



***Proyecto de orden de la Consejería de Educación y Juventud, por la que se establece el programa de especialización de formación profesional en Big Data en la Comunidad de Madrid y se define su plan de estudios.***

El artículo 24.1 del Decreto 63/2019, 16 de julio, del Consejo de Gobierno, por el que se regula la ordenación y organización de la formación profesional en la Comunidad de Madrid dispone que con el fin de atender las demandas formativas de especialización que trasladen los sectores productivos o las correspondientes asociaciones y organizaciones empresariales y entidades laborales, la Comunidad de Madrid podrá establecer, a través de programas de especialización, una formación complementaria de carácter especializado vinculada a los ciclos formativos que estén autorizados en el centro educativo. Asimismo, se determina que serán destinatarios de esta formación los titulados de formación profesional que se determinen en cada programa.

En 2018, Invest In Madrid ha realizado un estudio exhaustivo relativo al desarrollo del sector de las Tecnologías de la Información y la Comunicación, en adelante TIC, del informe resultado de dicho estudio se desprende que, en España, el subsector de Big Data muestra una tasa de crecimiento del 30% cada año, siete veces más que la inversión tecnológica tradicional lo que ha convertido a nuestro país en la cuarta región de la Unión Europea con más empleos tecnológicos. Además, más de 10.000 personas están trabajando en esta área. La Comunidad de Madrid cuenta con más de 23 centros de datos en todo su territorio y es el sexto centro más activo en desarrollo de proyectos basados en *blockchain* (cadena de bloques), con un 43% del empleo tecnológico español concentrado en nuestra comunidad autónoma. Estos datos propician establecer el programa de especialización de formación profesional en Big Data para atender las demandas formativas en nuestra región, de esta forma se colabora con el crecimiento de este sector.

De conformidad con el artículo 24.4 del Decreto 63/2019, de 16 de julio, corresponde al titular de la consejería con competencia en materia de educación determinar por orden la oferta de cada programa de especialización con la concreción de, al menos, su denominación, duración, currículo y ciclos formativos de formación profesional vinculados a dicho programa. Asimismo, de acuerdo con el apartado 2 del citado artículo, estos programas tendrán una estructura modular, de tal forma que estos módulos de formación complementaria, presentarán un claro componente de especialización con el sector productivo con el que se vincule y tendrán un carácter eminentemente práctico.

En el marco de lo dispuesto en el artículo 129 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas, la presente orden se ajusta a las exigencias del principio de necesidad y eficacia puesto que regula las condiciones para establecer la oferta del programa de especialización de formación profesional en Big Data que permita atender las demandas del sector productivo y faciliten hacer efectivas estas enseñanzas para que puedan ser impartidas en el ámbito de la Comunidad de Madrid, con el fin de mejorar la cualificación de los egresados de formación profesional con formación en TIC. La norma no se extralimita en sus disposiciones respecto a lo establecido en el Decreto 63/2019, de 16 de julio, atiende a la necesidad de ampliar la oferta de acuerdo con las demandas de los sectores productivos. Por otro lado, el cumplimiento de estos principios contribuye, además, a lograr un ordenamiento autonómico sólido y coherente en materia curricular que garantiza el principio de seguridad jurídica, además de cumplir con el principio



de transparencia, eficiencia y de estabilidad presupuestaria y sostenibilidad financiera, tanto por lo exhaustivo y transparente de su tramitación como por su publicación en el Boletín Oficial de la Comunidad de Madrid.

En el proceso de elaboración de esta orden se ha dado cumplimiento a los trámites de audiencia e información pública a través del Portal de Transparencia de la Comunidad de Madrid, conforme a lo dispuesto en el artículo 133.2 de la Ley 39/2015, de 1 de octubre, y el artículo 26.6 de la Ley 50/1997, de 27 de noviembre, del Gobierno, respetando así el principio de transparencia normativa.

El objeto de la presente orden es establecer en la Comunidad de Madrid el programa de especialización de formación profesional en Big Data con la regulación de todos los aspectos que se requieren para definir y concretar su organización, estructura y desarrollo curricular, así como las condiciones para que se facilite la oferta educativa correspondiente, en relación al acceso y la admisión del alumnado en estas enseñanzas y los requisitos para la autorización de centros educativos.

Por todo ello, en el ejercicio de las competencias que le atribuye el artículo 41.d), de la Ley 1/1983, de 13 de diciembre, de Gobierno y Administración de la Comunidad de Madrid, y de conformidad con lo establecido en el Decreto 288/2019, de 12 de noviembre, del Consejo de Gobierno, por el que se establece la estructura orgánica de la Consejería de Educación y Juventud,

## **DISPONGO**

### **CAPÍTULO I**

#### **Aspectos generales**

##### **Artículo 1.** *Objeto y ámbito de aplicación.*

1. La presente orden establece en la Comunidad de Madrid el programa de especialización en Big Data y regula su organización, estructura y desarrollo curricular, así como las condiciones para que se facilite la oferta educativa correspondiente, en relación al acceso y la admisión del alumnado en estas enseñanzas y los requisitos para la autorización de centros educativos.
2. Esta norma será de aplicación en los centros públicos y privados de la Comunidad de Madrid que, debidamente autorizados, impartan estas enseñanzas.

##### **Artículo 2.** *Finalidad.*

El programa de especialización en Big Data tiene la finalidad de atender las demandas de profesionales cualificados en el sector de las Tecnologías de la Información y la Comunicación en la Comunidad de Madrid, de tal forma que se articule una respuesta formativa adecuada y flexible que permita a los egresados de formación profesional en el campo de la informática especializarse en este ámbito.



