Online Ethics Center
FOR ENGINEERING AND SCIENCE

Database Discredited

Author(s)

R. Waldo Roth

Year

2006

Description

A credit bureau realizes that their credit file contains errors. A consultant is hired to analyze the problem and make recommendations. Many questions on the numerical and ethical aspects are discussed in the scenario.

Body

Introduction

As an industrial engineering consultant hired by a credit bureau ("RWT"), you have been asked to analyze problems that have occurred with their 20 million record

credit file. RWT management became concerned when the following situation came to their attention. 1, 2

A couple moving to a retirement community has an eye on their "dream home." Because they have a good credit history, they assume that they will have no trouble obtaining a mortgage to purchase this home through a local bank in their new community. A routine credit check through RWT reveals that in fact they are a bad credit risk. When a representative from the local bank pursues the case, she discovers that the couple has been mis-identified in the RWT database, which has confused them with another party having a very bad credit history. In making amends, the local bank approves the loan, but by now the home has already been sold to someone else. The couple is heartbroken and worse yet, they continue to experience credit problems for some time.

Back to Top

Numerical and/or Design Problem(s)

1. RWT has called you in as a consultant to make recommendations. Where do you begin?
2. What design flaws in the database have allowed this problem to occur?
3. Management at RWT claims they have only 1 error per 100,000 records in their databank. How would you develop an experimental design test this claim?
4. After determining whether RWT is correct in their claim that they have only one error per 100,000 records, determine how many bad records they have in their major file at the present time. What implications does this have, if any?
5. At what cost per record do you decide they need to rework the database? What other data or assumptions do you need to make before recommending a solution?

Assume you need:

1. a cost per record to update,
2. the number of errors per year (which can be estimated), and
3. the cost per error found by users.
4. You also need to know the rate of updates or new records per month or year. Assume there are 300,000 accesses or updates per month to the database.

5. Also assume it costs $10/record to clean up the database plus $50,000 in fixed costs. Finally, assume the cost for insurance, lawyers. etc. for each bad record found is $100,000.
6. Compare the two costs, draw conclusions, and recommend a course of action in a one page memo to RWT management. Or write a dialogue for a discussion of the matter with RWT.
7. Estimate the time required to clean up the database. Can you design a solution that would not take the database off-line for that amount of time?

Back to Top

Questions on Ethics and Professionalism

1. What are the relevant facts?
2. List the "stockholders" -- those with something to lose (or win) in this case.
3. Who benefits here? Who is harmed here? (There may be multiple answers.)
4. What ethical principles from your textbook apply to this case?
5. How could you prevent this situation from happening again?

Back to Top

Solutions to the Numerical Problems

1. You would first want to find out how often such errors occur and what the source of the typical error is in the system (e.g. data entry, updates, software, indices, etc.)
2. This assumes there is a design flaw. There may not be such a flaw at all.
3. First, verify how many records they actually have. Then you would take approximately 50 random records from the 20,000,000 records they claim to have. Then, every 20 mea/X = 50 records should be chosen. Hence, selecting "every 400,000th" record (X = 400K) in some random fashion would yield 50 records. Later, if a pattern emerges as to what type of records are in error, a subset of those could be randomly tested.

4. This would mean they have 200 bad records. That is really rather impressive if it turns out to be correct. Also, the type of error would be significant. A small address error may be trivial as compared to a pointer to the wrong person's record. If all 200 errors are bad pointers, RWT needs a major software overhaul.
5. If they have 300,000 transactions per month and that is where the errors occur, then there will be 3 x 12 or 36 errors per year entered into the system. However, the far tougher problem will be in finding these errors!
6. Using the assumptions made earlier, one may conclude that it will take about 60 years to check and verify the entire *static* file and to justify the cost of clean-up. If one assumes $1 million in cost per loss, it will still take 6 years to break-even. One can ignore the $50,000 fixed cost as irrelevant in the calculation.
7. A software solution might be implemented by taking the system down over a weekend. Otherwise, the snail's pace of the cleanup and the use of any other pseudo-manual system may immobilize employees.

Back to Top

Possible answers to ethical questions:

1. Responsibility for this data rests with the credit bureau (RWT).
2. Everyone listed in the problem and the public as well.
3. Probably not. These mistakes will occur in any file system.
4. No one benefits.

5. GIGO means errors must be dealt with and resolved in an ethical manner. Good software engineering design from the outset SHOULD solve the pointer problems for errors in the data. Ongoing spot checking is probably the only way to GIGO in the file, plus consistent cleanup of errors as they are found.

6. One probably can't!

Appendices, Attachments, and Bibliography

- 1. This case is based on an idea from *Ethical Decision Making and Information Technology,* Kallman and Grillo, McGraw-Hill, 1994. The case is called "Credit

Woes," see p. 59.

- 2. Note: There is some similarity in this case to the famous Ford Pinto case. See Harris, Pritchard, and Rabins, *Engineering Ethics: Concepts and Cases*, p. 205.
- There is now legal recourse for the couple in such a case, but the law focuses on the responsibilities of the credit supplier (how the data can be corrected) rather than on how individuals can resolve follow-up errors.

Back to Top

Notes

Author: R. Waldo Roth, Computing and System Sciences Department, Taylor University, 500 West Reade Avenue Upland, Indiana 46989-1001.

These problems were originally developed as part of an NSF-funded project to create numerical problems that raise ethical issues for use in engineering and other course assignments. The problems presented here have been edited slightly for clarity.

Rights

Resource Type

Case Study / Scenario

Parent Collection

Numerical & Design Problems With Ethical Content

Topics

Privacy and Surveillance
Public Well-being

Discipline(s)

Computer Sciences
Computer, Math, and Physical Sciences
Information Sciences
Social and Behavioral Sciences

Publisher

Online Ethics Center