



**SIMPOSIO
TECNOLOGÍA DISRUPTIVA
para la SEGURIDAD HÍDRICA**

La Revolución de los Datos: Retos para la Ciencia y la Sociedad

PANEL

Chris Stephens

C3-Centro de Ciencias de la Complejidad y
Instituto de Ciencias Nucleares, UNAM

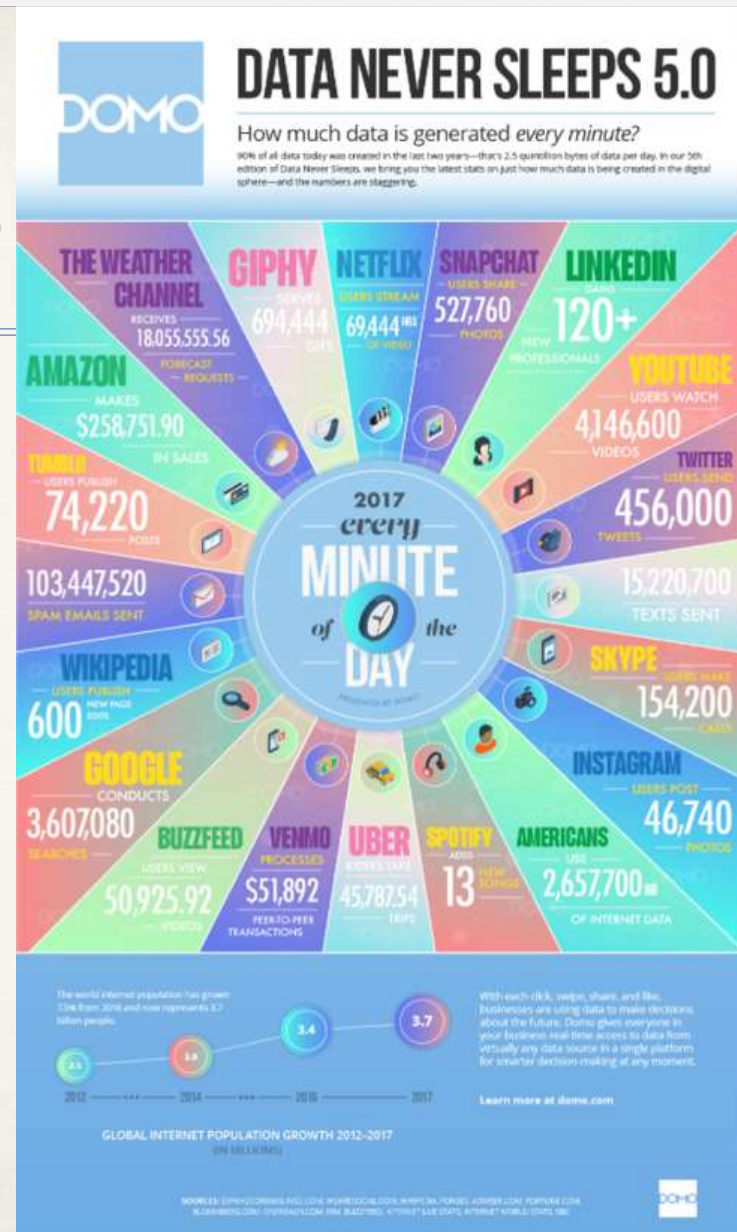
25 y 26 de octubre de 2018



Ha habido una “Revolución de los Datos”

Afecta profundamente
a nuestras vidas como
científicos y ciudadanos

¿Pero, que exactamente
es revolucionario?





