Volume 12, No. 5
August 2014
ISSN: 1545-679X

INFORMATION SYSTEMS EDUCATION JOURNAL

Special Issue: Teaching Cases

In this issue:

The **Information Systems Education Journal** (ISEDJ) is a double-blind peer-reviewed academic journal published by **EDSIG**, the Education Special Interest Group of AITP, the Association of Information Technology Professionals (Chicago, Illinois). Publishing frequency is six times per year. The first year of publication is 2003.

ISEDJ is published online (http://isedjorg) in connection with ISECON, the Information Systems Education Conference, which is also double-blind peer reviewed. Our sister publication, the Proceedings of ISECON (http://isecon.org) features all papers, panels, workshops, and presentations from the conference.

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the ISEDJ journal. Currently the target acceptance rate for the journal is about 45%.

Information Systems Education Journal is pleased to be listed in the 1st Edition of Cabell's Directory of Publishing Opportunities in Educational Technology and Library Science, in both the electronic and printed editions. Questions should be addressed to the editor at editor@isedj.org or the publisher at publisher@isedj.org.

INFORMATION SYSTEMS EDUCATION JOURNAL

Editors

ISEDJ Editorial Board

Teaching Case

Cleaning Data Helps Clean the Air

Kelley Donalds
kelley.donalds@bridgew.edu

Xiangrong Liu
Xiangrong.liu@bridgew.edu

Bridgewater State University
Bridgewater, MA

Abstract

In this project, students use a real-world, complex database and experience firsthand the consequences of inadequate data modeling. The U.S. Environmental Protection Agency created the database as part of a multimillion dollar data collection effort undertaken in order to set limits on air pollutants from electric power plants. First, students explore the database to identify design limitations from the perspective of a data analyst with a specific goal. Second, students create a new database design which overcomes identified problems. Through this case study, students develop the skill to infer usage implications by studying the design of an existing database. This is important since developers often inherit databases designed by others. Students also learn how to prepare data stored in a relational database for a data analysis project. By experiencing the consequences of an inadequate design from a user perspective, students can better appreciate the importance of relational database design principles and become more committed to using them.

Keywords: database design, data modeling, data cleaning, referential integrity, normalization

1. INTRODUCTION

John and Kayla had just started their new jobs as Data Analysts at the Utility Research Institute (URI), a non-profit organization that conducts research on behalf of its funding organizations -- primarily electric utility companies operating within the United States. John had an M.S. in Computer Science and had worked as a Database Management Administrator for the past five years. Kayla had just graduated with an M.S. in Mathematics with a concentration in Statistics. They had been assigned to work together on a project analyzing data that was compiled by the U.S. Environmental Protection Agency (EPA). The data was collected as part of a process for establishing the first ever national standards limiting emissions of hazardous air pollutants such as mercury from coal and oil fired power plants. The EPA had made the data available to the public in the form of a Microsoft Access database and the Institute wanted to use this data to determine boiler features and pollution control equipment that would satisfy emission standards for all of the newly regulated pollutants.

2. BACKGROUND

To kick off the project, Kayla and John's manager, Ravi, briefed them on the regulatory history of air pollutants within the U.S. utility industry. He said

that the recent 2011 ruling, known as the Mercury and Air Toxics Rule (MATR), imposed the first ever national limits on heavy metals such as mercury and acid gas emissions from coal and oil power plants (EPA 2011b). The rule specifically limited air emissions of mercury, filterable particulate matter and hydrochloric acid from coal and oil fired plants with at least 25 megawatt hours of generating capacity. The EPA had decided to use filterable particulate matter and hydrochloric acid as surrogates for all non-mercury metals and all acid gases respectively. To show the importance of the project, Ravi shared an article from a trade journal (Neville 2012) in which industry representatives described MATR as the most expensive regulation under the Clean Air Act (CAA) in terms of direct costs and the most extensive intervention into the power market that the EPA had ever attempted. EPA's own detailed analysis estimated that the rule would affect about 500 coal-fired plants and 100 oil fired plants at an annual cost of $9.6 billion (EPA 2011a). Given the significant compliance costs, it was likely that some utilities would be making "invest or retire" decisions for many plants -- especially older ones.

John asked Ravi why the electric utilities hadn't been subject to earlier regulation of these air pollutants. Ravi explained that while electric utilities were no stranger to regulation under the CAA, they had been treated differently than other industries in the major 1990 amendments (EPA 2013a). Congress passed these major revisions to better control urban air pollution (Title I), pollutants from mobile sources (Title II), toxic air emissions (Title III), acid rain (Title IV) and ozone-depleting chemicals (Title VI). Title V delegated responsibility for regulatory oversight to individual states via a permitting process. Title IV had imposed significant regulations on the utility industry to better control emissions of sulfur dioxide which contributes to acid rain. Title I had imposed limits on emissions of nitrous oxides and particulate matter which contribute to urban area smog and also impacted the utility industry.

However, Ravi explained that the electric utility industry had successfully forestalled regulation under Title III of the amendments (e.g. Lemonick 1990). Title III listed 189 air toxins for which the EPA was required to identify source categories that would be subject to future regulation under section 112 of the CAA. Standards under section 112 were based on what was referred to as maximum achievable control technology (MACT). For existing sources, MACT sets a minimum level of stringency called the floor which is the average emission "achieved by the best performing twelve percent of existing sources in the category or the best performing five sources for source categories with less than thirty sources" (EPA 2013b). Quoting the CAA, congress had required the EPA to perform a study of the "hazards to public health reasonably anticipated to occur" as a result of emissions of listed air toxins and to regulate electric utilities under Title III *only* "if the Administrator finds such regulation is appropriate and necessary after considering the results of the study" (EPA 2013c). A general report regarding all the listed air toxins by utilities was due in three years and an additional report addressing health effects of mercury emissions from utilities and other industries was due in four years.

Kayla asked why congress had given utility companies a reprieve; it didn't seem to make sense if they were significant sources of the listed air toxins. Ravi surmised that congress may have been more lenient with utility companies under Title III since they were already primary targets of regulation under Titles I and IV of the 1990 amendments. Both John and Kayla were surprised that emissions of heavy metals such as mercury, arsenic and lead had never been regulated within the utility industry. Noting that some individual states did limit power plant emissions of heavy metals such as mercury, Ravi agreed that it was surprising that so many air toxins from power plants had not been regulated at the federal level – at least until now.