## CAPÍTULO II

### **Identificación del programa de especialización, destinatarios, acceso y admisión al programa, perfil y entorno profesional**

#### **Artículo 3.** *Identificación del programa de especialización.*

El programa de especialización en Big Data queda identificado en el ámbito de la Comunidad de Madrid por los siguientes elementos:

- a) Denominación: Big Data
- b) Código: INFP01
- c) Familia profesional: Informática y Comunicaciones
- d) Ciclos formativos de grado superior regulados al amparo de la Ley Orgánica 2/2006, de 3 de mayo, de Educación, vinculados:
  - 1º Administración de Sistemas Informáticos en Red.
  - 2º Desarrollo de Aplicaciones Multiplataforma.
  - 3º Desarrollo de Aplicaciones Web.
  - 4º Sistemas de Telecomunicaciones e Informáticos.
  - 5º Automatización y robótica industrial.
  - 6º Mecatrónica Industrial.
- e) Duración: 360 horas

#### **Artículo 4.** *Requisitos de acceso.*

Para poder cursar el programa de especialización en Big Data será requisito estar en posesión del título de Técnico Superior correspondiente con alguno de los ciclos formativos de grado superior vinculados a esta formación que se relacionan en el artículo 3.d) o títulos de formación profesional equivalentes.

#### **Artículo 5.** *Admisión y matrícula en el programa de especialización en centros sostenidos con fondos públicos.*

1. La Dirección General con competencia en materia de formación profesional del sistema educativo establecerá con carácter anual el procedimiento de admisión a estas enseñanzas así como, en el caso de ser admitido para cursar estas enseñanzas, el procedimiento para formalizar la matrícula en los centros sostenidos con fondos públicos que, debidamente autorizados, las impartan.
2. De conformidad con los criterios de admisión a los programas de especialización establecidos en el artículo 36.2.d) del Decreto 63/2019, de 16 de julio, del Consejo de Gobierno, por el que se regula la ordenación y organización de la formación profesional en la Comunidad de Madrid, los criterios prioritarios para establecer el baremo en los procedimientos de admisión contemplarán la calificación final obtenida en alguno de los ciclos formativos requeridos para acceder a estas enseñanzas y el año en que obtuvo la titulación que se aporte como requisito de acceso. Asimismo, tendrán prioridad quienes



hayan cursado un ciclo formativo de grado superior perteneciente a la familia profesional de Informática y Comunicaciones.

3. La matrícula para estas enseñanzas en centros públicos debidamente autorizados requerirá el abono de los precios públicos correspondientes.
4. Los procedimientos, términos y condiciones relativos a la cancelación y anulación de la matrícula serán los establecidos con carácter general y que resulten de aplicación al alumnado que cursa ciclos formativos de formación profesional.

#### **Artículo 6. Perfil profesional.**

El perfil profesional del programa de especialización en Big Data queda determinado por su competencia general y sus competencias profesionales, personales y sociales.

#### **Artículo 7. Competencia general.**

La competencia general de este programa de especialización consiste en conocer y manejar aplicaciones específicas para el tratamiento, integración, gestión y análisis de macrodatos estructurados y no estructurados, así como los principios de funcionamiento del Big Data para abordar actividades empresariales relacionadas con el desarrollo de productos, el mantenimiento predictivo, la experiencia del cliente, la lucha contra el fraude y la eficiencia operativa.

#### **Artículo 8. Competencias profesionales, personales y sociales.**

Las competencias profesionales, personales y sociales de este programa de especialización son las siguientes:

- a) Conocer las principales técnicas para la integración y depuración de conjuntos de Big Data obtenidos de distintas fuentes y aplicaciones.
- b) Conocer y utilizar adecuadamente los motores de procesamiento de datos de uso más frecuente.
- c) Elaborar documentos e informes estadísticos como resultado del tratamiento de Big Data.
- d) Seleccionar y establecer los filtros y consultas más idóneos para extraer información relevante a partir de un conjunto de macrodatos.
- e) Resolver situaciones, problemas o contingencias con iniciativa y autonomía en el ámbito de su competencia, con creatividad, innovación y espíritu de mejora en el trabajo personal y en el de los miembros de un equipo de trabajo.
- f) Organizar y coordinar equipos de trabajo con responsabilidad, supervisando el desarrollo del mismo, manteniendo relaciones fluidas y asumiendo el liderazgo, así como aportando soluciones a los conflictos grupales que se presenten.
- g) Comunicarse con sus iguales, superiores, clientela y personas bajo su responsabilidad, utilizando vías eficaces de comunicación, transmitiendo la información o conocimientos adecuados y respetando la autonomía y competencia de las personas que intervienen en el ámbito de su trabajo.



- h) Generar entornos seguros en el desarrollo de su trabajo y el de su equipo, supervisando y aplicando los procedimientos de prevención de riesgos laborales y ambientales, de acuerdo con lo establecido por la normativa y los objetivos de la empresa.
- i) Supervisar y aplicar procedimientos de gestión de calidad, de accesibilidad universal y de «diseño para todas las personas», en las actividades profesionales incluidas en los procesos de producción o prestación de servicios.
- j) Ejercer sus derechos y cumplir con las obligaciones derivadas de su actividad profesional, de acuerdo con lo establecido en la legislación vigente, participando activamente en la vida económica, social y cultural.

**Artículo 9. Entorno profesional.**

Las personas que hayan obtenido el certificado que acredita la superación de este programa de especialización podrán ejercer su actividad en el sector de las Tecnologías de la Información y la Comunicación y el análisis de datos a gran escala como operadores y asesores en el tratamiento de macrodatos.

