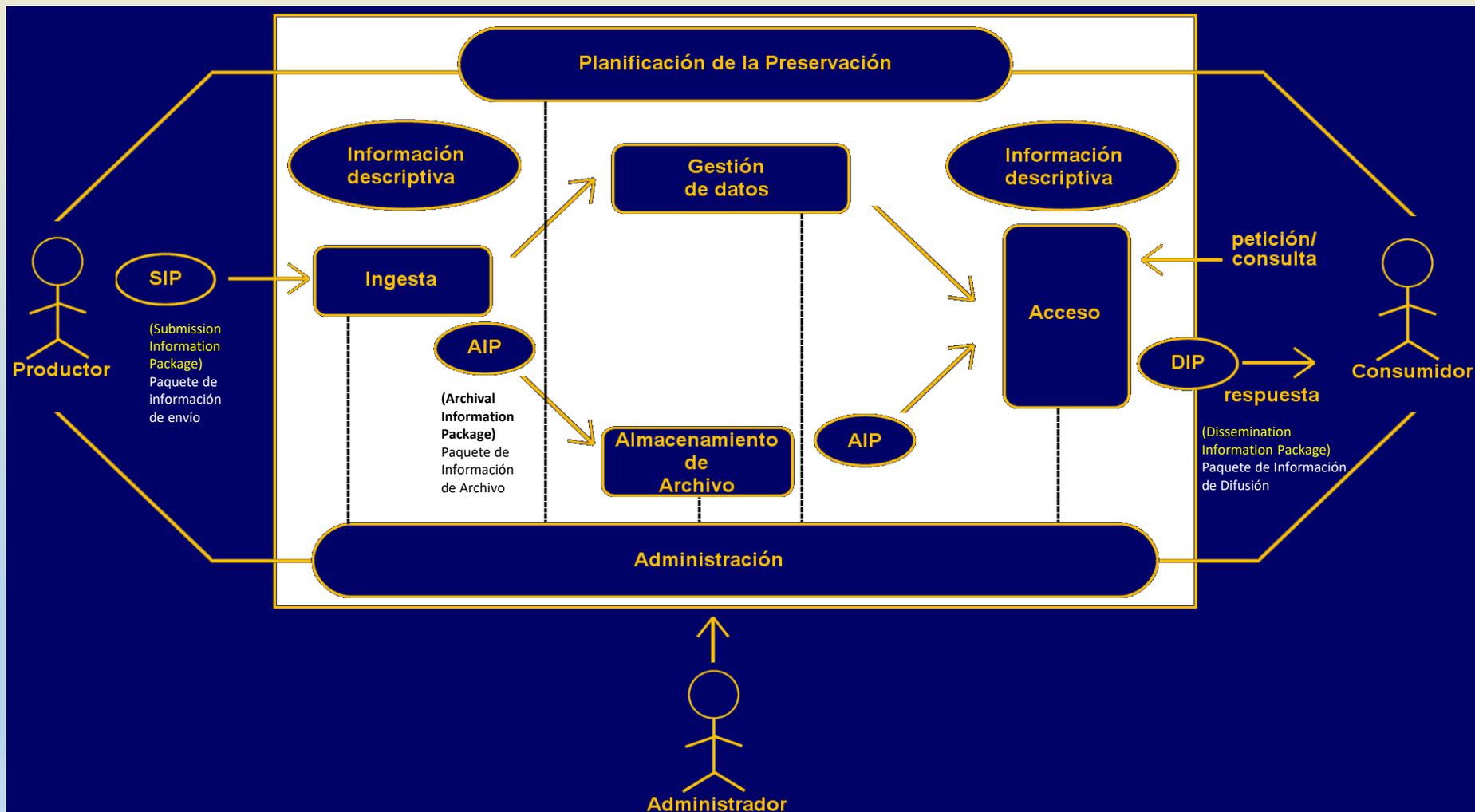
^{3/} <http://www.preservaciondigital.unam.mx/txt/base-comunidades.html>



Directrices y Estándares

La norma ISO 14721:2003 regula las funciones de preservación de archivos: almacenamiento de archivos, gestión de datos, migración de la información digital a otros formatos, ingesta, acceso y difusión. Esta norma define que el modelo de referencia para un Sistema Abierto de Archivo de Información (modelo OAIS) abarca todos los aspectos de la preservación, a largo plazo.