Mercury, in particular, had received significant attention (e.g. EPA 1997, Center for Disease Control 1999, Physicians for Social Responsibility 2004). As explained in the 1997 EPA report, mercury released by industrial sources into the air can circulate in the atmosphere for up to a year and can be deposited on land and water thousands of miles from the original source. When heavy metal mercury is consumed by living organisms, it is converted to bioaccumulative methyl-mercury which becomes more concentrated in organisms higher in the food chain. A fact sheet issued by the Physicians for Social Responsibility (2004) describes mercury as a "potent neurotoxin" that affects the functioning of the central nervous system and explains that most Americans are exposed to mercury through the consumption of fish – especially of higher food chain predatory fish like swordfish and tuna. In its 1997 report, the EPA had estimated annual emissions of mercury within the U.S. to be about

158 tons of which 87% came from waste and fossil fuel combustion. However, since waste combustion had been subject to earlier regulation, fossil fuel combustion (primarily coal) was now the dominant source of mercury emissions in the United States. Ravi summed up the discussion by stating that two decades after the 1990 amendments, the 2011 MATR had listed electric utilities as a source category under section 112 of the CAA and that the long delay was a result of years of litigation between industry, non-governmental organizations, states and the EPA.

3. RESEARCH PURPOSE

The discussion then switched to the purpose of the research and the EPA data. Ravi explained that in order to gather the data needed to set the standards, the EPA had issued a two-phase information collection request in 2009 (EPA 2009). In the first phase, electric generating units (EGUs) subject to the new regulation completed a twenty-five page paper survey providing the most recent twelve months of emissions test and fuel analysis data since 2005 as well as data about plant equipment (e.g. boiler characteristics, pollution controls) and permitting requirements. In the second phase, the EPA selected EGUs who were believed to be the best performing units within specified pollutant categories. These EGUs were required to conduct emissions stack testing to measure flue gas entering the atmosphere and to conduct analyses of fuel used during testing. The cost of data collection and quality assurance was estimated to be about $10 million and the cost of stack and fuel testing was estimated to be about $66 million (EPA 2009). In order to leverage this investment, the Institute wanted to gain as much knowledge as possible from the EPA data which was made available to the public in the form of two MS Access databases -- one for each collection phase. They would start with the data from the first collection phase. Ravi was sure that this task alone would be very challenging. After they had mastered the Phase I database, they would consider integration of Phase II data. Links to the original data and descriptive information are provided in Table 1.

The purpose of the current project is to determine which combinations of equipment provide the best overall control of multiple pollutants. Certain boilers can remove pollutants during combustion or while coal is being burned. For example, fluidized bed boilers float and tumble burning coal on upward jets of air. The tumbling allows solids such as limestone to be mixed in with the coal and absorb pollutants such as sulfur dioxide. Other boilers are designed to burn coal at lower temperatures which inhibit the formation of nitrous oxides. In addition, different types of post combustion controls can remove pollutants from the flue gas before it is released into the air through the smokestack.

Kayla had one nagging question: What was the value of analyzing equipment that wasn't intended to control emissions of the newly regulated pollutants? Ravi explained that the EPA (2011a) had argued that the new standards were based on "existing, commercially proven technologies that are...frequently used in this industry such as electrostatic precipitators, fabric filters (bag houses), flue gas desulfurization (scrubbers) or dry sorbent injection." In other words, equipment used to control sulfur dioxide, nitrous oxides, particulate matter also controlled emissions of the newly regulated pollutants – at least according to the EPA. Indeed as Ravi pointed out, the EPA was using particulate matter as a proxy for all non-mercury metals. "So does this mean, the newly regulated pollutants were – in effect -- already being regulated" Kalya asked? Ravi wasn't so sure stating that "these are the kinds of questions we need to answer with our research" and that "controls for different pollutants may interact in ways that do not simultaneously reduce all regulated pollutants".

4. A DATA NARRATIVE

John had spent the last week studying the Phase I EPA database and was meeting with Kayla to give her an overview of what he had learned so far. He also wanted to get a better understanding of what data and what format would be required to conduct statistical analyses. Referring to the EPA database diagram, John convinced Kayla that the EPA Phase I was complex involving many dimensions. It contained forty different tables which were linked together by almost as many relationships. He showed her a sketch (Figure 1) of the data entity relationships which he had created based on the EPA database diagram.

In order to get a better understanding of the content of the database, John explained to Kayla that he had created a smaller "test" M.S. Access database by deleting some of the tables and fields from the first phase EPA database. He believed that the smaller database contained the most important data for their research project and that

the simplification would facilitate their preliminary analysis. All relationships were those created by the EPA and no records had been deleted from the remaining tables. A screen shot of MS Access relationships in the test database is shown in Figure 2.

John had many questions but would do his best to explain the Figure 2 diagram to Kayla. A facility, described in the `facility_information` table, is all the property, plant and equipment that resides at single geographic location and that has a legal owner. A configuration is a set of equipment components ordered by their physical location within the electricity generation process. A facility can have multiple configurations which are possibly operated concurrently at a given point in time or possibly which have changed over time due to the addition, modification or removal of particular equipment. Configurations are described in the `configuration_components` table.

Each configuration starts with one or more of what was labeled as a "unit". Each unit is in turn mapped to one or more boilers in the `unit_boilers` table and boilers are described in the `boiler_ information` table. John knew that the information in the `boiler_information` table would be important but he did not know what a "unit" represented. It seemed that the label "unit" was so generic that it could represent any kind of equipment. Question 15 of the EPA survey required "identification (or designation) of all coal- and oil-fired steam generating units (boilers) (as defined by Clean Air Act section 112(a)(8)) located at this facility" The question parenthetically indicates that a steam generating unit is a "boiler" and a footnote indicates that either a `Boiler ID` or a `Generator ID` can be provided:

> Boiler ID as reported on U.S. DOE/EIA Form EIA-860 (2007), "Annual Electric Generator Report", schedule 6, part A, line 1, page 53 OR on schedule 6, part B, line 1, page 54 OR Generator ID as reported on "U.S. DOE/EIA From EIA-923 (2008)"

John wondered whether allowing the interchangeable use of boiler and generator ids was a source of design problems. According to the language of the CAA, the EPA is required to regulate steam generating units and the CAA defines an "electric utility steam generating unit'' as "any fossil fuel fired combustion unit of more than 25 megawatts that serves a generator that produces electricity for sale". Based on this definition, the steam-generating unit is not the same as the generator that produces electricity. The former generates steam and the latter generates electricity. Like the term "unit", "generate" also had multiple meanings. Further adding to the confusion, the term "steam" was often use to describe a generator as indicated on EIA (Energy Information Administration) Form 860:

> Enter the identification (ID) code for each boiler that provides steam to each combustible-fuel steam generator … and for each combined cycle steam turbine generator. Boilers may be associated with multiple generators.

It was also apparent that there is a many-to many relationship between boilers and generators. In order to clarify the terminology, John conducted some research and settled on the following definitions:

> A boiler is a vessel which burns fuel to boil water and create expanding, pressurized steam which is transferred to at least one turbine. The thermal energy will be converted into rotating kinetic energy.
>
> A turbine is a rotor with blades that is connected to the shaft of a generator. It uses rotary motion to convert kinetic to mechanical energy.
>
> A generator is copper wire coiled around a shaft that is surrounded by a giant magnet. When the shaft is rotated, electric current is created on the wire, converting mechanical energy to electrical energy.