### CAPÍTULO III

#### **Estructura y organización curricular. Documentación académica.**

**Artículo 10. Objetivos generales.**

Los objetivos generales de este programa de especialización son los siguientes:

- a) Valorar y comprender la importancia estratégica del análisis de macrodatos en la toma de decisiones empresariales.
- b) Obtener una visión global del impacto de los proyectos basados en Big Data en el desarrollo y gestión empresarial.
- c) Instalar, configurar y administrar sistemas Big Data.
- d) Conocer y utilizar de forma adecuada herramientas de procesamiento de datos en infraestructuras Big Data.
- e) Conocer las bases de los modelos, arquitecturas y lenguajes de programación habituales en Big Data.
- f) Conocer y manejar diferentes opciones de almacenamiento masivo de datos seleccionando la herramienta más adecuada en cada caso.
- g) Elaborar conclusiones e informes del tratamiento de datos masivos estructurados y no estructurados, una vez integrados y depurados adecuadamente.

**Artículo 11. Estructura del programa de especialización.**

1. De conformidad con el artículo 24.2 del Decreto 63/2019, de 16 de julio, el programa de especialización en Big Data se ordena con una estructura modular, de tal forma que los módulos de esta formación complementaria, presentarán un claro componente de



especialización con el sector productivo con el que se vincule y tendrán un carácter eminentemente práctico.

2. Los módulos profesionales que constituyen el currículo del programa de especialización en Big Data son los siguientes:
  - a) CM01-PEBD “Programación en entornos Big Data”
  - b) CM02-PEBD “Frameworks para Big Data”
  - c) CM03-PEBD “Procesamiento de datos”
  - d) CM04-PEBD “Bases de datos no relacionales”

### **Artículo 12.** *Currículo.*

1. El currículo del programa de especialización se sustenta en el perfil profesional y los objetivos generales regulados en la presente orden.
2. Los objetivos expresados en términos de resultados de aprendizaje, los criterios de evaluación, la duración, los contenidos y las orientaciones pedagógicas de los módulos profesionales recogidos en el artículo 11.2, son los que se especifican en el anexo I.
3. Los centros educativos concretarán y desarrollarán el currículo de este programa de especialización mediante programaciones didácticas, en el contexto del proyecto educativo del centro de acuerdo con lo establecido para las programaciones didácticas de los ciclos formativos de formación profesional en el artículo 32 del Decreto 63/2019, de 16 de julio.
4. Las programaciones didácticas se establecerán teniendo en cuenta las características socioeconómicas del sector y potenciarán la cultura de calidad, la excelencia en el trabajo, así como la formación en materia de prevención de riesgos laborales y de respeto medioambiental, atendiendo a la normativa específica del sector productivo o de servicios correspondiente.
5. Tanto en los procesos de enseñanza y de aprendizaje como en la realización de las actividades que desarrollen las programaciones didácticas se integrará el principio de igualdad de oportunidades entre mujeres y hombres y la prevención de la violencia de género, el respeto y la no discriminación por motivos de orientación sexual y diversidad sexual e identidad o expresión de género.
6. Los centros desarrollarán el currículo establecido en esta orden integrando el principio de “diseño universal o diseño para todas las personas”. En las programaciones didácticas se tendrán en consideración las características del alumnado, prestándose especial atención a las necesidades de quienes presenten una discapacidad reconocida, posibilitando que desarrollen las competencias incluidas en el currículo, así como el adecuado acceso al mismo.

### **Artículo 13.** *Organización y distribución horaria.*

1. Las actividades formativas que formen parte del programa de especialización en Big Data se impartirán dentro del calendario escolar establecido para cada curso académico.
2. Los centros docentes, con carácter general, organizarán el desarrollo de las actividades formativas con una duración de un curso académico y podrán establecer un calendario de



evaluaciones parciales análogo al correspondiente al primer curso de los ciclos formativos de formación profesional, que en todo caso quedará reflejado en la programación didáctica de estas enseñanzas. En este caso se seguirá el cuadro de distribución horaria que se recoge en el anexo II.

3. Asimismo, dentro del ámbito de su autonomía, los centros podrán organizar la impartición del programa de especialización de forma intensiva a lo largo del primer o segundo cuatrimestre del curso académico, pudiendo ofertar el programa en uno o ambos períodos. En este caso se seguirá el cuadro de distribución horaria que se recoge en el anexo III.

#### **Artículo 14. Enseñanza semipresencial.**

1. Los centros podrán organizar estas enseñanzas en modalidad semipresencial, de tal forma que su organización y distribución horaria se recogerá en la programación didáctica y atenderá los siguiente:
  - a) En el módulo profesional CM01-PEBD “Programación en entornos Big Data” se programarán, al menos, 54 horas con actividades presenciales, en las 36 horas restantes podrán programarse actividades a distancia.
  - b) El módulo profesional CM02-PEBD “Frameworks para Big Data” será impartido de forma presencial en su totalidad.
  - c) En el módulo profesional CM03-PEBD “Procesamiento de datos” se programarán, al menos, 72 horas con actividades presenciales, en las 18 horas restantes podrán programarse actividades a distancia.
  - d) En el módulo profesional CM04-PEBD “Bases de datos no relacionales” se programarán, al menos, 54 horas con actividades presenciales, en las 36 horas restantes podrán programarse actividades a distancia.
2. La asistencia a las actividades presenciales será obligatoria. Las actividades que se programen a distancia contarán con el soporte de una plataforma virtual y su seguimiento se llevará a cabo mediante, al menos, una tutoría lectiva semanal por cada módulo profesional que deberá impartirse durante el período que duren las actividades a distancia.
3. Sin perjuicio de lo recogido en los apartados anteriores las actividades formativas se impartirán dentro del calendario escolar establecido para cada curso académico y se programarán con una duración, con carácter general, de un curso académico o de forma intensiva a lo largo del primer o segundo cuatrimestre del curso académico, pudiendo ofertar el programa en uno o ambos períodos.

