

# Datos científicos abiertos en Chile

Claudio Gutiérrez  
FCFM, Universidad de Chile

# Un nuevo escenario

El rápido desarrollo de la tecnología computacional y de Internet ha generado nuevas y poderosas formas de capturar, almacenar y procesar gigantescas cantidades de datos.

Fenómeno conocido como  
“diluvio de datos”.

# Diluvio de datos: cuál es el cambio cualitativo

1. Cantidad de datos >> human I/O
2. Complejidad datos >> human CPU

En resumen: capacidad de “comprender” y “abarcar” los datos disponibles sobrepasa los límites de la capacidad humana.

# All Thinks, Great and Small

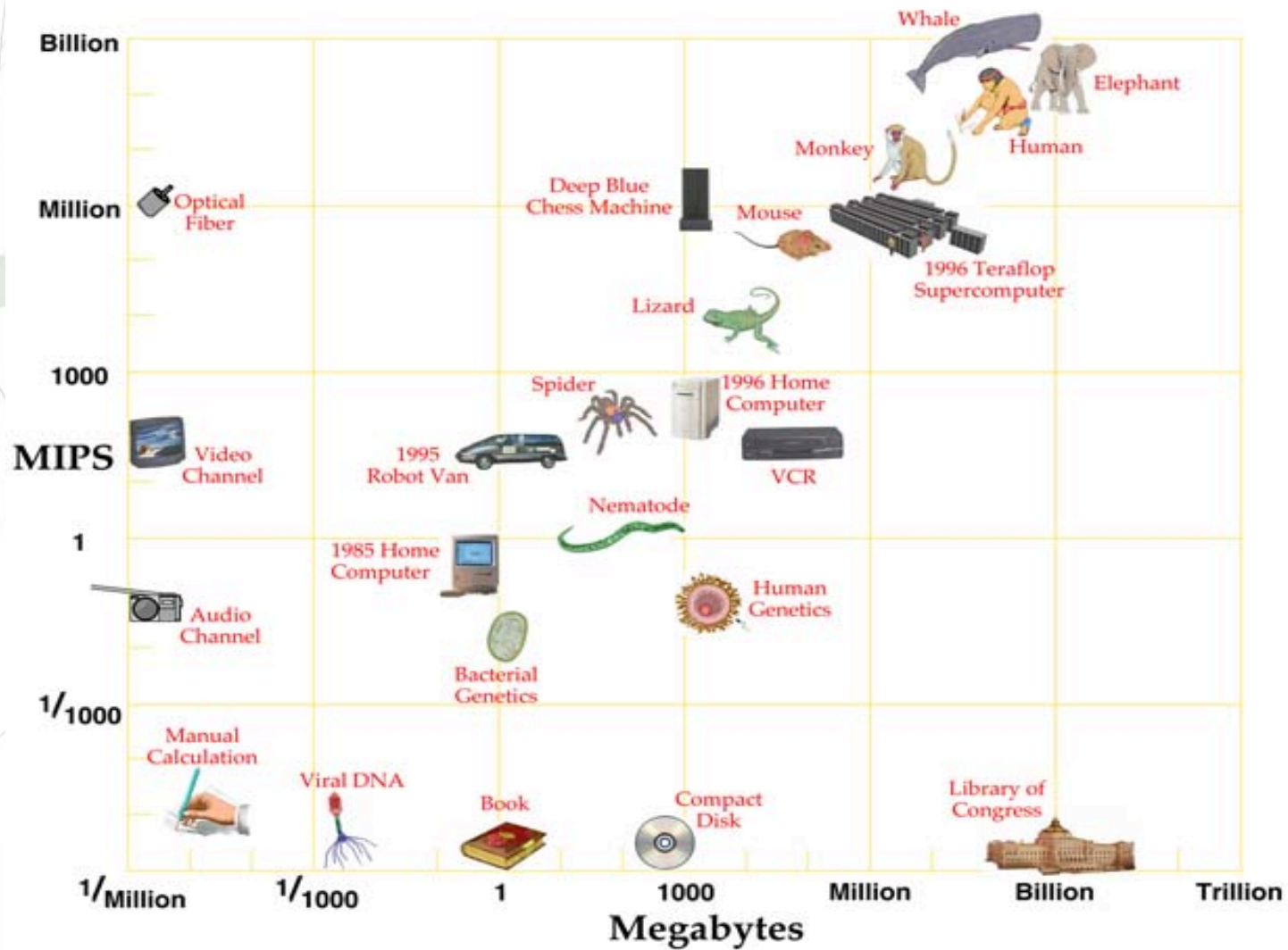
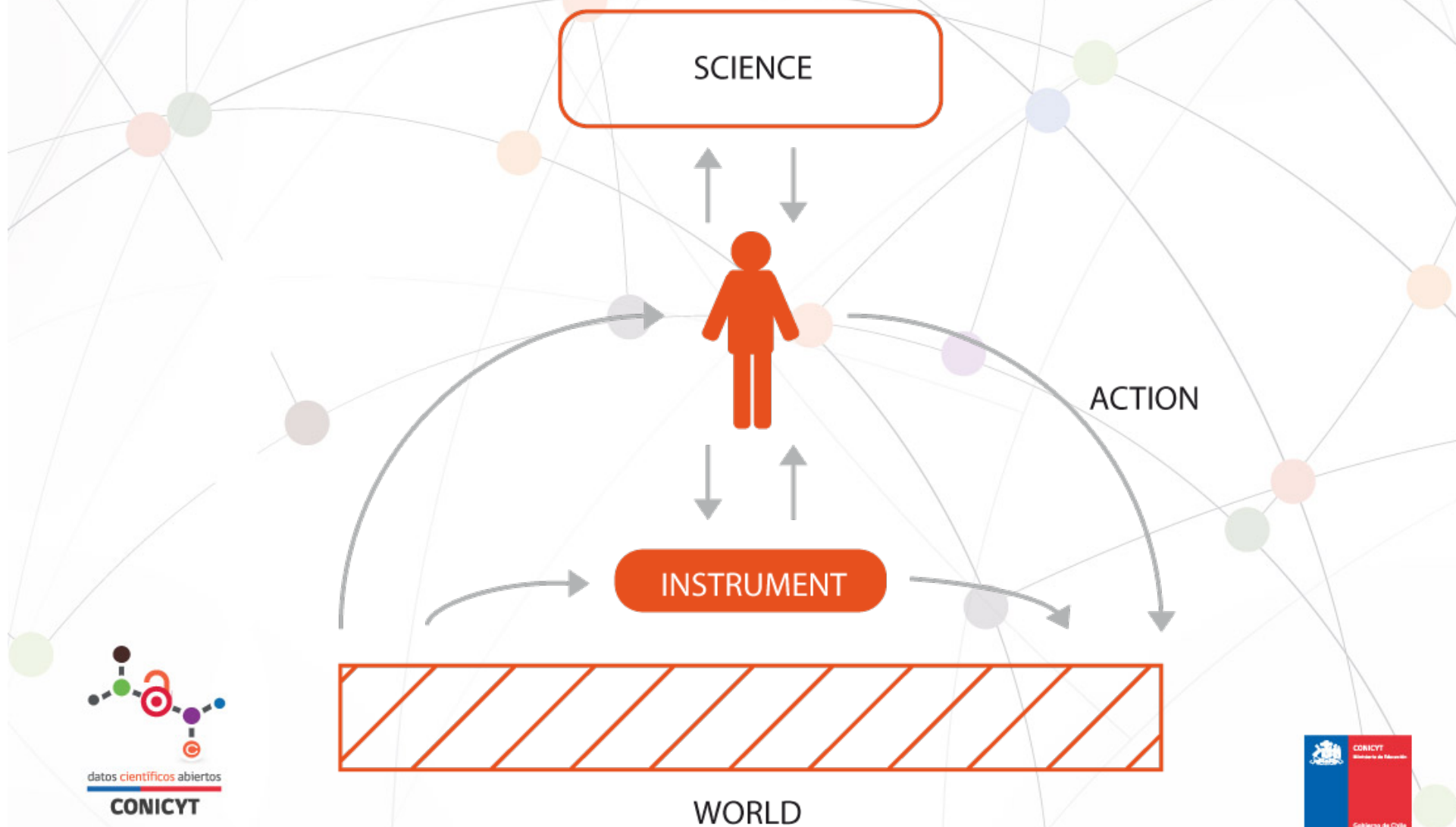
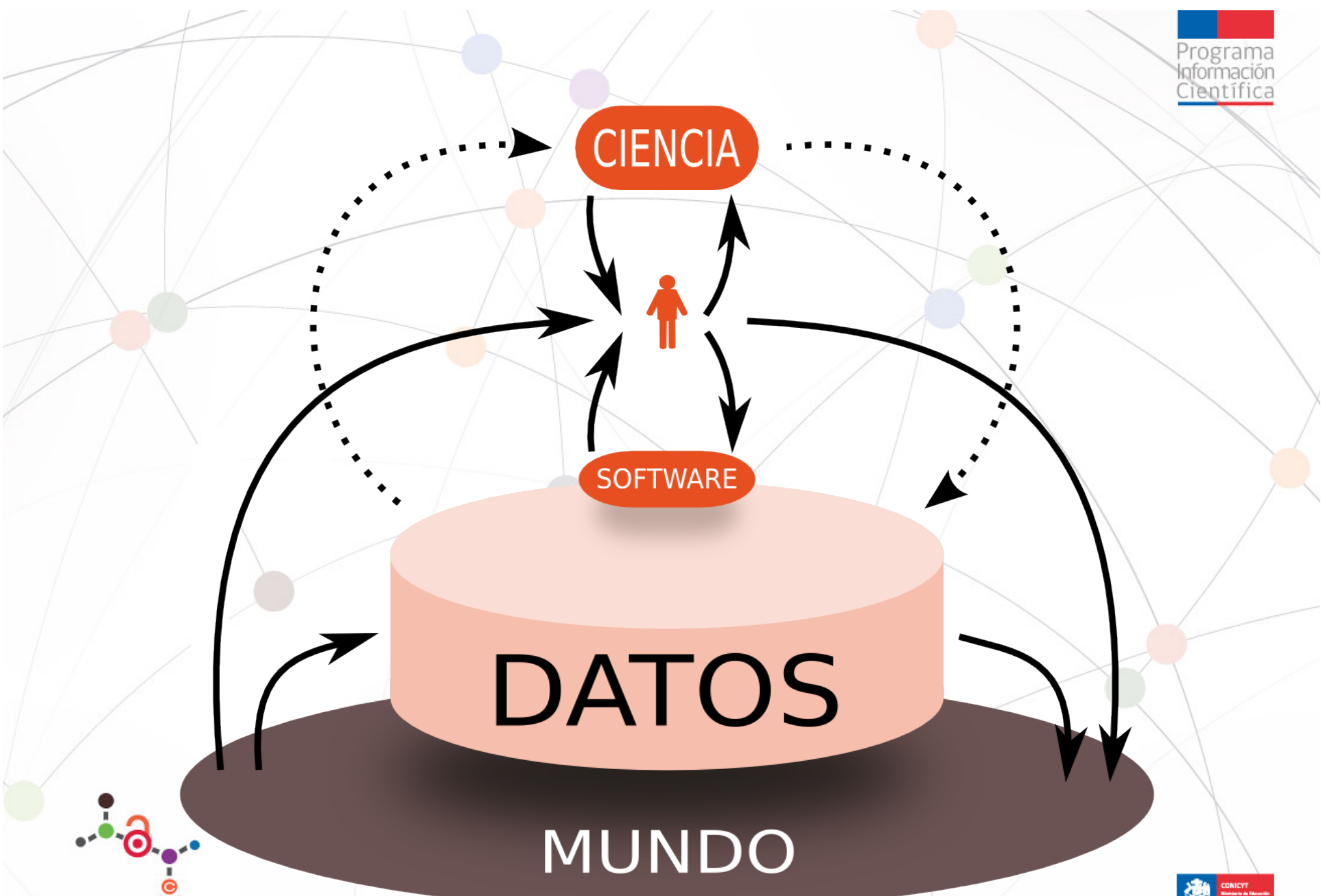