He was confident that these definitions provided much needed semantic clarity. And he had also discovered that the qualifier "steam" was used to distinguish the type of turbine which in addition to steam included water, wind and gas types.

Based on the survey instructions, `unit_id` was possibly meant to refer to a generator but he was still unsure. In the entire phase I database, there was NO additional information stored about units beyond the id itself. He was puzzled why the configuration table included units and not boilers.

CAA regulatory rules apply to boilers and not generators. This issue required further research; he had a nagging concern that it would be a cause of problems for their research.

Each configuration also has at least one chimney – called a stack – where gas exits the process. One or more pollution control devices may be installed after the unit and before a stack. The database contained four major groups of such post-combustion controls devices including particulate matter (PM) controls, nitrous oxide (NOx) controls, sulfur dioxide (SO2) and other controls. The "other" category contained mercury (Hg) control devices and Kayla and John agreed these would need to be separated. Control devices which are relatively independent (e.g. can be removed and relocated within a configuration or installed within another configuration) are referred to as "facility" controls and are described in the `facility_controls` table. In addition, boilers have design features to control NOx pollution which are described in the `boiler_nox_control` table. Air is sampled through ducts called sampling ports which can be placed at different locations within the process as well as at the exhaust stack. John noted that only pollution controls that were located *upstream* of (e.g. before) the sampling location should be associated with pollutant measurements at that location.

In the survey, utilities provided historical emissions data in the form of test reports. Each test report often corresponded to a compliance reporting requirement and each report in turn consists of multiple sampling runs where measurement devices collect and analyze samples of air during a discrete period of time. Multiple sampling runs might be used to ensure that measurements reflect steady state conditions of the electricity generation process. Each sampling run is in turn associated with one or more pollutants for which emissions are reported. The database contained emissions data for 106 different pollutants -- although many of these were infrequently reported. Kayla and John decided to focus only on the following pollutants: filterable particulate matter, sulfur dioxide (SO2), nitrogen oxide (NOx), total mercury (Hgt) and hydrogen chloride (HC1). John was initially confused about which type of mercury he should use but he had verified that total mercury is the sum of elemental mercury (Hg0), particulate bound mercury (Hgp) and oxidized mercury (Hg++). So for now, they would extract only Hgt. To further complicate matters, emissions were reported using different units of measurements including emissions rates (e.g. weight emitted per time period), emission factors (weight per heating fuel content) and concentrations (parts per air volume). These units of measurement are interdependent in that one may be derived from others given additional data. Kayla had done some initial investigation on converting emissions to a common unit of measurement and found that it was not straightforward. There were multiple conversion formulas which each made different assumptions and required different additional data. So to begin their analysis, Kayla and John agreed to use only sampling runs which reported emissions factors as pounds per million British Thermal Unit (lb/MMBtu) since this was the most frequent unit of measurement in the `sampling_run_ pollutants` table. Emissions data is contained in `test_reports`, `sampling_runs` and `sampling_run_pollutants` tables.

5. THE EXTRACT

John wanted to know what data format would be required for statistical analysis. Kayla explained that the typical input for statistical software is a two dimensional file or table. Each row represents an observation and each column represents a variable. John referred to this type of input as a flat "denormalized" table. Kayla continued explaining that it is usual in statistical analyses that some variables are dependent (those to be predicted or explained) and others are independent (those that form the basis for explanation or prediction). Computers scientists might more easily understand dependent and independent variables as output and input variables. In the current project, dependent variables are pollutant emissions and independent variables are boiler characteristics and control equipment.

They needed to determine the unit of analysis or observation and tentatively decided to define the observation as a unique combination of boiler characteristics and pollution controls at a particular facility. They would average pollutant emissions to this level of analysis. John suspected there might be some situations where emissions measurements could not be unambiguously associated with unique equipment and in these cases the emissions data should be excluded from the analysis. It was also important that a single emissions measurement was not averaged into multiple observations since this

would bias results by weighting some measurements more heavily than others. Also, multiple configurations of identical equipment at a specific facility should be merged into a single observation.

The discussion switched to data types – which was more straightforward than level of aggregation. Kayla suggested coding `boiler_firing_type` as a categorical data type with the following possible values: tangential, wall, cyclone, fluidized bed, integrated gas combustion cycle (IGCC) and other. Although not all statistical procedures handled categorical data types, initially she would conduct descriptive analyses by `boiler_firing_type`. The mapping of specific boiler firing type values which exist in the EPA database to the extract categories is shown equipment classification hierarchy shown in Table 3. For example, "front wall", "rear wall", "opposed wall", and "other" boiler firing types should be mapped to "wall" firing type.

Since a configuration can have a varying number of controls within a single category, Kayla suggested coding facility and boiler pollution controls as Boolean data types with 1 indicating presence of the control and 0 otherwise. Kayla and John drafted a preliminary structure for a data extract shown in Table 2. Like boiler firing types, the equipment classification hierarchy in Table 3 maps specific controls to the general controls in the extract file.

6. PRELIMINARY DATA ANALYSIS

John had some concerns about possible data anomalies which would affect the integrity of the data used for their research project. Their goal was to unambiguously relate emissions measurements to boiler characteristics and control equipment that was operational at the time of the test. He recognized that parts of the database did not meet normalization principles and some referential integrity constraints were missing. He came up with a plan to systematically investigate these issues. First, he would manually try to create extract records for some sample facilities. He had successfully done this for facility 663. MS-Access screenshots and the extract records for this facility are shown in figures 3 and 4 respectively. He had identified five additional facilities which he thought might present problems and would manually try to construct extract data for these facilities. The identifiers for the test facilities are: 56, 898, 1073, 1507 and 2324. For example, a potential problem for facility 1073 is that units 1 and 2 were each mapped to four boilers (1-4) in the `unit_boilers` table. He was concerned that the boilers would have different characteristics and had begun researching this plant using data at the Energy Information Administration web site. He had learned that in fact only boilers 1 and 2 should both be mapped to units 1 and 2. Second, in the process of creating extract data for the five facilities, he would make a list of problems in terms of relating emissions to equipment data. Third, he would design a new database to overcome any problems and input the data for the five facilities as a means of testing the new design. He hoped that this would demonstrate the viability of reformatting and importing all of the EPA data into the new design. He knew that the data would be used by the Institute for years to come and he was concerned that researchers would again and again need to deal with data anomalies for each analysis. Undoubtedly, assumptions would need to be made to resolve certain data ambiguities but at least they would be made explicit and uniformly applied to all future analyses.

7. REFERENCES

Center for Disease Control (1999). Public Heath Statement Mercury, Agency for Toxic Substances and Disease Registry. CAS#: 7439-97-6, March 1999.