#### **Artículo 15. Evaluación y calificación.**

1. La evaluación atenderá lo establecido en el artículo 38 del Decreto 63/2019, de 16 de julio, deberá ser objetiva, atenderá a la diversidad, las pruebas correspondientes a las evaluaciones finales serán siempre presenciales y el profesorado dará a conocer al alumnado las características principales y el procedimiento de la misma.
2. Los referentes de la evaluación serán los objetivos que expresados en resultados de aprendizaje se recogen para cada módulo profesionales en la presente orden.



3. Cada módulo profesional del programa de especialización será objeto de evaluación un máximo de cuatro convocatorias, en el caso de alumnado con necesidades específicas de apoyo educativo el número máximo de convocatorias será de seis.
4. En el programa de especialización será de aplicación el procedimiento establecido para la renuncia de convocatorias en los ciclos formativos de formación profesional en la normativa vigente.
5. Cada módulo profesional del programa de especialización será evaluado en una única convocatoria de evaluación final al término de las actividades formativas.
6. La calificación de cada módulo profesional será numérica entre uno y diez puntos, sin decimales, dicha calificación se tomará conforme a los criterios establecidos en la programación didáctica del programa de especialización y la programación del módulo profesional correspondiente, se considerará superado aquel módulo profesional que obtenga una calificación igual o superior a cinco puntos.
7. Para superar el programa de especialización se requiere la superación de todos los módulos profesionales que incorpore. El programa de especialización superado contará con una calificación final obtenida como resultado de la media aritmética de todos los módulos profesionales que incluya.
8. El alumnado que haya superado alguno o algunos de los módulos profesionales incluidos en el programa de especialización sin superar el mismo podrá repetir estas enseñanzas sin necesidad de cursar los módulos profesionales superados, de los cuales trasladará la calificación obtenida.
9. Los módulos profesionales incluidos en el programa de especialización en Big Data no podrán ser objeto de convalidación ni exención.

#### **Artículo 16.** *Documentación académica.*

1. Los documentos de evaluación del programa de especialización en Big Data serán las actas de evaluación final, el expediente académico del alumno y la certificación académica de estudios completos o incompletos.
2. La Dirección General con competencia en materia de formación profesional establecerá los modelos de los documentos de evaluación a los que se refiere el apartado anterior.
3. Las actas de evaluación recogerán los resultados obtenidos en cada uno de los módulos profesionales que componen dicho programa del alumnado matriculado en el grupo de referencia y, en su caso, la decisión de superación del programa de especialización. El acta irá firmada por el profesorado que haya impartido docencia en los diferentes módulos profesionales, con el sello del centro y el visto bueno del director del centro.
4. El expediente académico del alumno incluirá el conjunto de calificaciones finales obtenidas en cada módulo profesional, así como cuantas incidencias hayan podido concurrir durante el período en el que se encuentre cursando el programa de especialización. Junto con los datos personales del alumno figurarán los datos de identificación del centro, sus antecedentes académicos con indicación expresa del requisito de acceso y la información relativa, en su caso, a los cambios de centro educativo o domicilio del alumno.



5. La certificación académica se expedirá en soporte digital, siempre que figure el código de verificación electrónica (CVE) que permita comprobar su autenticidad. Esta certificación acreditará los estudios cursados ya sean completos o parciales, reflejarán los resultados obtenidos por el alumno y tendrán validez en el ámbito de la Comunidad de Madrid.

#### **Artículo 17.** Profesorado.

1. Las especialidades y, en su caso, las titulaciones del profesorado con atribución docente en los módulos profesionales incluidos en el artículo 11.2 son las que se determinan en el anexo IV.
2. Además de estas titulaciones requeridas, con las que el profesorado tendrá que acreditar la formación inicial y cualificación específica que garantice la capacitación adecuada para impartir el currículo de los módulos profesionales, se deberá acreditar la formación pedagógica y didáctica necesaria para ejercer la docencia, según se establece en el artículo 100 de la Ley Orgánica 2/2006, de 3 de mayo, de Educación.

#### **Artículo 18.** Espacios y equipamientos.

1. Los espacios necesarios para el desarrollo de las enseñanzas de este ciclo formativo son los establecidos en el Anexo V.
2. Los espacios formativos establecidos podrán ser ocupados por diferentes grupos de alumnos que cursen el mismo programa de especialización u otras enseñanzas de formación profesional.
3. Los equipamientos que se incluyen en cada espacio han de ser los necesarios y suficientes para garantizar al alumnado la adquisición de los resultados de aprendizaje y la calidad de la enseñanza.
4. Además, los espacios y equipamientos deberán cumplir la normativa sobre diseño para todos y accesibilidad universal, así como sobre prevención de riesgos laborales y seguridad y salud en el trabajo.

#### **Artículo 19.** Oferta y autorización.

1. De conformidad con el artículo 27 del Decreto 63/2019, de 16 de julio, el programa de especialización en Big Data podrá impartirse en los centros docentes, públicos y privados, que sean debidamente autorizados por la consejería competente en materia de educación, según el procedimiento establecido con carácter general para la autorización de oferta de enseñanzas de formación profesional.
2. Los centros docentes, para poder ser autorizados a ofertar el programa de especialización en Big Data, deberán contar previamente con autorización para impartir algunos de los ciclos formativos de grado superior de la familia profesional de Informática y Comunicaciones, así como disponer de los recursos humanos y materiales dispuestos en esta orden.
3. Los centros podrán matricular hasta 30 alumnos en cada unidad que tengan autorizada.

**DISPOSICIÓN FINAL PRIMERA.** *Aplicabilidad de otras normas.*



**Comunidad  
de Madrid**

En aquellas cuestiones que no estén reguladas en la presente orden se aplicarán de forma supletoria las normas de ordenación y organización de la formación profesional de la Comunidad de Madrid.

**DISPOSICIÓN FINAL SEGUNDA.** *Habilitación para la aplicación y ejecución.*

Se autoriza al titular de la Dirección General con competencias en materia de formación profesional a dictar las resoluciones e instrucciones que se consideren necesarias para la aplicación de lo dispuesto en la presente orden.