El Modelo de Referencia OAIS (ISO 14721), es el estándar más utilizado para la Preservación Digital:





Propuesta de modelo de preservación digital para la DGPe

PLAN DE PRESERVACIÓN DIGITAL: Define las políticas y lineamientos de los objetos a preservar, autoriza usuarios y roles, e identifica los recursos humanos, materiales y financieros.

INGESTA

Preparación de los objetos a preservar y gestionar:

- ▶ Identificación de formatos
- ▶ Verificación de integridad
- ▶ Definición de metadatos
- ▶ Generación del SIP

GESTIÓN Y ARCHIVADO

Transformación de los objetos y aplicación de procesos para su conservación a largo plazo:

- ▶ Generación del AIP
- ▶ Procesos de verificación y mantenimiento
- ▶ Actualización de objetos
- ▶ Seguridad
- ▶ Almacenamiento

CONSULTA

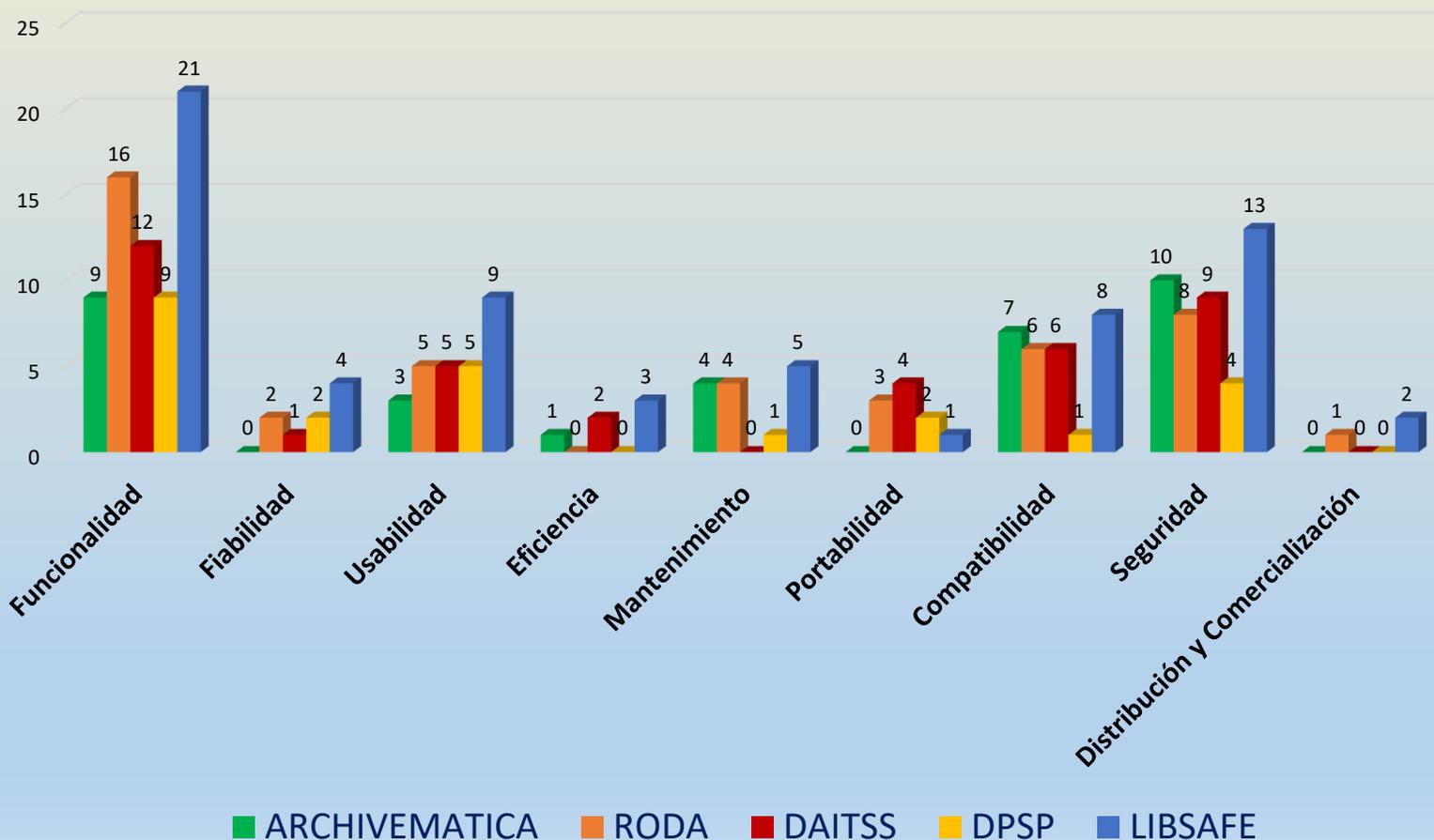
- ▶ Interfaz gráfica de acceso
- ▶ Búsqueda de los objetos digitales
- ▶ Acceso a los objetos preservados
- ▶ Generación del DIP
- ▶ Medidas de seguridad para la visualización y manipulación