EPA (2011a). Fact sheet: mercury and air toxics standards for power plants. Retrieved March 7, 2014 from http://www.epa.gov/mats/pdfs/20111221MATSsummaryfs.pdf

EPA (2011b). National emission standards for hazardous air pollutants from coal- and oil-fired electric utility steam generating units and standards of performance for fossil-fuel-fired electric utility, industrial-commercial-institutional, and small industrial-commercial-institutional steam generating units. 1 -1117. Retrieved March 7, 2014 from http://www.epa.gov/mats/pdfs/20111216MATSfinal.pdf

EPA (2013a). Overview – The Clean Air Act Amendments of 1990. Retrieved March 7, 2014 from http://epa.gov/air/caa/caaa_overview.html

EPA (2013b). 112(n) - Studies: Utilities, Coke Ovens, POTWs, Mercury, Hydrogen Sulfide.

Retrieved March 7, 2014 from http://www.epa.gov/ttnatw01/112n/112npg.html

EPA (2013c). Overview by Section of CAA: Introduction to CAA and Section 112 (Air Toxics). Retrieved March 7, 2014 from http://www.epa.gov/ttnatw01/overview.html

EPA (1997). Mercury Study Report to Congress. *Executive Summary*. Retrieved March 7, 2014 from http://www.epa.gov/ttn/atw/112nmerc/volume1.pdf

EPA (2009). ICR Final Part A. Retrieved March 7, 2014 from http://www.epa.gov/ttn/atw/utility/g1/eu_mact_icr_part_a.pdf

Lemonick, M. D.(1990, November 5). Forecast: Clearer Skies: The Revised Clean Air Act is Costly but well worth the price. *Time*, *136(20)*,33.

Neville, A. (2012). Debate heats up over new mercury and air toxics rule. *Power Magazine.* Retrieved March 7, 2014 from http://www.powermag.com/debate-heats-up-over-new-mercury-and-air-toxics-rule/

Physicians for Social Responsibility. (2004). Mercury in Fish, Fact Sheet #3. Retrieved March 7, 2014 from http://action.psr.org/site/DocServer/Mercury_Fact_Sheet__1.pdf?docID=710

Note: Teaching Notes and Case Supplements are available by contacting the authors

Appendix

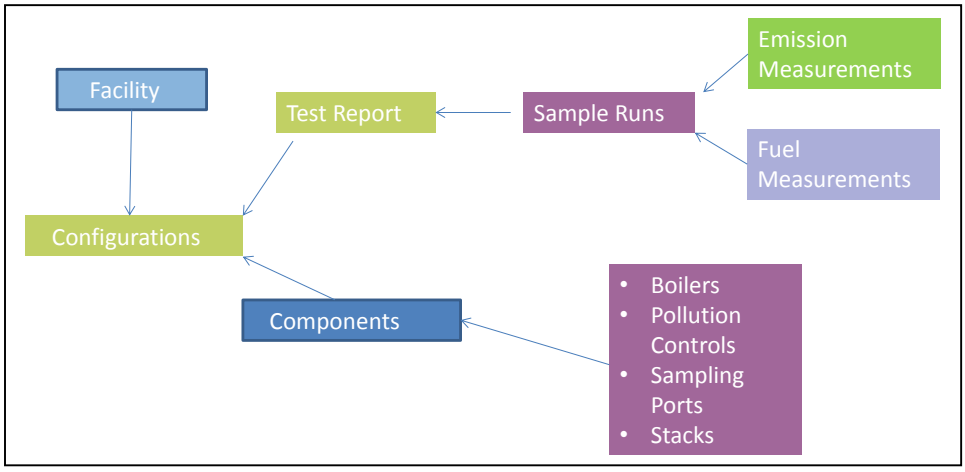


Figure 1 – Sketch of Entity Relationships

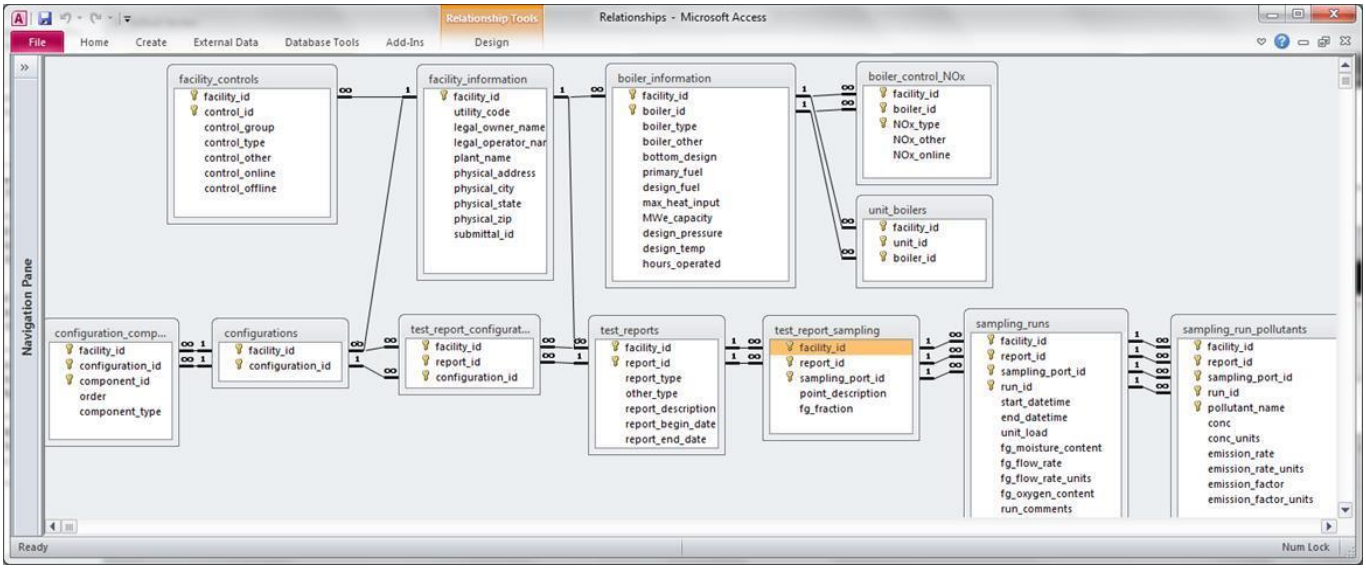


Figure 2 – M.S. Access Relationships in Test Database

Survey (see enclosure 1)	www.epa.gov/ttn/atw/utility/g1/eu_mact_icr_part_b.pdf
Data dictionary	www.epa.gov/ttn/atw/utility/pro/eu_mact_icr_part-i_ii-data_dictionary.pdf
Data Diagram	www.epa.gov/ttn/atw/utility/pro/eu_mact_icr_part-i_ii-db_erd.pdf
MS Access Database	www.epa.gov/ttn/atw/utility/eu_icr_parti_partii.mdb
Other Related Links	www.epa.gov/ttn/atw/utility/utilitypg.html

Table 1 - Links to Original Data and Descriptive Information

Potential Identifiers

1. **Facililty_ID**
2. **Configuration_ID**
3. **Boiler_ID**
4. **Unit_ID**
5. **Sampling_Port_ID**