Image by H. Moravec

# El mundo antes del diluvio (de datos)





# Un nuevo desafío para la ciencia

1. Cambios en el quehacer y tareas de los científicos
2. Necesidad de ayuda en soporte, archivo e infraestructura de datos científicos
3. Nuevas responsabilidades de agencias de C&T y de gobiernos e instituciones

# Aclaración necesaria

1. **Información abierta:** acceso libre a información (gobierno, salud, educación, etc.) ej. Transparencia.
2. **Conocimiento abierto:** acceso abierto a los *productos procesados* del conocimiento (libros, papers, obras, etc.) ej. Scielo.
3. **Datos abiertos:** acceso abierto a *insumos y productos no procesados* del conocimiento (datos “crudos”)



# Grandes desafíos

1. **Motivar/incentivar** a los científicos. Dar a conocer alcances de la iniciativa y los beneficios y protección que conlleva.
2. **Crear política/infraestructura** para:
  1. Facilitar e incentivar archivar y curar los datos de investigaciones
  2. Disponibilizar esos datos de manera abierta y masiva.

# Desafíos técnicos

- ***Físicos***: formatos y repositorios
- ***Conceptuales***: Indexación, metadatos, catalogaciones
- ***Sociales/uso***: Interoperabilidad, búsqueda, disponibilización
- ***Privacidad/Seguridad***: protocolos, roles, etc.

# El mayor desafío

## *El valor de los datos*

- Cuáles son los datos importantes?
- ¿Calidad de los datos (técnica, de contenidos, etc.)?
- ¿Qué priorizar?
- Relevancia

# Casos de éxito en Chile

- Gobierno Transparente
- Data.gob.cl
- Sesiones de los Cuerpos Legislativos de la Rep. Chile (1810-1845)
- Biblioteca Digital, Univ. Chile
- Diversas experiencias sectoriales

# ¿Cómo empezar?

1. Observa cómo lo hacen tus pares
2. Asesórate con gente del área de la información
3. Identifica y describe tus datos
4. Publica tus datos donde puedan perdurar
5. Difunde y comparte esa publicación
6. Considera usar una base de datos
7. Más sugerencias en

<http://datoscientificos.cl>

# ¿Cómo estamos?

★ disponga su material en la web (cualquier formato)

★★ dispóngalo como datos estructurados (ej. Excel en vez de una imagen escaneada)

★★★ dispóngalos en formato no propietario (ej. CSV en vez de excel)

★★★★ use URLs para identificar objetos, de tal manera que otra gente pueda referenciarlo

★★★★★ enlace sus datos a los de otra gente para proveerle contexto a ellos

Muchas gracias!

*Hazte parte.  
propone, opina, participa*

La ciencia la hacemos entre todos

# Materiales de apoyo

- Sitios Web:  
<http://datoscientificos.cl>  
<http://datosabiertos.cl>
- OECD Principles and Guidelines for Access to Research Data  
<http://www.oecd.org>



# Principios OCDE

- Apertura
- Flexibilidad
- Transparencia
- Conformidad legal
- Protección de propiedad intelectual
- Responsabilidad formal
- Profesionalismo

# Beneficios públicos

El acceso a los datos de investigaciones:

1. incrementa el retorno de las inversiones públicas en esta área;
2. Refuerza la investigación científica abierta
3. Incentiva la diversidad de estudios y de opiniones
4. Promueve nuevas áreas de trabajo y permite la exploración de tópicos no vistos por los investigadores originales

# Beneficios individuales

El acceso a los datos de investigaciones:

1. Mayor visibilidad y credibilidad a la investigación y los investigadores
2. Incentiva nuevas investigaciones sobre el tema (y luego más citas y relevancia del tópico)
3. Facilita investigaciones cruzadas y colaboración con otros colegas