Datos profundos: La Revolución de los Datos y la Toma de Decisión



Una revolución en la generación de datos

No ha habido una revolución en:
Tipo de datos
Velocidad
Propósito



Cerebro humano
10-100 Terrabytes

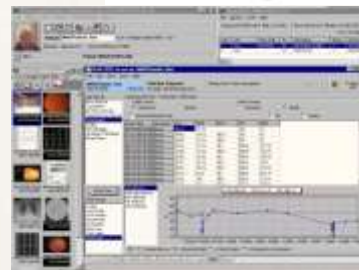
Todos los libros en el mundo
30-50 Terrabytes



Una revolución en el análisis de datos

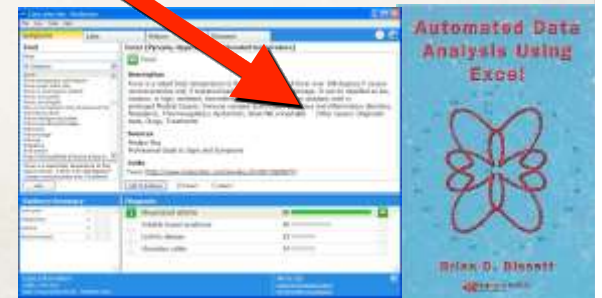


1 genoma humana
= 1GB (200)
Imagen de CT
= 10MB
Imagen de MRI
= 40MB



En forma electrónica
1 zettabyte

Una revolución en el almacenamiento de datos



¿Porqué la Revolución de los Datos es y será tan importante?



El propósito principal de los seres vivos (y un propósito principal de la ciencia) es... **Predecir** para la

Toma de Decisiones

Todos los seres vivos son “mineros de datos” y minamos BI G data

¿Dónde están estos datos?



Antes



Después

de la Revolución de los Datos

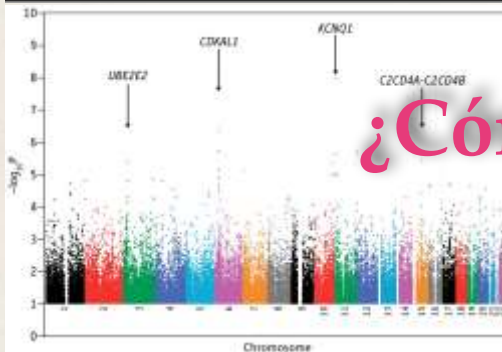




¿Qué representan los datos de la Revolución?



Sistemas Complejos
Adaptativos



¿Cómo los analizamos?

Food Database

Group Search by Name Favorites Saved Meals

1st Food Restaurants
its, oils, and salad dressings
uits

Citrus fruits, juices

- Apple juice, fresh [1 c]
- Apple-orange juice [1 c]
- Apicol-orange juice [1 c]
- Grapefruit & orange juice, canned, unsweetened [1 c]
- Grapefruit & orange juice, canned, w/ sugar [1 c]
- Grapefruit & orange juice, fresh [1 c]
- Grapefruit & orange juice, frozen, (reconstituted) [1 c]
- Grapefruit & orange juice, nfs [1 c]
- Grapefruit juice, canned, bottled, carton, unsweet [1 c]
- Grapefruit juice, canned, bottled, carton, w/ sugar [1 c]
- Grapefruit juice, freshly squeezed [1 c]
- Grapefruit juice, frozen, unsweetened (reconst) [1 c]
- Grapefruit juice, frozen, w/ sugar (reconstituted) [1 c]
- Grapefruit juice, nfs [1 c]
- Grape-tangerine-lemon juice [1 c]
- Lemon juice, canned or bottled [1 c]
- Lemon juice, fresh [1 c]
- Lemon juice, frozen [1 c]

Food Information

Name:
Apicol-orange juice [1 c]

Carbohydrates:	30.3	g
Fiber:	1	g
Sugar:	27.8	g
Protein:	1.2	g
Fat:	0.3	g
Saturated Fat:	0	g
Calories:	122.5	cal
Cal from Fat:	2.7	cal
Sodium:	7.5	mg
Cholesterol:	0	mg

Show in favorites list

Save Changes Cancel Changes

Collapse All Mark as Favorite New Group New Food Delete Item



The worldview of the last 3 centuries:

NO EXCEPTIONS.

How... es?