**DISPOSICIÓN FINAL TERCERA.** *Entrada en vigor.*

La presente orden entrará en vigor el día siguiente al de su publicación en el “Boletín Oficial de la Comunidad de Madrid”.

EL CONSEJERO DE EDUCACIÓN Y JUVENTUD

Fdo.: Enrique Ossorio Crespo



## ANEXO I

Elementos curriculares del plan de estudios del programa de especialización

### 01 Módulo Profesional: Programación en entornos Big Data.

**Código: CM01-PEBD.**

**Duración: 90 horas.**

<i>Resultados de aprendizaje</i>	<i>Criterios de evaluación</i>
Conoce y comprende las características de los lenguajes de programación de uso más frecuente en proyectos Big Data.	<ul style="list-style-type: none"> <li>• Se han utilizado los lenguajes de programación Big Data más extendidos.</li> <li>• Se han instalado, configurado y utilizado entornos de trabajo en local y en remoto.</li> <li>• Se han reconocido las principales diferencias con los lenguajes de programación tradicionales.</li> <li>• Se han identificado los paradigmas de programación habituales en programación científica y Big Data.</li> </ul>
Desarrolla aplicaciones aplicando características avanzadas de los lenguajes de programación Big Data	<ul style="list-style-type: none"> <li>• Se han manejado las estructuras de datos más habituales en programación científica.</li> <li>• Se han instalado y utilizado las librerías más habituales.</li> <li>• Se han entendido las diferentes partes de un proyecto Big Data.</li> <li>• Se han reconocido las posibilidades de entrada/salida de los diferentes lenguajes.</li> </ul>
Analiza los datos y realiza gráficos y visualizaciones relevantes.	<ul style="list-style-type: none"> <li>• Se ha realizado el análisis estadístico de los datos.</li> <li>• Se han realizado diferentes tipos de visualizaciones.</li> <li>• Se han seleccionado los tipos de visualización más apropiados según los datos que se quieran mostrar.</li> <li>• Se han utilizado programas y librerías específicos para la visualización de datos.</li> </ul>

### Contenidos

#### 1. Lenguajes para Big Data:

- Características de los lenguajes de programación para Big Data.
- Tipos de datos.
  - Por categorías: estructurados, no estructurados y multi-estructurados o híbridos.
  - Por origen: web y redes sociales, comunicación entre máquinas, transacciones, biométricos y generados por personas.
- Estructuras de control: de selección, iterativas. Funciones. Recursividad.
- Librerías específicas.

#### 2. Entrada y salida de datos:

- Opciones de entrada y salida habituales.
- Formatos habituales en Big Data.
- Funciones para trabajar con diferentes formatos.

#### 3. Entornos de desarrollo:

- Entornos de desarrollo para programación Big Data.
- Instalación y configuración de entornos de desarrollo local.
- Entorno de desarrollo para trabajo en remoto.



#### 4. Análisis de datos:

- Análisis descriptivo. Utilización de un histórico de datos. Identificación de comportamientos.
- Análisis predictivo. Creación de modelos. Establecimiento de un patrón e identificación de las principales tendencias.
- Análisis prescriptivo. Técnicas de simulación y optimización.

#### 5. Visualización de datos:

- Librerías y herramientas para visualización de datos.
- Tipos de gráficas y visualizaciones.
- Gráficos y mapas interactivos. Gráficos *responsive*.

#### *Orientaciones pedagógicas.*

Este módulo profesional contiene la formación necesaria para el desempeño de actividades relacionadas con el entorno de trabajo en el que el alumnado va a ejercer su profesión.

Las actividades centrarán la atención en el ámbito profesional relacionado con los lenguajes y librerías que permiten el tratamiento de macrodatos, así como la obtención de gráficos y visualizaciones relevantes para el análisis.

Las líneas de actuación en el proceso de enseñanza y de aprendizaje que permiten alcanzar los objetivos del módulo, versarán sobre:

- Lenguajes más utilizados en Big Data como *Python*, *R* o *Scala*.
- Librerías específicas para análisis de datos como *numpy*, *pandas*, *scikit*.
- Librerías y herramientas para la visualización de datos.
- Instalación de entornos de desarrollo; locales y remotos.
- Integración en un entorno Big Data.

Las estrategias metodológicas se basarán en fundamentalmente en el desarrollo de actividades prácticas que simulen situaciones reales de trabajo en las que se facilite la aplicación y puesta en práctica de los contenidos.

## 02 Módulo Profesional: *Frameworks* para Big Data.

**Código: CM02-PEBD.**

**Duración: 90 horas.**

<i>Resultados de aprendizaje</i>	<i>Criterios de evaluación</i>
Identifica los conceptos básicos de arquitectura de sistemas para la adquisición y procesamiento distribuido de datos.	<ul style="list-style-type: none"> <li>• Se han estudiado diferentes métodos de almacenamiento y computación de datos.</li> <li>• Se han identificado los componentes básicos de un sistema de procesamiento distribuido.</li> <li>• Se han identificado herramientas adiciones para la gestión y procesamiento de información.</li> </ul>
Diseña, instala y administra arquitecturas y tecnologías para el manejo de grandes volúmenes de datos.	<ul style="list-style-type: none"> <li>• Se ha instalado una infraestructura de almacenamiento distribuido.</li> <li>• Se han instalado las aplicaciones básicas de un entorno de procesamiento distribuido.</li> <li>• Se han tenido en cuenta los criterios de capacidad, redundancia y alta disponibilidad de los datos.</li> </ul>
Conoce técnicas de diseño y desarrollo de aplicaciones de procesamiento de grandes volúmenes de información.	<ul style="list-style-type: none"> <li>• Se han identificado las herramientas para interacción con bases de datos relacionales y <i>NoSQL</i></li> <li>• Se han configurado herramientas para la adquisición y procesamiento de flujos de datos.</li> </ul>



