

Introducció al Big Data

Codi: 104748

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503873 Comunicació Interactiva	OB	3	2

La metodologia docent i l'avaluació proposades a la guia poden experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Professor/a de contacte

Nom: Santiago Giraldo Luque

Correu electrònic: Santiago.Giraldo@uab.cat

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: espanyol (spa)

Grup íntegre en anglès: No

Grup íntegre en català: No

Grup íntegre en espanyol: No

Altres indicacions sobre les llengües

La docència és desenvoluparà en castellà. Alguna part es pot fer en català o en anglès.

Prerequisits

L'assignatura no té requisits obligatoris, però es recomana la superació prèvia de les següents assignatures de l' Grau:

Sistemes d'informació

Emmagatzematge i recuperació d'informació

Serveis Web Avançats

Objectius

L'objectiu principal de l'assignatura és introduir els i les estudiants en els conceptes bàsics i en les pràctiques principals del Big Data.

Així mateix, el curs té els següents objectius específics:

1. Desenvolupar un coneixement propedèutic per al desenvolupament posterior d'aplicacions de Business Intelligence: el desenvolupament de solucions de big data per a la intel·ligència de negoci i la seva influència en la presa de decisions.
2. Introduir els conceptes de fonts i tipus de dades (estructura, classificació, integració i qualitat).
3. Promoure l'exploració de les peticions i el treball amb fonts d'Open data.
4. Realitzar les primeres aproximacions a l'anàlisi de bases de dades en entorns d'eines de treball amb fulls de càlcul i altres eines pràctiques.

Competències

- Cercar, seleccionar i jerarquitzar qualsevol tipus de font i document útil per a la elaboració de missatges, treballs acadèmics, exposicions, etc.
- Gestionar el temps de manera adequada i ser capaç de planificar tasques a curt, mitjà i llarg terminis.
- Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia.
- Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
- Que els estudiants sàpiguin aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.
- Reconèixer i planificar la infraestructura tecnològica necessària per a la creació, emmagatzematge, anàlisi i la distribució de productes multimèdia interactius i de l'internet social.

Resultats d'aprenentatge

1. Compartir les experiències en grup com a forma d'aprenentatge per treballar posteriorment en grups multidisciplinaris.
2. Contrastar i verificar la veracitat de les informacions aplicant criteris de valoració.
3. Descriure la infraestructura necessària per a l'emmagatzemament del big data.
4. Diferenciar allò substancial d'allò rellevant en tots els tipus de documents de l'assignatura.
5. Diferenciar les varietats de tipus d'arquitectures existents per treballar amb big data.
6. Explicar la infraestructura necessària per al tractament del big data.
7. Explicar les característiques de la infraestructura necessària per a la recuperació del big data.
8. Planificar i executar projectes acadèmics en l'àmbit del big data.
9. Presentar els treballs de l'assignatura en els terminis previstos i mostrar-ne la planificació individual o grupal aplicada.
10. Solucionar problemes bàsics del big data.

Continguts

Unitat 1. El Big Data: Introducció a l'assignatura. Concepte del Big Data, els seus processos i les seves característiques.

Unitat 2. Fonts, captura i emmagatzematge de dades: Presentació de fonts de dades (principalment fonts obertes), dels processos d'accés i sol·licituds d'informació pública i lleis de transparència. Processos de recerca, descàrrega i emmagatzematge de diferents tipus de dades (formats).

Unitat 3. Processament i anàlisi de dades: Ús d'eines i funcions de neteja i anàlisi de dades per a la presa de decisions.

Unitat 4. Visualització de dades: Presentació d'eines de visualització de dades per a la presentació d'informes orientats a la presa de decisió.

Unitat 5. Cartografia amb dades: Presentació de diferents eines i possibilitats de representació cartogràfica de la informació.

(*) El calendari detallat amb el contingut de les diferents sessions s'exposarà el dia de presentació de l'assignatura. Es penjarà també al Campus Virtual on l'alumnat podrà trobar la descripció detallada dels exercicis i pràctiques, els diversos materials docents i qualsevol informació necessària per a l'adequat seguiment de l'assignatura. En cas de canvi de modalitat docent per raons sanitàries, el professorat informarà dels canvis que es produiran en la programació de l'assignatura i en les metodologies docents.

El contingut d'aquesta assignatura serà sensible als aspectes relacionats amb la perspectiva de gènere.