ADMINISTRACIÓN: Implementa y administra los elementos tecnológicos, para ofrecer un servicio seguro, integro, disponible y confiable



Evaluación de la tecnología

Análisis de características de los SPD evaluados



Con base "Guía de criterios básicos para valorar sistemas de preservación digital" elaborada por el Grupo de Preservación Digital (GPD) de la UNAM, se concluyó que:

- ❖ LIBSAFE contempla 66 de 73 criterios
- ❖ RODA contempla 45 de los 73 criterios.

El origen de RODA es open source, razón por la que existe una versión básica y otra ampliada (Enterprise)



PROTOTIPO "BÓVEDA DIGITAL"



Comparativa de funcionalidades y servicios de las plataformas de Preservación Digital

Elemento del modelo	Detalle de elementos	Libsafe	Roda Comercial	Roda Open Source
Pre-ingesta	Pre-Ingesta automática	✓	X	X
	Pre-Ingesta manual	✓	✓	✓
Ingesta	Validación	✓	✓	✓
	Caracterización	✓	✓	✓
	Identificación del formato	✓	✓	✓
	Conversión	✓	✓	X
	Reindización	✓	✓	✓
	Metadatos personalizados	✓	✓	✓
Archivado	Copias ilimitadas	✓	✓	✓
	Distribuidas geográficamente	✓	✓	✓
	Compatibilidad de almacenamiento	✓	✓	X
	Registro de incidencias en B.D.	✓	X	X
Diseminación	Acceso/Consulta	✓	✓	X
	Emulación	✓	X	X

Elemento del modelo	Detalle de elementos	Libsafe	Roda Comercial	Roda Open Source
Aspectos de soporte	Asesoría	✓	✓	X
	Capacitación	✓	✓	X
	Instalación y configuración	✓	✓	X
	Soporte técnico	✓	✓	X
Administración y control	Granularidad de usuarios y roles	✓	X	X
	Informes básicos	✓	✓	✓
	Informes robustos y gráficos	✓	X	X
	Plan de preservación técnico	✓	✓	X
	Gestión de Riesgos	✓	✓	✓
Seguridad	Firma Digital	✓	X	X
	Auditorías	✓	X	X
Experimentales	Procesos de optimización y relacionados con IA	✓	X	X



PROTOTIPO “BÓVEDA DIGITAL”