	<ul style="list-style-type: none"><li>• Se han utilizado herramientas para consulta de datos.</li></ul>
Monitoriza y analiza el rendimiento de un sistema Big Data, identificando y resolviendo incidencias de funcionamiento.	<ul style="list-style-type: none"><li>• Se ha configurado una herramienta de coordinación para las aplicaciones básicas de un entorno de procesamiento distribuido.</li><li>• Se ha realizado una ingesta de datos sobre el sistema de archivos distribuido.</li><li>• Se ha realizado el procesamiento de datos mediante la metodología Map/Reduce.</li><li>• Se ha analizado la capacidad del sistema de procesamiento y el rendimiento del mismo.</li><li>• Se han tenido en cuenta los criterios de seguridad de los datos.</li></ul>

### Contenidos

#### 1. Infraestructuras para Big Data:

- El paradigma Big Data. Modelo 5V: velocidad, volumen, variedad de los datos, veracidad y valor.
- Sistemas de computación distribuida: *clusters*, *cloud*.
- Componentes básicos y requisitos para la arquitectura de un sistema Big Data.

#### 2. Componentes de la arquitectura Big Data:

- Infraestructura en almacenamiento básica para la ingesta de datos.
- Infraestructuras para el procesamiento de datos.
- Herramientas adicionales.

#### 3. Almacenamiento distribuido:

- Características de los sistemas de ficheros distribuidos.
  - Almacenamiento de información permanente.
  - Sistemas multiusuario.
  - Transparencia en la identificación y en la ubicación.
  - Escalabilidad.
  - Consistencia y seguridad.
- Capacidad, redundancia y alta disponibilidad. Tolerancia a fallos.
- Despliegue de nodos de almacenamiento,
- Administración y acceso a los datos.
- Monitorización de sistemas de ficheros distribuidos. Herramientas para monitorizar aplicaciones y servidores.

#### 4. Administración de entornos Big Data:

- Paradigma de computación distribuida y en paralelo, proceso cercano a los datos.
- Aplicaciones y técnicas de procesamiento distribuido.
- Ingesta de datos.
- Metodología *MapReduce*.
- Rendimiento y monitorización de los procesos.
- Administración de seguridad de los datos.

### Orientaciones pedagógicas.

Este módulo profesional contiene la formación básica que debe poseer un administrador de sistemas Big Data para realizar la correcta configuración de una infraestructura de almacenamiento y procesamiento distribuido de datos.



El administrador debe garantizar que el sistema funciona en condiciones adecuadas de capacidad, rendimiento y seguridad de los datos.

Se buscará que el alumnado sea capaz de configurar el sistema operativo realizando la parametrización adecuada a las especificaciones técnicas de rendimiento y seguridad que permitan la instalación de un entorno de aplicaciones Big Data así como que disponga de los conocimientos adecuados para configurar el subsistema de almacenamiento y procesamiento distribuido en clúster, garantizando la seguridad y redundancia de los datos.

Esta responsabilidad incluye la configuración de parámetros básicos del sistema operativo anfitrión, así como del subsistema de procesamiento y otras herramientas adicionales que se instalan sobre él, como por ejemplo, en el ecosistema Cloudera Hadoop:

Las líneas de actuación en el proceso de enseñanza y de aprendizaje que permiten alcanzar los objetivos del módulo, versarán sobre:

- Bases de datos *NoSQL* y herramientas como *Cassandra* y *HBase*.
- Herramientas para conectar flujos de datos estructurados como *Flume* y *Sqoop*.
- Plataformas de almacenamiento, procesamiento y consulta de datos como *Pig*, *Hive*, *Impala* y *Oozie*.
- Herramientas para la coordinación de procesos de un clúster como *Zookeeper*.

Las estrategias metodológicas se basarán en fundamentalmente en el desarrollo de actividades prácticas que simulen situaciones reales de trabajo en las que se facilite la aplicación y puesta en práctica de los contenidos.

### 03 Módulo Profesional: Procesamiento de datos.

**Código: CM03-PEBD.**

**Duración: 90 horas.**

<i>Resultados de aprendizaje</i>	<i>Criterios de evaluación</i>
Desarrolla aplicaciones para la extracción de información.	<ul style="list-style-type: none"> <li>• Se ha implementado una aplicación para extraer información de varias fuentes.</li> <li>• Se han implementado aplicaciones para extraer datos en distintos formatos.</li> <li>• Se han actualizado datos de una anterior ingesta.</li> </ul>
Escribe programas que manipulan y almacenan información utilizando tipos de datos avanzados.	<ul style="list-style-type: none"> <li>• Se han escrito programas que transforman y limpian los datos obtenidos de una ingesta.</li> <li>• Se han escrito programas que almacenan en uno o varios destinos los datos de una ingesta.</li> <li>• Se han escrito programas que obtienen datos de calidad.</li> <li>• Se han escrito programas que respetan la normativa de seguridad en cuanto a protección de datos sensibles.</li> </ul>
Escribe programas de procesamiento de datos.	<ul style="list-style-type: none"> <li>• Se han identificado y clasificado los distintos tipos de procesamiento en Big Data.</li> <li>• Se han escrito programas de procesamiento de datos por lotes.</li> <li>• Se han escrito programas de procesamiento de datos en <i>streaming</i>.</li> </ul>

#### Contenidos

##### 1. Ingesta de datos:

- Tecnologías de ingesta de datos.



- Origen y formato de los datos. *Opendata*.
- Latencia y disponibilidad de los datos.
  - Causas de la latencia.
  - Funciones que dependen de la latencia.
  - Formas de mitigar la latencia.
- Actualizaciones de los datos. Histórico de cambios.
- Extracción de datos.