Metodologia

L'estructura de l'assignatura, en la qual es realitzen diferents activitats pràctiques, busca la interiorització de competències relacionades amb l'ús del Big Data (recerca, extracció, anàlisi i publicació de dades per a la presa de decisions). La seva metodologia és completament pràctica. A través de les activitats de laboratori, de tallers i de la realització de la feina final de curs, s'avalua tant el component teòric de l'assignatura, com l'aplicació pràctica dels continguts estudiats.

L'avaluació continuada de l'assignatura, en la qual es realitzen pràctiques puntuals i contínues de curta durada, permet realitzar un seguiment precís sobre l'aprenentatge i la progressió de l'estudiant. Així mateix, es treballa de forma progressiva sobre l'adquisició dels coneixements que, pas a pas, són involucrats en les pràctiques següents.

L'assignatura d'Introducció al Big Data contempla tres tipus o categories d'activitats formatives avaluable:

Pràctiques de laboratori: treballs individuals o en equip en els quals es realitzen activitats pràctiques amb un entregable puntual amb temps límit. Els estudiants han d'aplicar els coneixements, distribuir el temps i preparar els lliuraments dins de l'aula i en les hores destinades a la pràctica sota la guia de professor.

Seminaris: treballs individuals o en equip en què es realitzen activitats pràctiques més extenses i amb entregables oberts a la creativitat dels estudiants. No hi ha temps limitats a l'aula, però sí dates límits de lliurament. Els estudiants han d'aplicar els coneixements, distribuir el temps i preparar els lliuraments iniciant el seu treball dins de l'aula, però continuant sota la modalitat d'activitats supervisades per l'equip docent.

Treball final de curs: exercici pràctic d'avaluació grupal en el qual els estudiants han de resoldre, durant el desenvolupament del curs, un problema d'aplicació pràctica vinculat al Big Data. Els estudiants han de plantejar el problema i realitzar els quatre processos per brindar una proposta de solució basada en grans quantitats de dades: recerca, extracció, anàlisi i publicació de l'informe de dades que inclogui una proposta de decisió basada en la informació recol·lectada i analitzada.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Pràctiques de laboratori	33	1,32	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Sessions teòriques	15	0,6	1, 2, 4, 5
Tipus: Supervisades			
Seminaris	10	0,4	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Tutories	10	0,4	
Tipus: Autònomes			
Treball autònom: realització de lectures, preparació de treballs i pràctiques i estudi personalcurs	60	2,4	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Avaluació

Les activitats d'avaluació són:

Activitat A: Pràctiques de laboratori, que tenen un pes del 40% sobre la qualificació final

Activitat B: Pràctiques d'aula, que tenen un pes del 30% sobre la qualificació final

Activitat C: Entrega de treballs, que té un pes del 30% sobre la qualificació final

Per poder aprovar l'assignatura, cal treure una nota mínima aprovatòria (5,0) en cadascuna de les activitats.

RECUPERACIÓ: En les tres últimes setmanes del curs els estudiants que no hagin superat l'assignatura podran presentar-se a una prova de síntesi de reavaluació que consistirà en un test teòric i en un exercici pràctic. La condició obligatòria per poder optar a la recuperació de l'assignatura és haver fet, com a mínim, 2/3 del total de les pràctiques del curs (activitats A, B i C) i haver obtingut una nota mitjana igual o superior a 3,5 (i menor que 5) en totes les activitats d'avaluació.

D'acord amb els criteris assenyalats anteriorment, si un estudiant no realitza, al menys, el 66% de les pràctiques de les activitats d'avaluació es considerarà com no avaluable d'aquesta assignatura.

PLAGI:

L'estudiant que realitzi qualsevol irregularitat (còpia, plagi, suplantació d'identitat ...) es qualificarà amb 0 aquest acte d'avaluació. En cas que es produeixin diverses irregularitats, la qualificació final de l'assignatura serà 0.

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Entrega de treballs	30%	6	0,24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Pràctiques d'aula	30%	8	0,32	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Pràctiques de laboratori	40%	8	0,32	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Bibliografia

Alcalde, Ignasi. (2015). *Visualización de la información. De los datos al conocimiento*. Editorial UOC.

Bounegru, Liliana; Gary, Jonathan (Eds.). (2020). *The Data Journalism Handbook II. Towards a Critical Data Practice*. European Journalism Centre and Google News Initiative.