Boiler characteristics

6. **Boiler_Firing_Type**
7. **Boiler_MaxHeatInput**
8. **MWe_Capacity**
9. **Primary_Fuel**

NOx Boiler Controls

10. **LoNox_Burner**
11. **Ovr_Fire** (Over air fire)
12. **Other_BoilerNOx**

NOx Facility Controls

13. **SCR** (selective catalytic reduction
14. **SNCR** (selective noncatalytic reduction)
15. **Other_Nox**

Mercury Facility Controls

16. **ACI** (activated carbon injection)
17. **DSI** (dry sorbent injection)

PM Facility Controls

18. **ESP** (Electrostatic precipitator)
19. **PM_Filter**
20. **PM_Scrubber**
21. **PM_Cyclone**
22. **PM_Other** (all other PM)

SO2 Facility Controls

23. **Wet_Fgd** (Wet Flue Gas Desulfurization)
24. **Dry_Fgd** (Dry Flue Gas Desulfurization)

Pollutant Emissions

25. **PM_F** (PM - Filterable)
26. **SO2** (Sulfur Dioxide - SO2)
27. **NOx** (Nitrogen Oxide - NOx)
28. **Hgt** (Total Mercury Hgt)
29. **HC1** – (Hydrogen Chloride HCl)

Table 2 – Structure of Extract

Boiler Firing Types

1. Tangential Firing
2. Wall Firing
   - 2.1. Front Wall Firing
   - 2.2. Rear Wall Firing
   - 2.3. Opposed Wall Firing
   - 2.4. Other Wall Firing
3. Cyclone Firing
4. Fluidized Bed Firing
5. Stoker Firing
   - 5.1. Stoker Underfeed
   - 5.2. Stoker Overfeed
   - 5.3. Stoker Spreader
   - 5.4. Stoker Other
6. Integrated Gas Combustion Cycle (IGCC)
7. Other Boiler Firing Type

Pollution Control Types

1. Particulate Matter (PM) Controls
   - 1.1. Electrostatic Precipitator (ESP)
     - 1.1.1. Cold Side ESP with Flue Gas Conditioning
     - 1.1.2. Cold Side ESP without Flue Gas Conditioning
     - 1.1.3. Hot Side ESP with Flue Gas Conditioning
     - 1.1.4. Hot Side ESP without Flue Gas Conditioning
   - 1.2. PM Filter
     - 1.2.1. Pulse Filter
     - 1.2.2. Reverse Air Filter
     - 1.2.3. Shake and Deflate Filter
   - 1.3. PM Scrubber
     - 1.3.1. Syngas
     - 1.3.2. Wet
     - 1.3.3. Venturi
   - 1.4. PM Cyclone
     - 1.4.1. Multiple Cyclone
     - 1.4.2. Single Cyclone
   - 1.5. PM other
2. Nitrous Oxide (NOx) Controls
   - 2.1. Facility Nox Controls
     - 2.1.1. Selective Catalytic Reduction
     - 2.1.2. Selective Non-Catalytic Reduction
     - 2.1.3. Facility Nox Other
   - 2.2. Boiler NOx Controls
     - 2.2.1. Boiler Nox Controls
     - 2.2.2. Low NOx Burner
     - 2.2.3. Overair fire (including advanced)
     - 2.2.4. Boiler NOx Other
3. Sulfur Dioxide(SO2) Controls
   - 3.1. Wet Flue Gas Desulfurization (WFGD)
     - 3.1.1. Wet FGD – Disk
     - 3.1.2. Wet FGD Flooded Disk
     - 3.1.3. Wet FGD Jet Bubbling Reactor
     - 3.1.4. Wet FGD Spray
     - 3.1.5. Wet FGD Tray
     - 3.1.6. Wet FGD Spray and Tray
     - 3.1.7. Wet FGD Other
   - 3.2. Dry Flue Gas Desulfurization (DFGD)
     - 3.2.1. Dry FGD Sorbent Injection
     - 3.2.2. Dry FGD Spray
     - 3.2.3. Dry FGD Other
4. Mercury Controls
   - 4.1. Activated Carbon Injection
   - 4.2. Dry Sorbent Injection
   - 4.3. Other Facility Controls
   - 4.4. Boiler Controls