**2. Preprocesado de datos:**

- Transformaciones de los datos.
- Destino de los datos. Consumidores de datos.
- Calidad de los datos. Formato, origen, veracidad y valor.
- Seguridad de los datos.

**3. Procesamiento de datos:**

- Introducción a los sistemas de procesamiento Big Data.
- Procesamiento por lotes. Aplicaciones y características.
- Procesamiento en *streaming*. Aplicaciones y características.

*Orientaciones pedagógicas.*

Este módulo profesional contiene la formación necesaria para poder procesar datos en un entorno Big Data.

Las actividades centrarán la atención en los componentes de software utilizados para la ingesta de datos y posterior extracción y procesamiento de los mismos.

Las líneas de actuación en el proceso de enseñanza y de aprendizaje que permiten alcanzar los objetivos del módulo, versarán sobre:

- La ingesta de datos de distintos tipos de fuentes y con distintos formatos.
- La preparación de dato, en la que se realiza un filtrado de la información, la transformación de los datos obtenidos en bruto y el envío a su destino. Todo este proceso incluye respetar las normativas vigentes en cuanto a protección de datos.
- El procesado de datos por lotes y por *streaming*. Análisis de la información obtenida y su visualización mediante el empleo de herramientas.

Las estrategias metodológicas se basarán en actividades prácticas en las que se utilicen las herramientas de procesamiento de datos por lotes y en *streaming* en infraestructuras Big Data y se apliquen las principales técnicas de preprocesado de datos.

**04 Módulo Profesional: Bases de datos no relacionales.**

**Código: CM04-PEBD.**

**Duración: 90 horas.**

<i>Resultados de aprendizaje</i>	<i>Criterios de evaluación</i>
Evalúa las diferentes arquitecturas disponibles para el almacenamiento de datos.	<ul style="list-style-type: none"> <li>• Se han reconocido las características de las diferentes arquitecturas para el almacenamiento de datos.</li> <li>• Se han interpretado diferentes situaciones reales de implantación de las distintas arquitecturas para el almacenamiento de datos.</li> </ul>



	<ul style="list-style-type: none"><li>• Se han identificado ejemplos para el desarrollo de las diferentes opciones propias de cada arquitectura.</li></ul>
Desarrolla tipos de casos de uso para cada modelo de arquitectura.	<ul style="list-style-type: none"><li>• Se han identificado las características de cada uno de los casos de uso de cada modelo de arquitectura.</li><li>• Se han reconocido ejemplos aplicados a los diferentes casos de uso para cada arquitectura.</li><li>• Se han resueltos ejemplos reales de aplicabilidad de casos de uso.</li></ul>
Instala y administra bases de datos no relacionales.	<ul style="list-style-type: none"><li>• Se han instalado diferentes tipos de bases de datos no relacionales.</li><li>• Se han identificado las técnicas para la administración de bases de datos no relacionales.</li><li>• Se han administrado distintas clases de bases de datos no relacionales.</li></ul>
Realiza consultas de datos utilizando herramientas propias para su tarea.	<ul style="list-style-type: none"><li>• Se han identificado diferentes herramientas para la consulta de datos.</li><li>• Se han reconocido distintas situaciones que permiten el uso y explotación de los datos.</li><li>• Se ha resuelto la ejecución de determinados ejemplos para la consulta de datos en una base de datos no relacional.</li></ul>

### Contenidos

#### 1. Arquitecturas disponibles para el almacenamiento masivo de datos:

- Características principales, criterios de selección.
- Documental. Catálogos, perfiles de usuario y sistemas de administración de contenido.
- Clave-valor. Escalado horizontal.
- Basado en columnas. Familias de columnas.
- Grafos. Nodos y bordes.
- Series temporales.
- Datos de objetos.
- Datos de índice externo.

#### 2. Casos de uso para cada arquitectura:

- Documental.
- Clave-valor.
- Basado en columnas.
- Grafos.

#### 3. Bases de datos no relacionales:

- Características principales: flexibilidad, escalabilidad, alto rendimiento y funcionabilidad.
- Ejemplos de uso más frecuente.
- Instalación y administración de bases de datos no relacionales. Técnicas y procedimientos.

#### 4. Consulta de datos:

- Principales herramientas. Características y criterios de selección.
- Programación de consultas y búsquedas en función de la arquitectura utilizada.

### Orientaciones pedagógicas.

Este módulo profesional contiene la formación necesaria para el desempeño, instalación y explotación de bases de datos no relacionales.



Las actividades centrarán la atención en el conocimiento, instalación y explotación de bases de datos no relacionales con sus correspondientes particularidades.

Las líneas de actuación en el proceso de enseñanza y de aprendizaje que permiten alcanzar los objetivos del módulo, versarán sobre:

- Arquitecturas disponibles para el almacenamiento de datos.
- Casos de uso planteados para cada arquitectura como: sistemas de flujo de eventos, gestores de contenido y plataformas *Blogging*, analíticas web y datos en tiempo real, aplicaciones *eComerce*, información sobre la sesión de navegación, perfiles de usuario y preferencias, datos del carrito de compra, cachear datos, sistemas de flujo de eventos, gestores de contenido y plataformas *Blogging*, contadores, datos conectados, enrutamiento y entrega o servicios basados en la posición, así como motores de recomendaciones.
- Instalación y administración de bases de datos no relacionales.
- Herramientas para realizar la consulta de datos en los esquemas relacionales.

Las estrategias metodológicas se basarán en actividades prácticas que permitan conocer las diferentes arquitecturas disponibles para almacenamiento masivo diferenciando los casos de uso apropiado para cada tipo de arquitectura, así como la instalación y administración e bases de datos no relacionales y la utilización de herramientas de consulta de datos adecuadas.