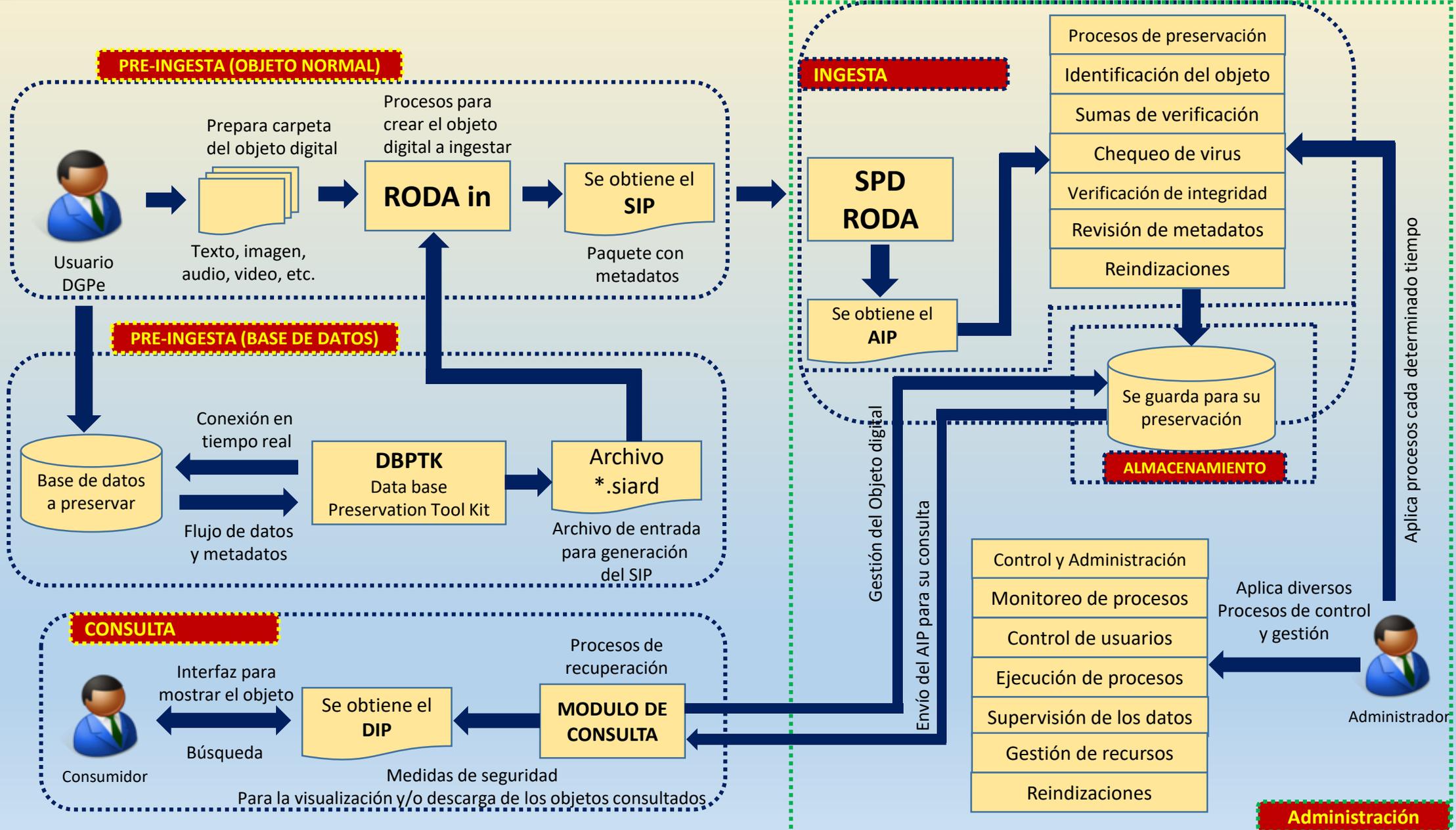


Funcionalidades del Prototipo de la DGPe con RODA Open Source

INGESTA	PROCESOS DE PRESERVACIÓN	ACCESO/CONSULTA
<p>Pre-ingesta Herramienta RODA-in para preparación del objeto que se ingesta (incluye plantilla de metadatos para los diferentes objetos)</p> <ul style="list-style-type: none">▶ Soporte de diversos formatos incluyendo los más comunes (PDF, TIFF, MPEG-2, WAV)▶ Autenticación de usuarios (Open LDAP)▶ Google Web Toolkit (interfaz de usuario)▶ Asignación de identificadores únicos▶ Identificación del objeto digital para validación automática (JHove / JHove2 y DROID)▶ Verificación de virus▶ Validación de metadatos▶ Herramienta DBPTK para bases de datos (Mysql, Postgresql, Oracle, Sybase, etc.)	<ul style="list-style-type: none">▶ Monitoreo del AIP para virus▶ Chequeo de integridad del AIP▶ Validación, modificación y actualización de metadatos.▶ Procesos de indexación▶ Reconstrucción de repositorio▶ Procesos batch▶ Identificación de riesgos▶ Almacenamiento replicado y distribuido▶ Granularidad de comunidades y usuarios básica	<p>Módulo desarrollo en JAVA que permite:</p> <ul style="list-style-type: none">▶ Recuperar los archivos preservados (creación del DIP)▶ Búsqueda básica de los objetos▶ Procesos de conversión y visualización▶ Aplicación de medidas de seguridad en la consulta de objetos.



