

Algunas Observaciones sobre Investigación (en Bases de Datos)

Leo Bertossi

Carleton University

School of Computer Science

bertossi@scs.carleton.ca

www.scs.carleton.ca/~bertossi

Faculty Fellow of the IBM Centre for Advanced Studies,
Toronto Lab.

Qué investigar ...

En qué se debe investigar?

Creo en la investigación motivada por la curiosidad; lo que incluye:

- La natural necesidad humana de comprender el mundo

La noción de computación es connatural al ser humano y, de hecho, previa al advenimiento de los computadores

Los computadores (en un sentido extendido), aún siendo artefactos inventados por el ser humano, son parte de este mundo y por consiguiente objetos de estudio tan válidos como otros

- El deseo y voluntad (y necesidad) de querer resolver problemas y hacer que ciertas cosas “funcionen”
(creo menos en que sea la “necesidad” altruista la que motive a la mayoría de los científicos aplicados)

La curiosidad es y debe ser la fuerza motriz detrás de la investigación

Calidad ...

Entonces, da lo mismo en qué se investigue?

No, si es que el investigador está inserto en instituciones que financian/apoyan, ... su investigación

El investigador debe investigar en aquellas áreas donde pueda demostrar un razonable grado de impacto y aporte

Básicamente ante y por sus pares de la comunidad **internacional**

En suma, el investigador debe tener la meta y voluntad de superar los estándares mínimos de calidad de su área

Todo esto se aplica a bases de datos; y ahí hay una más o menos clara medida de calidad

Hay journals y conferencias de distintos niveles de calidad, pero que pueden ser comparados

Algunas vías de publicación

Hay una serie de journals generales que también son muy buenos lugares para publicar en base de datos

- Journal of the ACM
- Journal of Computer and Systems Science
- Theoretical Computer Science
- Information and Computation
- ...

Otros journals específicos de bases de datos

- ACM Transactions on Database Systems
- IEEE Transactions on Knowledge and Data Engineering
- VLDB Journal
- Information Systems
- ...

Hay también algunos muy buenos específicos de ciertas áreas (information retrieval, data mining, web, ...)

Conferencias internacionales

- ACM Symposium on Principles of Database Systems (PODS)
- International Conference on Database Theory (ICDT)
- ACM SIGMOD
- International Conference on Very Large Databases (VLDB)
- Extended Database Technology (EDBT)
- International Conference on Data Engineering (ICDE)
- ...

También muy buenas conferencias en otras áreas que también publican sobre bases de datos: IJCAI, LICS, ...

Dudas? Casi siempre pueden ser resueltas mirando:

<http://citeseer.nj.nec.com/impact.html>

(lista jerarquizada; 1220 CS/CE journals y conferences)

En CS/CE mejor indicador que ISI (que nadie en CS conoce fuera de Chile) Peor es nada?

Qué investigar?

- Cualquier cosa bien hecha ...
- Mirar qué se está investigando, mirando los principales lugares de publicación
 - Primero los proceedings (usualmente on-line) de las principales conferencias: SIGMOD, PODS, VLDB, ICDT, EDBT, ICDE, ...
 - Segundo, los principales journals
 - Otras publicaciones más ligeras, pero muy informativas y rápidas:
 - * ACM SIGMOD Record
 - * IEEE Bulletin on Data Engineering
 - * IEEE Intelligent Systems
 - * Communications of the ACM
 - * IEEE Computer

- Varios informes de comités sobre dónde falta investigación

Partir con el Lowell Report (Mayo 2003) y las referencias en él:

<http://research.microsoft.com/~gray/lowell/>

Cómo partir

- Definitivamente pensando en publicar al más alto nivel internacional

La meta máxima sobre lugar de publicación no puede ser local

No sólo asegura calidad, sino también exposición del trabajo y del **investigador** a la comunidad donde está la acción ...

- Sin atacar la publicación “local”, **nunca he visto funcionar** la idea del

“partir publicando en medios locales o malos para llegar (por milagro?) a publicar en medios de transcendencia”

- Buscar asociación en investigación con investigadores de reconocida calidad fuera de lugar de uno, ojalá en el extranjero

- Mantenerse informado sobre qué está pasando en los lugares donde está la verdadera acción (ver transparencia anterior)
- Ser MUY cuidadoso en la elección de supervisores de tesis, mentores, ...

Chequear cómo están posicionados en investigación a nivel internacional

- Relevancia de los temas de investigación
- Reconocimiento transitivo por los pares
- Red de contactos
- Hábitos de trabajo, visión de los problemas, formas de ataque de los problemas, ...
- Ambiente nutritivo en torno al supervisor
- Cómo escribir papers
- ...

Si no tiene tenure, es una estrella?

No dejarse impresionar por lista de publicaciones en la pagina web (ISI no es garantía de calidad)

Numerosa gente con potencial de investigador se ha perdido por mala elección del supervisor

Por dónde va la cosa ...

Menos por el lado de “core DB technology”

Ver, e.g. Call for paper VLDB 2004:

- Core Database Technology
- **Infrastructure for Information Systems**
- Industrial and Applications

BDs tienden a ser vistas como una “commodity”

Hay que darles valor agregado

Agregarles “inteligencia”

- Data mining
- OLAP
- Nuevas capacidades de consulta, e.g. con información incompleta, incierta, inconsistente, ...
- Uso de metainformación
- Capacidad de integración con otras bases de datos
- Seguridad en BDs: políticas e implementaciones declarativas
- ...

Insertarlas en aplicaciones complejas

- Web services (nuevos modelos transaccionales)
- Semantic web
- Knowledge Based Systems (ontologies)
- Workflows
- Aplicaciones científicas
- Mobile computing, pervasive computing
- ...