Jacob L.A. Moravitz | Invention

The... we are slaves of the law... ne

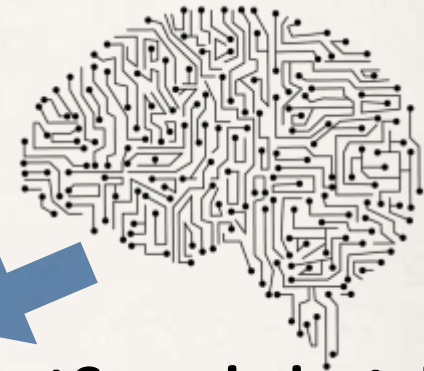
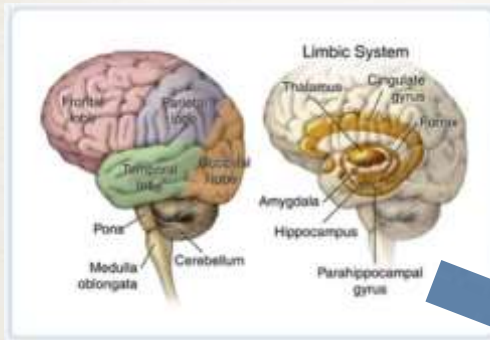


YOU WILL GET A TICKET FOR NOT WEARING YOUR SAFETY BELT.

Safety belt enforcement is being stepped up everywhere. It doesn't matter where you drive, they'll be looking for you. Simply, law enforcement writes tickets to save lives. So buckle up or you will get a ticket. No exceptions.



Representación Algorítmica de una Predicción/Decisión



Conocimiento

$$P(C | X(t))$$

¿Conocimiento?

Heurística

Datos + Información

Conocimiento

¿Qué formato?

Lenguaje natural, numérico, categórico, espacio-temporal

Los $P(C|X)$ son "maquinas" como maquinas físicas hacen una cosa bien
Cada decisión/acción que tomamos requiere y usa una maquina
DIFFERENTE, independientemente de si o no usamos HIA



La Importancia de Conocimiento de Dominio



Hay muchas cajas de herramientas



Seleccionando un (C|X) es como seleccionar una herramienta. Pero...



Para problemas complejos - multi-factorial y multi-escala - se requiere múltiples disciplinas y trabajo en equipo

Misión del C3 -Centro de Ciencias de la Complejidad



... realizar investigación científica transdisciplinaria de frontera en las ciencias de la complejidad, creando un espacio en donde expertos de muy diversas áreas puedan interactuar y contribuir a la solución de problemas trascendentes y de importancia nacional

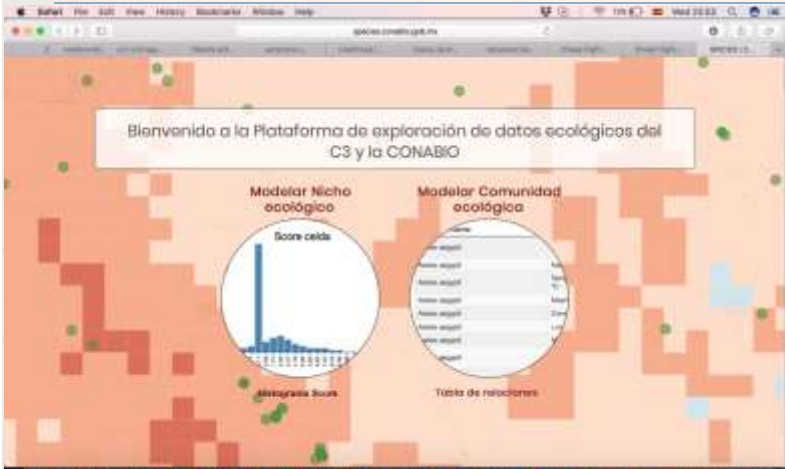
Es también misión del Centro formar científicos entrenados en el trabajo transdisciplinario en equipo y en el fortalecimiento de los métodos modernos asociados a la ciencia computacional.

Importancia de buscar sinergias y colaboración con el CViCom



Algunos Proyectos del C3.