Esquema de operación del prototipo

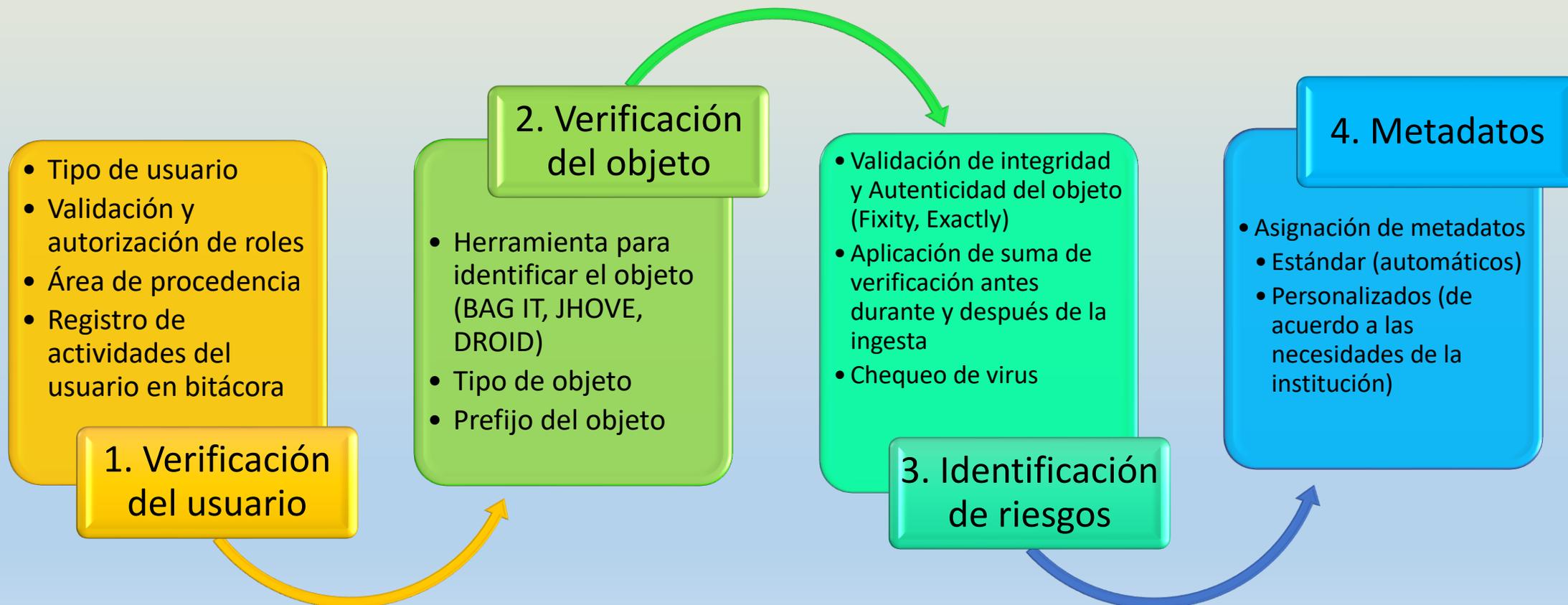




Video Demo del prototipo

Propuesta de Procesos para el Plan de Preservación Técnico (Parte 1)