## ANEXO II

Cuadro de distribución horaria del programa de especialización impartido a lo largo de un curso académico

Familia profesional: <b>INFORMÁTICA COMUNICACIONES</b>				
Programa de especialización: <b>Big Data</b>			Duración: <b>360 horas</b>	Código: <b>INFP01</b>
MÓDULOS PROFESIONALES				Curso. 3 trimestres (horas semanales)
Clave	Código	Denominación	Duración del currículo (horas)	
01	CM01-PEBD	Programación en entornos Big Data	90	3
02	CM02-PEBD	Frameworks para Big Data	90	3
03	CM03-PEBD	Procesamiento de datos	90	3
04	CM04-PEBD	Bases de datos no relacionales	90	3
<b>HORAS TOTALES</b>			<b>360</b>	<b>12</b>

## ANEXO III

Cuadro de distribución horaria del programa de especialización impartido en un cuatrimestre

Familia profesional: <b>INFORMÁTICA COMUNICACIONES</b>				
Programa de especialización: <b>Big Data</b>			Duración: <b>360 horas</b>	Código: <b>INFP01</b>
MÓDULOS PROFESIONALES				Cuatrimestre (horas semanales)
Clave	Código	Denominación	Duración del currículo (horas)	
01	CM01-PEBD	Programación en entornos Big Data	90	6
02	CM02-PEBD	Frameworks para Big Data	90	6



03	CM03-PEBD	Procesamiento de datos	90	6
04	CM04-PEBD	Bases de datos no relacionales	90	6
<b>HORAS TOTALES</b>			<b>360</b>	<b>24</b>

## ANEXO IV

### Especialidades y titulaciones del profesorado con atribución docente en los módulos profesionales del programa de especialización en Big Data

Módulo profesional	Cuerpo docente y especialidad <sup>(1)</sup>		Titulaciones <sup>(3)</sup>
	Cuerpo <sup>(2)</sup>	Especialidad	
CM01-PEBD "Programación en entornos Big Data"	CS PS PTFP	Informática (107) Sistemas y aplicaciones informáticas (227)	<ul style="list-style-type: none"> <li>Graduado en Ingeniería Informática, Graduado en Matemáticas o títulos de formación universitaria equivalentes.</li> <li>Además se requiere una cualificación específica en la materia que podrá acreditarse mediante formación de postgrado en métodos analíticos para datos masivos o Big Data de, al menos, 24 créditos ECTS o una experiencia profesional en el sector de Big Data de duración equivalente, al menos, a un año a jornada completa.</li> </ul>
	Profesor especialista		
CM02-PEBD "Frameworks para Big Data"	CS PS PTFP	Informática (107) Sistemas y aplicaciones informáticas (227)	<ul style="list-style-type: none"> <li>Graduado en Ingeniería Informática, Graduado en Matemáticas o títulos de formación universitaria equivalentes.</li> <li>Además se requiere una cualificación específica en la materia que podrá acreditarse mediante formación de postgrado en métodos analíticos para datos masivos o Big Data de, al menos, 24 créditos ECTS o una experiencia profesional en el sector de Big Data de duración equivalente, al menos, a un año a jornada completa.</li> </ul>
	Profesor especialista		
CM03-PEBD "Procesamiento de datos"	CS PS PTFP	Informática (107) Sistemas y aplicaciones informáticas (227)	<ul style="list-style-type: none"> <li>Graduado en Ingeniería Informática, Graduado en Matemáticas o títulos de formación universitaria equivalentes.</li> <li>Además se requiere una cualificación específica en la materia que podrá acreditarse mediante formación de postgrado en métodos analíticos para datos masivos o Big Data de, al menos, 24 créditos ECTS o una experiencia profesional en el sector de Big Data de duración equivalente, al menos, a un año a jornada completa.</li> </ul>
	Profesor especialista		
CM04-PEBD "Bases de datos no relacionales"	CS PS PTFP	Informática (107) Sistemas y aplicaciones informáticas (227)	<ul style="list-style-type: none"> <li>Graduado en Ingeniería Informática, Graduado en Matemáticas o títulos de formación universitaria equivalentes.</li> <li>Además se requiere una cualificación específica en la materia que podrá acreditarse mediante formación de postgrado en métodos analíticos para datos masivos o Big Data de, al menos, 24 créditos ECTS o una experiencia profesional en el sector de Big Data de duración equivalente, al menos, a un año a jornada completa.</li> </ul>
	Profesor especialista		

(1) Profesorado de centros públicos.

(2) **CS** = Catedrático de Enseñanza Secundaria    **PS** = Profesor de Enseñanza Secundaria    **PTFP** = Profesor Técnico de Formación Profesional  
**PE** = Profesorado Especialista

(3) Profesorado de centros de titularidad privada o de titularidad pública de otras administraciones distintas de la educativa.

## ANEXO V

### Espacios y equipamientos mínimos

Espacios:

Espacio formativo	Superficie m <sup>2</sup> 30 alumnos <sup>(1)</sup>
-------------------	--



Aula polivalente	60
Aula técnica	60

- (1) En caso de autorizarse unidades escolares con ratios inferiores a 30 alumnos, el aula polivalente deberá tener una superficie de 2 m<sup>2</sup>/alumno, con un mínimo de 40 m<sup>2</sup>.

### Equipamientos mínimos:

Espacio formativo	Equipamiento
Aula polivalente	Mobiliario escolar. Equipos de proyección audiovisual.
Aula técnica	Equipos audiovisuales. Un equipo informático por alumno con, al menos, 32 Gb de RAM, conectado en red y con conexión a Internet. Sistemas gestores de bases de datos. Servidores y clientes. Entornos de desarrollo, compiladores e intérpretes, analizadoras de código fuente y empaquetadores, entre otros. Software para desarrollo de gestión empresarial y gestión de macrodatos, entre otros.