Table 3 - Equipment Classification Hierarchy

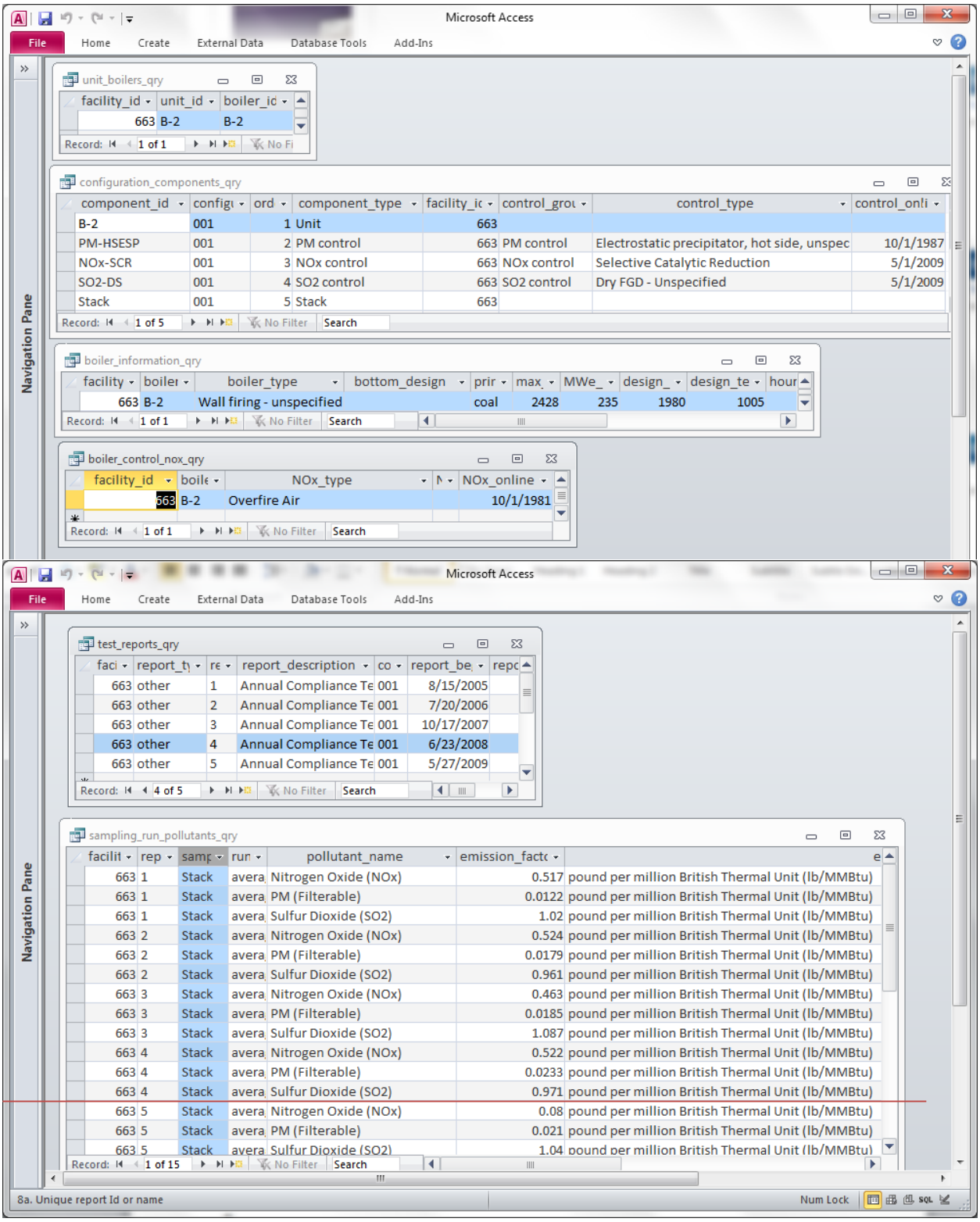


Figure 3 – MS Access Screen Shots for Facility 663

Extract Records[1]			
Potential Identifiers			
1. Facililty_ID	663	663	663
2. Configuration_ID	1	1	1a
3. Boiler_ID	B-2	B-2	B-2
4. Unit_ID	B-2	B-2	B-2
5. Sampling_Port_ID	Stack	Stack	Stack
Boiler characteristics			
6. **Boiler_Firing_Type**	Wall-firing	Wall-firing	Wall-firing
7. Boiler_MaxHeatInput	2428	2428	2428
8. MWe_Capacity	235	235	235
9. Primary_Fuel	coal	coal	coal
NOx Boiler Controls			
10. LoNox_Burner	0	0	0
11. **Ovr_Fire** (Over air fire)	1	1	1
12. **Other_BoilerNOx**	0	0	0
NOx Facility Controls			
13. **SCR** (selective catalytic reduction	0	0	1
14. **SNCR** (selective noncatalytic reduction)	0	0	0
15. **Other_Nox**	0	0	0
Mercury Facility Controls			
16. **ACI** (activated carbon injection)	0	0	0
17. **DSI** (dry sorbent injection)	0	0	0
PM Facility Controls			
1. **ESP** (Electrostatic precipitator)	0	1	1
2. **PM_Filter**	0	0	0
3. **PM_Scrubber**	0	0	0
4. **PM_Cyclone**	0	0	0
5. **PM_Other** (all other PM)	0	0	0
SO2 Facility Controls			
6. **Wet_Fgd** (Wet Flue Gas Desulfurization)	0	0	0
7. **Dry_Fgd** (Dry Flue Gas Desulfurization)	0	0	1
Pollutant Emissions[2]			
8. **PM_F** (PM - Filterable)	NULL	0.017975	0.021
9. **SO2** (Sulfur Dioxide - SO2)	NULL	1.00975	1.04
10. **NOx** (Nitrogen Oxide - NOx)	NULL	0.5065	0.08
11. **Hgt** (Total Mercury Hgt)	NULL	NULL	NULL
12. HC1 – (Hydrogen Chloride HCl)	NULL	NULL	NULL
Applicable Dates[3]	10/81-9/87	10/87-4/09	5/09-present
1) Configurations shown here in columns would be in rows in the actual statistical extract.			
2) The second configuration emissions are averages of first four reports 2005 - 2008. Only test report 5 emissions from 2009 should be associated with the third configuration.			
3) 10/81 OverAir Fire Control installed. 10/87 ESP installed. 5/09 SCR and FGD installed.			

Figure 4 Sample Extract Records for Facility 663

Teaching Case

Employing two factor authentication mechanisms: A Case Study

Cameron Lawrence
Cameron.Lawrence@business.umt.edu
School of Business Administration
The University of Montana

Eric Fulton
eric@subsectorsolutions.com
Subsector Solutions

Gerald Evans
Jerry.Evans@business.umt.edu

David Firth
David.Firth@business.umt.edu

School of Business Administration
The University of Montana

Abstract

This case study examines the life of a digital native who has her online accounts hacked, passwords reset, and is locked out of important online resources including her university email account and Facebook. Part one of the case study examines how the hack was perpetrated and the fallout of losing control of one's digital identity. Part two of the case study details how the main character recovered her accounts, simultaneously providing readers with the tools necessary to protect their own digital identities. Specifically, this case focuses on the use of two-step authentication schemes along with the generation, use and management of complex passwords. We then provide a set of discussion questions along with suggested lab activities that will show students how to implement the technologies discussed in the case. This case is intended to be used at both the undergraduate and graduate levels. This case complements the model curriculum objectives in IS 2010.1 and IS 2010.7.

Keywords: Information Security, Security, Privacy, Hacked, Two-Factor Authentication, Digital Native

1. Introduction

Kim West was born with newspapers in her blood. Her father was an acclaimed journalist who had authored a number of high-profile editorial pieces on government corruption, and Kim wanted nothing more than to follow in her father's footsteps. During her senior year in high school, Kim was admitted to a prestigious university with a top-five journalism program. As soon as Kim arrived on campus, she started working with the school paper. She began on the lowest rung of the paper's staff and started slowly working her way up. She received her big break during her junior year when she uncovered a story about how the student body president, Ryan Scott, had used student funds to pay for personal expenses and parties for his close friends. The article, along with subsequent stories, caused so much damage that the student senate impeached Ryan and forced him to resign after the winter break. Given the success of the story, West was hopeful she would win the coveted position as editor of the university newspaper, which would be announced at the end of the spring semester; however, there was no way for her to know that her breakthrough journalism work was about to lead to her digital life being turned upside-down.

2. The Take Down

Kim nestled into her favorite corner at the campus coffee shop to get some work done before her Advanced Reporting seminar. She had an hour to burn and wanted to put the finishing touches on an article for the school newspaper, which would be published the following day. As she had done hundreds of times during her college career, she powered up her Mac laptop seamlessly connecting to the coffee shop's wireless network, and logged into her university email account. This also gave her access to Google Docs, which she used to do all of her writing. A few years ago the university switched over to a service provided by Google that gave every student access to Google Apps, and she couldn't imagine how she worked without these tools. At that very moment, an anonymous student sitting in the same coffee shop was busy at work.

F4T4L (pronounced "fatal") occasionally hacked for money. His student loan debt was high while his interpersonal skills are low. The hacking world was a place where he can make extra money and get the support he hadn't found anywhere else. In hacking he found a community that accepted him and valued his skills. He tried to work a regular day job, but those didn't last very long; the low-level positions didn't pay nearly enough to support him while he was in college. People he knew occasionally approached him to "hack" someone they knew, but usually they couldn't afford the rates he charged. His services came at a high price because, if he was going to do something for which he might get caught, it was going to be worth his effort. It had been a while since anyone had approached F4T4L for his skills, which was why he was surprised when he received a message from the former student body president Ryan Scott. F4T4L had gotten to know Ryan in a freshman-level computer science class that all students were required to take. Ryan struggled with the class and came to rely on the quiet student, F4T4L, who sat next to him. It was clear to Ryan after a few weeks that this kid knew much more than the instructor and was a whiz with technology. Over the course of the semester, Ryan was surprised to learn that the quiet, low-key, computer whiz was proud to call himself a "hacker."