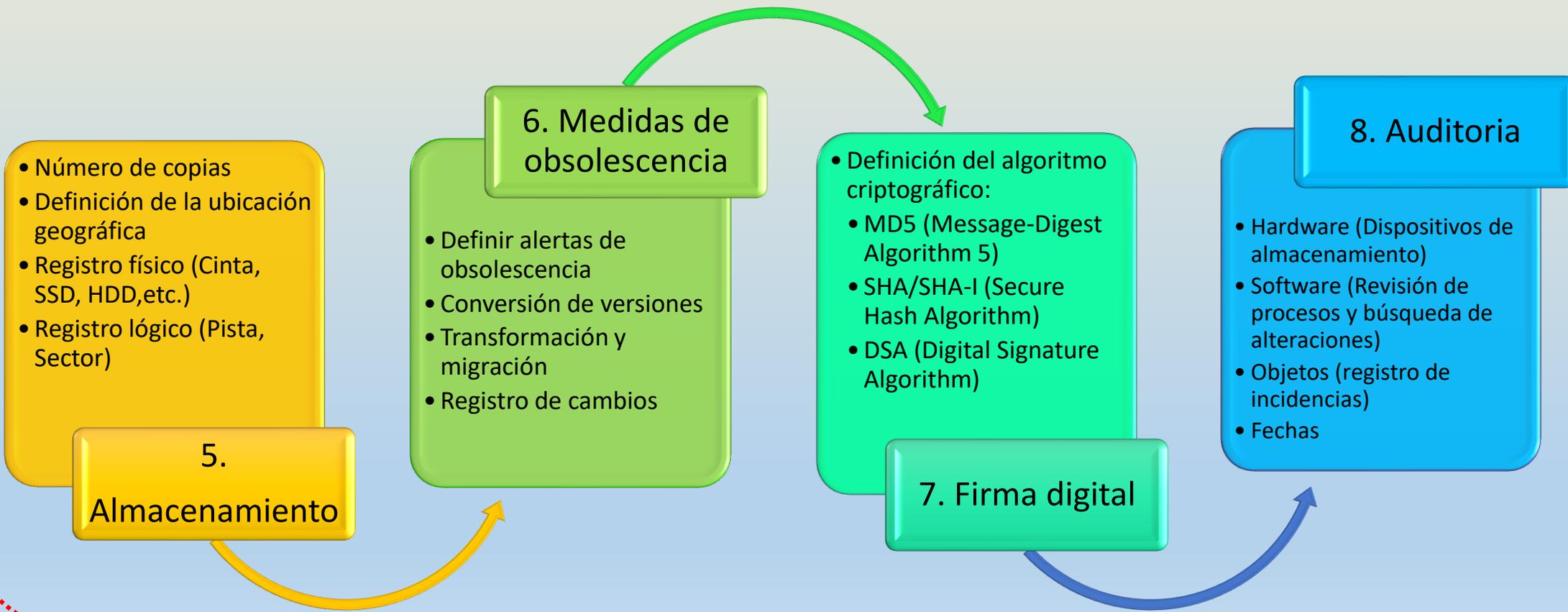
El análisis y evaluación de diversos sistemas de preservación, han servido de base para la definición de los procesos técnicos para los objetos digitales. A continuación se muestra el esquema de procesos considerados en el prototipo con el sistema RODA:



Propuesta de Procesos para el Plan de Preservación Técnico (Parte 2)

A continuación, se muestran los procesos propuestos que, de manera ideal, complementarían el sistema:

Módulos por implementarse (complementarios al prototipo)





Avances

- ❖ Investigación y análisis de los modelos existentes en preservación digital.
- ❖ Investigación y estudio de tecnología (SPD y Herramientas).
- ❖ Comparativa de los principales SPD y herramientas de preservación digital.
- ❖ Definición de la Plataforma Tecnológica a utilizar (RODA y DSpace).
- ❖ Implementación del prototipo.
 - ❖ Preparación de la infraestructura básica (Hardware y Software).
 - ❖ Instalación y configuración del sistema.
 - ❖ Desarrollo del módulo de consulta.
 - ❖ Ingesta de datos de diversos formatos (Texto, Imagen, Audio, Video, etc.).
- ❖ Pruebas de ingesta con:
 - ❖ Información de la DGPe (Base de Datos de Credenciales y de bitácoras).
 - ❖ Archivos digitales de hasta 200 GB.
 - ❖ Archivos de bases de datos.
 - ❖ Información en formato PDF de la Dirección de Relaciones Laborales de la DGPe.
- ❖ Elaboración de la guía del usuario (Se tiene una versión en borrador).
- ❖ Colaboración en el grupo de tecnología del GPD.