<https://datajournalism.com/read/handbook/two>

Bounegru, Liliana; Chambers, Lucy; Gary, Jonathan (Eds.) (2012). *The Data Journalism Handbook: How Journalists Can Use Data to Improve the News*, O'Reilly Media. <https://datajournalism.com/read/handbook/one>

Bradshaw, Paul. (2017). *Scraping for Journalists. How to grab information from hundreds of sources, put it in data you can interrogate - and still hit deadlines (2nd edition)*. Leanpub

Bradshaw, Paul. (2019). *Finding Stories in Spreadsheets. Recipes for interviewing data - and getting answers*. Leanpub

Bradshaw, Paul., Maseda, Bárbara. (2015). *Periodismo de datos: Un golpe rápido. Cómo entrar, obtener los datos, escabullirse con la noticia... ¡Y asegurarse de que nadie salga herido!* Leanpub.

Cairo, Alberto. (2011). *El arte funcional: Infografía y visualización de información*. Alamut.

Cairo, Alberto. (2016). *The Truthful Art: Data, charts, and maps for communication*. New Riders.

Cairo, Alberto. (2017). ¿Visualización de datos: una imagen puede valer más que mil números, pero no siempre más que mil palabras. *El profesional de la información*, 26(6), 1025-1028.

Carlberg, Conrad. (2011). *Análisis estadístico con Excel*. Anaya.

CARTO (2018). The Top Trends in Data Visualization for 2018. *Medium*.
<https://medium.com/@carto/the-top-trends-in-data-visualization-for-2018-54911e875375>

Charte Ojeda, Francisco (2016). *Excel 2016*. Anaya.

Ferrer-Sapena, Antonia; Sánchez-Pérez, Enrique. (2013). Open data, big data: ¿Hacia dónde nos dirigimos? *Anuario ThinkEPI*, 7, 150-156.

Fuchs, Christian. (2017). "Dallas Smythe Today - The Audience Commodity, the Digital Labour debate, Marxist Political Economy and Critical Theory. Prolegomena to a Digital Labour Theory of Value". En: Fuchs, C., Mosco, V. (Eds.). *Marx and the Political Economy of the Media*. Haymarket Books. pp. 522-599.

Giraldo-Luque Santiago; Fernández-Rovira Cristina (2021) Economy of Attention: Definition and Challenges for the Twenty-First Century. En: Park S.H., Gonzalez-Perez M.A., Floriani D.E. (Eds.). *The Palgrave Handbook of Corporate Sustainability in the Digital Era*. Palgrave Macmillan. pp. 283-305.

Greene, Derek. (2014). Practical Social Network Analysis With Gephi. [Practical Social Network Analysis With Gephi · Derek Greene](#)

Greene, Derek; Cunningham, Pàdraig (2013). Producing a Unified Graph Representation from Multiple Social Network Views. *Proc. ACM Web Science'13*

Jauset, J.A. (2007). *Estadística para periodistas, publicitarios y comunicadores* (1 ed.). Catalunya: UOC.

Kauffmann, Erick; Peral, Jesús; Gil, David; Ferrández, Antonio; Sellers, Ricardo; Mora Higinio (2020). A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making, *Industrial Marketing Management*, 90, 523-537.

Mayer-Schönberger, Viktor; Cukier, Kenneth (2013). *Big data. La revolución de los datos masivos*. Turner.

Moore, David.S. (2005). *Estadística aplicada básica*. Antoni Bosch.

O'Neil, Cathy. (2017). *Armas de destrucción matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*. Capitan Swing.

Patino, Bruno (2020). *La civilización de la memoria de pez. Pequeño tratado sobre el mercado de la atención*. Alianza.

Mierzejewski, Mikołaj; Lee, Maggie; Grün, Gianna-Carina; Bradshaw, Paul; Panuccio, Erika; Aldhous, Peter (2018). Data scraping for stories. *Conversations with Data: #12*.
<https://datajournalism.com/read/newsletters/data-scraping-for-stories>

Tascón, Mario (2013). Introducción. Big Data. Pasado, presente, futuro. *Telos: Cuadernos de comunicación e innovación*, 95, 47-50.

Turing, Alan. (1974). *¿Puede pensar una máquina?* Universidad de Valencia.

Programari

A l'tractar-se d'una assignatura completament pràctica el programari requerit és l'habitual per a la realització de les tasques de captura, tractament i anàlisi d'informació en diferents formats.

En concret, es requereixen de les següents eines:

Programari d'edició de textos: Word o similar

Programari d'anàlisi de dades: Excel o similar

Programari de visualització de dades: Infogram - Datawrapper - CARTO - Gephi

Com l'assignatura pot realitzar algunes sessions virtuals, els estudiants han de tenir accés també a un ordinador amb càmera i micròfon i connexió a Internet per poder participar de les sessions on-line.