F4T4L had not seen Ryan much since class their freshman year. It didn't hurt his feelings that Ryan didn't keep in touch. F4T4L was used to making friends who needed his help, and once they didn't need help, disappeared from his life. Ryan never fully dropped F4T4L, but due to his popularity and involvement with student government, Ryan just didn't have time for many people in his life. Thus, F4T4L's curiosity was more than piqued when Ryan asked him to coffee seemingly out of the blue, prompting him to accept the invitation and meet Ryan the next day.

F4T4L didn't pay much attention to the politics on campus, but Ryan's impeachment was such a big story that it was almost impossible to be on campus and not hear about it. Ryan was not the forgiving type. Over the course of their conversation, F4T4L learned that Ryan wanted revenge on the reporter who had dealt a crushing blow to his life. He wanted F4TL to take down her email account, her Facebook account, and anything else F4T4L could access. Ryan wanted F4T4L to post embarrassing material to Kim's Facebook page. In addition, he was instructed to send offensive messages to all of her contacts after which he was to delete her saved messages and along with anything else he could get his hands on. Ryan wanted to humiliate Kim just as he had been embarrassed. They ultimately agreed on a price of $1,000.00, with $500.00 up-front and the last $500.00 after the job was done.

F4T4L went to work and was confident Kim would not know what hit her.

Kim went to the same coffee shop at a little after 2:00 PM every Tuesday and Thursday. As Ryan and F4T4L met on a Tuesday, F4T4L planned his attack for Thursday. He arrived 15 minutes before Kim and powered up a sticker-covered ThinkPad X1 Carbon, which was connected to an Alpha wireless card with a 10db antenna booster. After his laptop booted, F4T4L's fingers start to fly. There was no way for the coffee shop patrons to know a master was in their midst, quietly plying his craft. This thought always gave F4T4L a lot of personal satisfaction and was one of his favorite parts of hacking. He booted up a number of tools, like ARPSpoof, SSLSniff, and Ettercap. Each tool would play a part in the interception and decryption of Kim's password. Actually, every user of the coffee shop was being hacked, but he was only interested in one: Kim West, student reporter extraordinaire. Like clockwork, Kim entered the coffee shop and settled in to get some work done. He quickly identified her machine on the network and applied a filter that just displayed the traffic from her computer.

"That didn't take long," F4T4L mumbled to himself as he took a sip of coffee. A few seconds later "Bingo!" floated off of F4T4L's lips as he saw Kim's credentials fly by on his screen while he idly spun a pen in his right hand. He continued to watch, waiting, hoping to see more passwords appear on his screen. After an hour F4T4L saw Kim stand up to leave. "*Well, I guess my work is done*," thought F4T4L as he stopped intercepting the coffee shop's network traffic. At this point he was partially successful because he had only secured the credentials for Kim's email account. He did know, however, that there was a very good chance that Kim, like most users, daisy-chained her accounts and that the credentials used to login to her email account were probably the same as those used to login to other services such as Facebook. He opened a new tab in his web browser and browsed to Facebook. He then used the same email address and password he had collected earlier. A split second after he hit the enter key he was in. "I can't believe these people are so stupid..." F4T4L mumbled. Once he obtained the credentials and logged in successfully he knew the rest would be easy. F4T4L logged out of Facebook, closed his laptop bag, and stood, beginning the walk home with a smile on his face. Later that evening it took less than 20 minutes to accomplish what came next.

2. The Response

Every writer develops habits around writing; Kim is no exception. Early on she learned that she did her best writing early in the day and saved the evenings for editing her work. Following her usual evening ritual, Kim sat in her favorite chair and booted up her laptop as her best friend and roommate, Jenn Humphries, watched her latest obsession, *Downton Abbey*. At first Kim was annoyed since she couldn't login to her email account. Her fingers were cold, and she was trying to type too fast. She slowed down and typed her username and password carefully, annoyed at having to type so slowly. She pressed enter and saw a "Password or Username Incorrect" page. She tried typing her password in again, even slower this time. She received the same result, a "Password or Username Incorrect" page stared mockingly back at her. She knew she had typed the right password and was still denied. Panic started to set in. "*This has got to be a mistake*," thought Kim as she repeatedly tried to login to her account.

"No, no, no, no!" she cried frantically. From across the room she shouted, "Jenn, stop messing with me."

"I'm not doing anything Kim, can't you tell I am watching *Downton Abby*?" retorted Jenn.

"No seriously, if you aren't messing with me, I need help right now!" responded Kim in a nervous voice.

Jenn was always surprised that someone as smart as Kim didn't know anything about the technology she was so dependent upon. Jenn was majoring in Management Information Systems and was Kim's go-to tech person. Jenn had played little jokes on Kim in the past such as setting up the "Upside-Down-Ternet" on their network turning Jenn's Internet upside down, but Jenn knew she hadn't done anything recently. Moving over to Kim and her laptop, Jenn asked, "What's up?"

At about that time Kim's mobile phone lit up. She answered a call from an old friend from high school who told her that there was some wild stuff going up on her Facebook page and she wanted to make sure Kim was OK. Kim quickly assured her that she was fine, but it was clear that something terrible was going on.

Jenn quickly determined that Kim was being attacked. They went through the password reset

process on her university email account, but that didn't work as the backup email address had been changed. It was clear to Jenn that whoever was behind this attack knew what they were doing. Fortunately, Jenn worked part time on the university's help desk and knew the campus IT administrators. In fact, she played on the same ultimate Frisbee team as the lead administrator and had his mobile number. She called and explained the situation. He was able to remotely access Kim's account and clearly see there had been suspicious activity. They noticed a password change and a few other account configuration changes that happened almost simultaneously a few hours previously. In addition, they could see a mass email containing an offensive message was sent out to all of Kim's contacts, which included high ranking university officials. For Kim however, the worst was yet to come. After the message was sent, the attacker deleted all of her saved email messages, her entire address book, and all of her Google Drive documents. Effectively, her digital life was just deleted.

The IT admin could easily see that the account had been compromised. Fortunately for Kim the university had adopted the Google apps for education platform. This means Google hosts the university's email system and provides cloud-based productivity tools, ranging from word processing and spreadsheet applications to a private, cloud-based hard drive for every student on campus. The system had been successfully deployed two years ago and it changed how many students, faculty and staff work on campus. In addition to the email and productivity applications, the Google education platform also provides administrators with powerful tools to manage the technology, including the ability to recover data in the event an account has been compromised. He told Kim not to worry as they could recover her data and have it available in the next 30 minutes. Upon hearing this news Kim sighed in relief. Losing control of a Facebook account was bad, but losing all of her documents and emails was catastrophic. She was happy to have her documents back, although she was still frustrated at losing her Facebook. She silently added a mental note to buy a nice thank-you gift for the IT admin who saved her digital life.