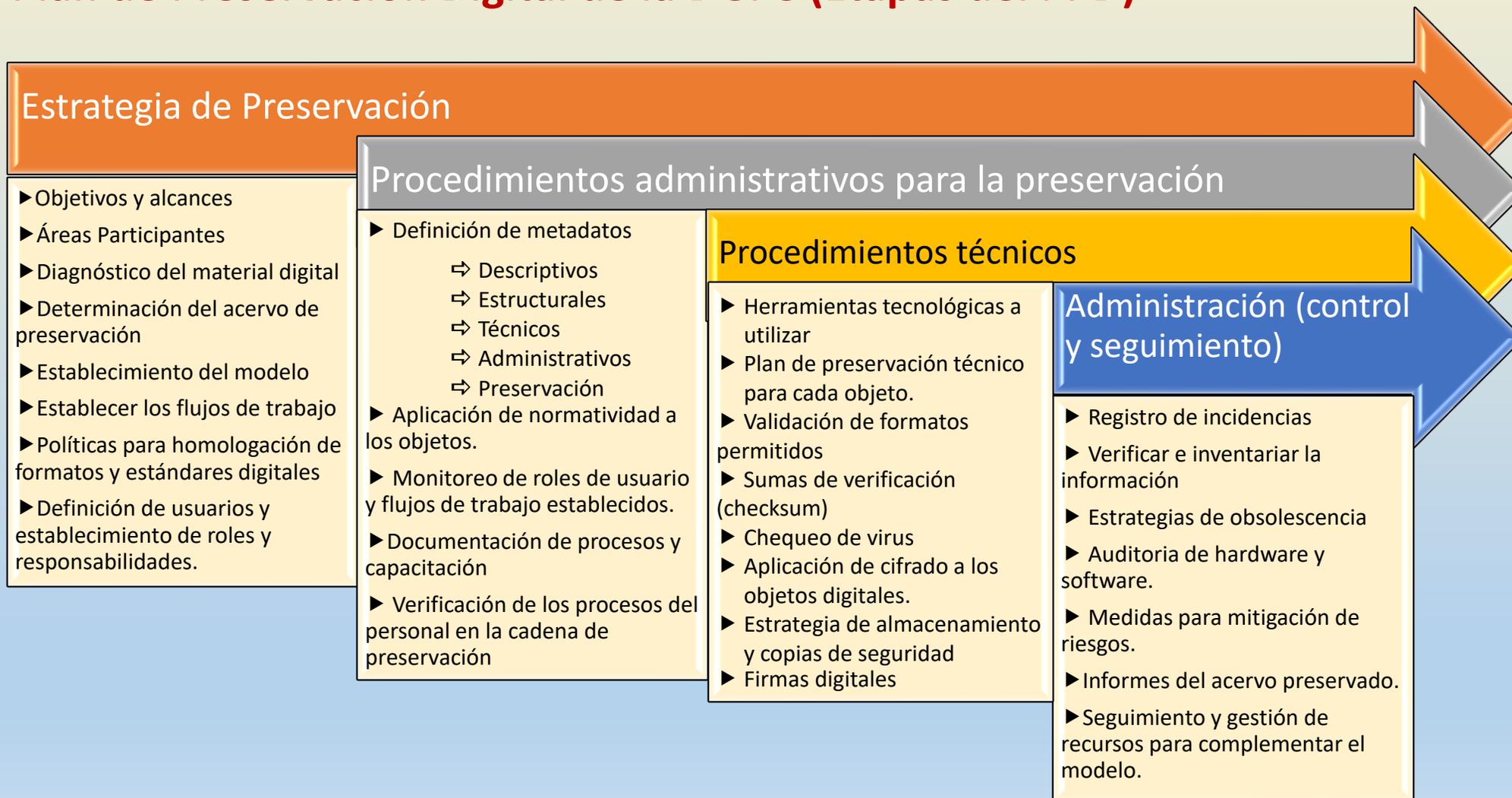


Pruebas de casos de uso

Requerimientos y/o actividades

- ❖ Seleccionar el área para la prueba de caso de uso.
- ❖ Entrevista con el área seleccionada para determinar el acervo a preservarse.
- ❖ Definición de políticas y procedimientos.
- ❖ Análisis de información, documentos y formatos.
- ❖ Determinar suficiencia tecnológica en la capacidad de almacenamiento.
- ❖ Elaborar un plan de preservación técnico para el acervo específico del área seleccionada.
- ❖ Presentación del modelo y definición de actores y roles.
- ❖ Pruebas de prototipo
- ❖ Elaboración de la guía de usuario.
- ❖ Capacitación para la ingesta de datos y consultas.
- ❖ Monitorización y seguimiento.
- ❖ Establecer mecanismos de evaluación.
- ❖ Ajustes al prototipo.

Plan de Preservación Digital de la DGPe (Etapas del PPD)





Requerimientos:

➤ Normativos

- ❖ Plan de preservación institucional
- ❖ Establecer lineamientos para los procesos de digitalización de acervo potencial de preservación para la DGPe.

