Peruvian Computing Society (SPC)
School of Computer Science
Sillabus 2021-I

1. COURSE

CS379. Tópicos Avanzados en Ciencia de Datos (Elective)

2. GENERAL INFORMATION

2.1 Credits	:	3
2.2 Theory Hours	:	1 (Weekly)
2.3 Practice Hours	:	2 (Weekly)
2.4 Duration of the period	:	16 weeks
2.5 Type of course	:	Elective
2.6 Modality	:	Face to face
2.7 Prerrequisites	:	CS272. Data Management II. (5^{th} Sem)

3. PROFESSORS

Meetings after coordination with the professor

4. INTRODUCTION TO THE COURSE

Information Management (IM) plays a leading role in almost every area where computers are used. This area includes the capture, digitization, representation, organization, transformation and presentation of information; Algorithms to improve the efficiency and effectiveness of access and update of stored information, data modeling and abstraction, and physical file storage techniques.

It also covers information security, privacy, integrity and protection in a shared environment. Students need to be able to develop conceptual and physical data models, determine which IM methods and techniques are appropriate for a given problem, and be able to select and implement an appropriate IM solution that reflects all applicable constraints, including scalability and Usability.

5. GOALS

- To make the student understand the different applications that the databases have, in the different areas of knowledge.
- Show appropriate ways of storing information based on their various approaches and their subsequent retrieval of information.

6. COMPETENCES

a) An ability to apply knowledge of mathematics, science. (**Usage**)

b) An ability to design and conduct experiments, as well as to analyze and interpret data. (**Assessment**)

d) An ability to function on multidisciplinary teams. (**Assessment**)

i) An ability to use the techniques, skills, and modern computing tools necessary for computing practice. (**Usage**)

7. SPECIFIC COMPETENCES

a14) Properly use files for storage and retrieval of information.

b4) Identify and efficiently apply various algorithmic strategies and data structures for the solution of a problem given certain space and time constraints.

b5) Identify and efficiently apply diverse algorithmic strategies and data structures for the solution of a problem in parallel and distributed environments.

d2) Developing group presentations and reports on specific topics.

d3) Develop group work on each course topic.

i3) Properly use the query optimization, performance, indexing and table fragmentation modules for distributed DBs using an open source database engine such as PostgreSQL, Cassandra or MongoDB

8. TOPICS

Unit 1: Physical Database Design (10)	
Competences Expected: b,j	
Topics	**Learning Outcomes**
• ... • ... • ...	• ... [Usage] • ... [Usage] • ... [Usage]
Readings : [Bur04], [Cel05]	

9. WORKPLAN

9.1 Methodology

Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

9.2 Theory Sessions

The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

9.3 Practical Sessions

The practical sessions are held in class where a series of exercises and/or practical concepts are developed through problem solving, problem solving, specific exercises and/or in application contexts.

10. EVALUATION SYSTEM

********* EVALUATION MISSING ********

11. BASIC BIBLIOGRAPHY

[Bur04] Donald K. Burleson. *Physical Database Design Using Oracle*. CRC Press, 2004.

[Cel05] Joe Celko. *Joe Celko's SQL Programming Style*. Elsevier, 2005.