

# I1 - IIC1005 - 2do semestre de 2020

Jueves 1 de Octubre de 2020, hasta las 22:00 hrs.

Lectura: Exploring Visual Information Flows in Infographics (2020)

Materia: Desde lenguajes de Programación hasta Ingeniería de Software

---

## Indicaciones

- Entrega: Jueves 1 de octubre de 2020, 22:00 hrs vía **canvas**. Debe subir sus respuestas en formato PDF a la [evaluación I1 oficial en canvas](#). **Recuerde responder las preguntas en orden**, si no sabe la respuesta de una pregunta, deje el espacio de respuesta en blanco. Use la siguiente plantilla para responder: [Enlace aquí](#)
- El control es **individual**. La copia será castigada con nota 1.1, además de las sanciones disciplinarias correspondientes.
- El control tiene cuatro partes. Las dos primeras son preguntas de la lectura. Las dos últimas son preguntas de materia.
- **Calificación**: Cada pregunta vale 1 punto.

$$Nota = \frac{puntos}{15} * 6 + 1 \quad (1)$$

---

## Parte 1: Preguntas de Lectura

1. (1 pto.) ¿Qué es VIF? ¿Cuál es la principal contribución del artículo en relación a VIF?
2. (1 pto.) Los autores indican un dataset público que es el más relacionado con infografías. A pesar de eso, sacaron sus imágenes de otros sitios. Indique el nombre del dataset especializado en infografías, así como los dataset que usaron finalmente y al menos 2 razones que motivaron a usar estos datasets.
3. (1 pto.) ¿Qué es YOLO, qué significa (la sigla) y para qué lo usaron en el proyecto? ¿cuál fue el rendimiento promedio de YOLO?

## Parte 2: Preguntas de Lectura de profundización

1. (1 pto.) ¿Qué es Gestalt? ¿Cuáles son los principios de Gestalt usados en esto trabajo? ¿para qué se usan cada uno de los principios en la identificación de VIF?
2. (1 pto.) ¿Cuántos patrones VIF se identificaron? ¿cuáles eran los 3 más frecuentes y los 3 menos comunes? dibuje los 6 patrones consultados (considere los datos de la sección VIF Taxonomy)
3. (1 pto.) Describa dos ideas de trabajo futuro mencionadas en el artículo.

4. (1 pto.) Explique a qué se refiere y por qué se realizó el “multipass manual coding” descrito en la sección *ANALYSIS OF VIF FROM DIVERSE WEBSITES*.

### Parte 3: Preguntas de Materia: Lenguajes, Visualización, Arquitectura, S.O.

1. (1 pto.) (Lenguajes) ¿Cuál es la diferencia entre un lenguaje compilado y un lenguaje interpretado? Explica con tus palabras y nombra al menos una ventaja de cada una. Suponga que deber crear una aplicación con gráficos 3D para una consola de juegos, ¿en qué tipo de lenguaje preferiría de programarlo y por qué?
  2. (1 pto.) (Arquitectura)
    - a) Realice una analogía completa entre el flujo en una app de delivery de comidas (o restaurant) y el flujo completo de una instrucción en una computadora. Como ejemplo, esta sería el primer paso:

1      Fetch -> App/Sistema de informacion (cable) manda orden (instruccion) emitida por usuario (memoria) a la CPU (local de cocina)

  - b) Señale cómo la diferencia entre una GPU y una CPU puede explicar el aumento de popularidad de las GPUs en la última década.
3. (1 pto.) (S.O.) ¿Qué es una syscall? ¿Cuál es su función en un sistema operativo? Nombre dos syscalls, distintas a open o close, explique que hacen y qué consecuencia podría tener en el sistema si no existiesen.
4. (1 pto.) (Visualización de Información) ¿Cuál es el aporte del Dataset de Anscombe’s Quartet (1973) para justificar el uso de visualización? Dibuje dos visualizaciones (creadas por usted mismo), la primera debe tener un alto lie factor, y la segunda un bajo lie factor, explique su elección.

Para representar los documentos multiplicamos la frecuencia de cada palabra **tf** por el peso calculado **idf**

$$\text{idf}(t, D) = \log \frac{N}{|\{d \in D : t \in d\}|}$$

- **Documento 1:** Un auto rojo, un auto verde
- **Documento 2:** Un tomate rojo y un globo amarillo.
- **Documento 3:** Un plátano amarillo y un tomate.

	1	2	3	4	5	6	7	8	9
Doc. 1									
Doc. 2									
Doc. 3									

ID	palabra	idf
1	amarillo	
2	auto	
3	globo	
4	plátano	
5	rojo	
6	tomate	
7	un	
8	verde	
9	y	

Figure 1: Documentos de ejemplo y matriz documento-término que debe completar.

## Parte 4: Preguntas de Materia: Recuperación de Información, Redes, Tecnología Web, Blockchain

1. (1 pto.) (Recuperación de Información) En el modelo de bolsa de palabras, ¿qué significa TF-IDF? Usando los documentos de la Figure 1, calcule los valores vacíos de la matriz documento-término que se indican. No olvide en la columna de la derecha detallar los valores de IDF para cada término. Pueder obviar puntos y comas (no considerarlos como tokens). Considere para TF = ocasiones en que aparece el término en el documento.
2. (1 pto.) (Ingeniería de Software) El profesor Navón indicó en su presentación que una característica importante de un proyecto de Ingeniería de Software es especificar claramente los requisitos. Suponga, como en el ejemplo de la clase, que le piden hacer un proyecto con una página web para reportar un dashboard con gráficos de la empresa, y que al final del proyecto le pagarán \$2.000.000.- Describa al menos 5 consultas que haría usted para clarificar requisitos (responda en función de aspectos mencionados por la presentación del profesor en clases)
3. (1 pto.) (Tecnología Web) Escribe el código HTML con el que podrías generar la página que se ve en la Figure 2.
4. (1 pto.) (Blockchain) (a) ¿Es posible reconstruir los datos de un bloque a partir de su hash y el hash del bloque anterior? (b) ¿Es posible verificar la integridad de los datos de un bloque a partir de su hash y el hash del bloque anterior? Explique, para cada caso, cómo lo haría o por qué no se puede.

### Este es un título

Este es un párrafo que no dice nada interesante.

[Link al github del curso \(si presionas aquí te vas a dirigir al syllabus de la página de github\)](#)

---

### Lista de cursos

A continuación, una lista con algunos cursos del DCC:

- IIC1005
- IIC1103
- IIC2413
- IIC2613

Figure 2: Página HTML que debe generar.

---