➤ Técnicos

- ❖ Elaborar un plan de preservación técnico para cada formato a preservarse en el prototipo.
- ❖ Mejorar las opciones de búsqueda y recuperación del módulo de consulta del prototipo, para mejorar el acceso al material preservado.
- ❖ Adicionar un módulo con DSpace que interactúe con RODA para ofrecer una mayor granularidad en los tipos de usuarios, comunidades y colecciones.
- ❖ Diseñar e implementar un módulo de **“Estrategias de obsolescencia”** que contemple procesos automatizados de conversión y migración de formatos para garantizar la actualización de las versiones de los objetos digitales.



Requerimientos:

➤ Técnicos (Cont.)

- ❖ Diseñar e implementar mecanismos de auditorías que permitan revisar las incidencias con los objetos digitales preservados, tanto a nivel hardware, software y contenido digital.
- ❖ Ampliar la infraestructura de almacenamiento en diferentes ubicaciones geográficas (Contar con un Plan de Recuperación de Desastres).
- ❖ Controlar, administrar y gestionar los objetos digitales que se vayan ingestando en el prototipo.

➤ Colaborativos

- ❖ Seguir participando en grupos de Preservación Digital dentro y fuera de la UNAM y avanzar sobre estrategias y lineamientos sólidos a nivel institucional y nacional.
- ❖ Ser parte de la comunidad RODA open source para seguir complementando y actualizando el prototipo planteado.



La teoría de la evolución de las especies de Charles Darwin establece que los individuos que tienen más probabilidades de sobrevivir no son los más fuertes, sino los que mejor se adaptan al cambio. La tecnología digital ha transformado la forma en que las empresas se relacionan con su entorno, de forma que, si aplicamos la teoría de Darwin al mundo empresarial, es probable que en un futuro próximo solamente sobrevivan las empresas que mejor se adapten a la Transformación Digital ^{1/}.

Conclusiones

Los recursos financieros y tecnológicos son importantes para la transformación digital, sin embargo, lo esencial para lograrla, es contar con un compromiso profundo de todas las áreas y niveles de la institución para cambiar la forma tradicional de realizar las cosas.

Propuesta para lograr la transformación Digital

En México y en particular en la UNAM, se debe conformar un ecosistema similar a los que se están gestando en Europa¹, integrado por las siguientes comunidades y recursos:

Humanos

Investigadores, desarrolladores, arquitectos y científicos de datos, responsables de las áreas de TI.

Socioculturales

Especialistas en ética, responsables de las áreas operativas, legales y educativas.

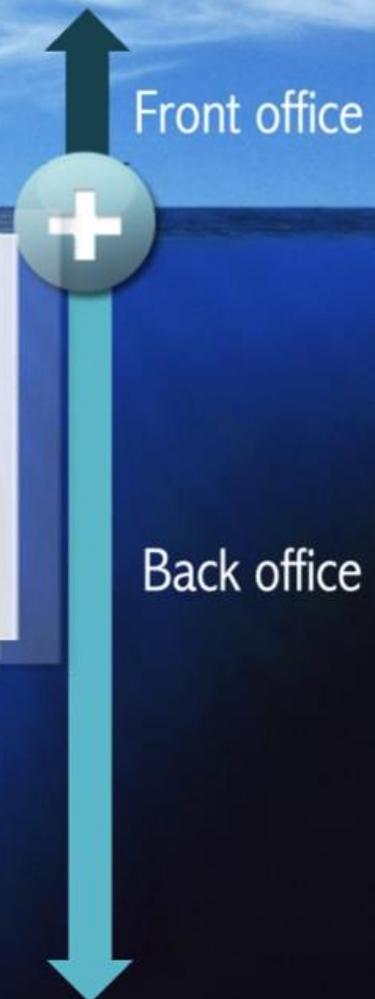
Económicos

Responsables que gestionen los recursos financieros o convenios de colaboración para el logro del proyecto.

Infraestructura

Hardware, telecomunicaciones, software, sistemas, herramientas, etc.

Cualquier esfuerzo para transformar la *front office* no llegará lejos a menos que nuevos sistemas digitales sean enganchados profundamente en el núcleo.



CONTACTO

Mat. Olivia Suárez Martínez
osuarez@dgp.unam.mx

Mtro. José Antonio Contreras Benítez
acontreras@dgp.unam.mx

Ing. Carlos Alberto Román Zamitiz
croman@dgp.unam.mx

Fuente: Webinar "Transformación Digital" por Lucas Adolfo Giraldo Ríos de la empresa RSM de Colombia

¡Gracias por su atención!



¿Preguntas?