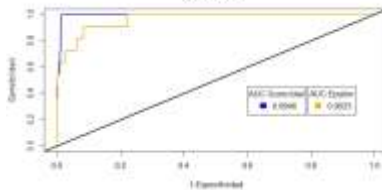
La computación juega un papel importante en todos



Detección de interacciones biológicas

- Depredación
- *Lynx rufus*

	Presa	Score total
<i>Sylvilagus auduboni</i>	1	32.9163696
<i>Sylvilagus bachmani</i>	0	32.9163696
<i>Nasua nasua</i>	0	30.5265046
<i>Reithrodontomys microdan</i>	1	29.2840247
<i>Dipodomys ordii</i>	0	29.1183394
<i>Peromyscus eremicus</i>	0	29.0314133
<i>Peromyscus lewipes</i>	1	28.1092097
<i>Chaetodipus penicillatus</i>	1	27.9424153
<i>Neotoma floridana</i>	0	26.9401445
<i>Sigmodon fulviventer</i>	0	26.8817102
<i>Sylvilagus graysoni</i>	0	26.4597068
<i>Reithrodontomys montanus</i>	0	26.3426345
<i>Callipepla gambelli</i>	0	26.3158638
<i>Pipoides stricklandi</i>	0	25.8401668
<i>Antrastomus orizonae</i>	0	25.5477411
<i>Columba inca</i>	0	25.5339324
<i>Spizella angustifrons</i>	0	25.0583737
<i>Pipoides villosus</i>	0	24.9179631
<i>Keraperomphlus spilosoma</i>	1	24.6575532
<i>Sorex emarginatus</i>	0	23.9509356



Enfermedades Crónicas

“Salud y enfermedad; un enfoque desde las Ciencias de la Complejidad en la búsqueda de alarmas tempranas”. PAPIIT

“Enfermedad y Salud: un enfoque desde las ciencias de la complejidad”. Fronteras de la Ciencia, CONACyT

“Salud y enfermedad y las Alertas Tempranas”. Fronteras de la Ciencia, CONACyT

“Obesidad y Diabetes y Médico en tu Casa - Colaboración C3, Lab CDMX, SEDESA

Enfermedades Emergentes

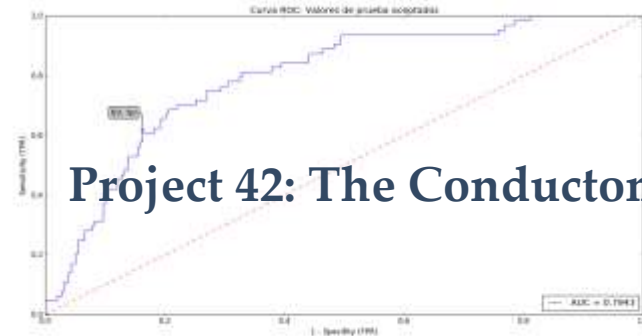
“Complejidad ecológica de las enfermedades emergentes”. PAPIIT

“SPECIES”, CONABIO

“Vacuna para la Enfermedad de Chagas”, Fundación Carlos Slim

Cancer (Biología de Sistemas)

“Cancer como una enfermedad compleja: leucemia y cáncer epitelial, de lo micro a lo macro”



Resultados de modelos predictivos basados en datos de un estudio de 1,800 no-académicos, académicos y estudiantes de la UNAM: 2,524 variables - Genético, epidemiológico, fisiológico,...

Conclusiones



La Revolución de los Datos esta generando ~ 1 Zettabyte de datos cada año. Estamos ahogando en datos.

No tenemos la más mínima idea de cuanto “conocimiento” hay en estos datos. No hablamos “base de datos”.

En esos datos hay posibilidad de salvar/ quitar vidas, evitar/ fomentar conflicto, reducir/ aumentar la pobreza,..., representar/ distorsionar realidad

El uso (y abuso) de estos datos requiere la Ciencia de la Computación, pero también requiere muchas más disciplinas (semántica de los datos).

Para la solución de los grandes problemas complejos que enfrentamos, la sociedad necesita producir nuevas maneras de hacer ciencia donde la computación juega un papel metodológico fundamental en proyectos transdisciplinarios de gran importancia.

Hay oportunidades para colaboraciones fructíferas entre el C3 y el IMTA en estos problemas.



SIMPOSIO
TECNOLOGÍA DISRUPTIVA
para la **SEGURIDAD HÍDRICA**



stephens@nucleares.unam.mx



Página ponente



@Twitter

www.aguadisruptiva.com

25 y 26 de octubre de 2018