The next challenge was to recover Kim's Facebook account. This was a little more difficult because they did not have a personal relationship with someone at Facebook to help recover the account. Visiting Facebook's help page, they realized account hacking is a fairly common problem on Facebook. Thankfully, Facebook has processes in place that enable legitimate account holders to recover their accounts in the event of profile hijacking. Jenn and Kim initiated the account-recovery process and by the next morning had regained control of Kim's account. It was easy for the Facebook staff to conclude the account had been compromised. First, there was the obvious clue of the inappropriate and obscene posts, but there were also other clues related to how the attacker accessed Facebook. The attacker had been very careful to cover his or her tracks so that security professionals could not track the perpetrator down. The attacker had used something called "Tor," a software add-on used to maintain anonymity online.

Kim was lucky. She had lost access to her accounts in the early evening and had recovered everything by the next morning. Over a late breakfast, Jenn and Kim pieced together what had happened the previous day. It was clear that Kim lost control of her university email and Facebook account at almost the same time. The last time she accessed her email account was at her favorite coffee shop, which must have been where it was stolen. She's had coffee there hundreds of times and written numerous newspaper articles in that coffee shop without having her accounts stolen. How and why did it happen this time?

Slowly, Jenn put it all together. There must have been someone in the coffee shop connected to the same Wi-Fi network who was monitoring Kim's network activity. Jenn had heard of students playing around with powerful Wi-Fi antennas and applications that intercept and monitor network traffic, but she had not heard of a specific student being targeted in such a vicious manner. When Jenn shared her hypothesis on what had happened, Kim didn't agree. Kim said she wasn't logged in to Facebook and was only briefly catching up on email while she was in the coffee shop.

"Let me guess," Jenn said, "you use the same password for all of your accounts, don't you?"

"Of course I do. Doesn't everyone? There is no way anyone can remember unique passwords for every site we log into," Kim responded.

"No, not everyone does!" said Jenn. "What you did is called "daisy-chaining", which means your various online accounts are linked together by a common email address and password. This

makes it MUCH easier for someone to hack into your account. I am going to help you configure your email and Facebook accounts so this will never happen again. The first thing we are going to do is set up two-step authentication on your accounts, then I am going to introduce you to an application called LastPass, which will help you manage all of your passwords. In under an hour we can make sure this will never happen again. The best news is these tools are free."

Kim looked at Jenn and said, "I am happy you know what you are doing because this is all kind of going over my head!" making a whoosh noise along with the hand-over-head gesture.

Jenn's first task for Kim was to set up two-step authentication on her email and Facebook accounts. "What in the world is two-step authentication?" asked Kim.

"It really is simple," Jenn replied. "I learned about this in my MIS security class last semester and set it up as soon as I could." Jenn continued, "Actually, I am surprised more people aren't using features like this. By default, when you login to your email or Facebook account you are using single-factor authentication, which is your password. Get it? It's only using one single thing to identify who you are. So if someone obtains your email address and figures out your password they can gain access to your account. Since most people use the same email address and password for all of their online accounts, this poses a serious security threat. When you implement two-factor authentication, the website requests two things from you to verify your identity. Usually this is something you know, like a password, and something you have, like a random number token or authentication application installed on your phone. When you activate the powerful two step-step feature built into our university email accounts, it almost eliminates any chance of your email account being hacked."

"I know it sounds a bit confusing, but it really is simple. Here, you can see for yourself. Let's watch this short video that Google put together that illustrates the concept" said Jenn as she played the video on her iPhone for Kim.

Link to Video:
http://www.youtube.com/watch?v=zMabEyrtPRg

After watching the video and spending 15 minutes getting two-step verification setup on Kim's email account, the two young women turned to securing Kim's Facebook account. Now that Kim was familiar with the two-step verification concept it didn't take any time to set up her Facebook account with the same level of protection. Facebook calls this level of security "Login approvals," and it works in almost the same manner as the two-step verification process that Google uses. Now Kim has Facebook's "Login Approvals" configured. When she logs in from a computer she hasn't used before with her Facebook account, she is required to enter not only her username and password but also a unique code that is sent to her cellphone via text message.

Once Kim had two-step activated for her Gmail and Facebook accounts, Jenn introduced her to the Google Authenticator application for her iPhone. Using the app Kim went through a short setup process with Gmail and Facebook. For both apps she used the Authenticator application to take a picture of a QR code. As soon as the phone snaps the QR code, it adds a random number generator for the application that changes every minute. Now when Kim wants to log into her account from a new computer, she has to type her password and the currently displayed random number. Thus, even if her password were stolen, the attacker would still not have the random number required to log into the account.

Kim had learned a lot in the past hour about her online security and was getting a little tired, but as long as Jenn was teaching, Kim was going to keep learning. "Next up is LastPass. It is a browser add-on that allows you to manage all of the passwords you use in your online life. LastPass not only manages your passwords, but even enters them for you!" explained Jenn.

"So LastPass just saves all of my passwords and logs in for me? Why not just remember four or five passwords?" asked Kim.
"Well," said Jenn, "it's a bit more complicated. When you sign up for new websites or change existing passwords, you can have LastPass create a complex password and remember it for you. For example, this is a password I just had LastPass generate: **7XacK9V4eWtvbzQV88Fc**. While two-step authentication is the best approach to security, not all sites have that capability. Ideally, all of the sites you login to should be protected by separate and individual passwords like the one I just generated. Since there is no way we can remember numerous complex passwords, we need a tool such as LastPass to do it for us. The

really good news is that even if someone hacks into a service you use frequently such as Twitter and steals your password, it won't be a big problem for you because that password is only used on Twitter and can't be used to login to any of your other accounts." Having satisfied most of Kim's questions, Jenn helped Kim install the LastPass application and configure it for use.

"That's about everything I can teach you right now," said Jenn. "There's a lot more you could do to increase your security, but without extra research they would sound like a bunch of random acronyms to you. VPN, YubiKey, Whole Disk Encryption, Tor, VMWare and BSD Virtual Machines, HTTPS Everywhere: these all don't make any sense to you, do they? asked Jen.

"Not really," replied Kim.

"Well, don't worry about it. The few steps we have taken in the last hour have increased your security beyond what the vast majority of technology users employ and, more importantly, you have the essentials down. If you want to learn more about some of the things I just listed off we can meet up later and I can show you," said Jenn with a smile. Jenn continued hurriedly, "But finals are just around the corner, and I need to get to the library to study," and then she headed out the door.

With her roommate gone and her online life recovered, Kim sat and reflected on how vulnerable she had been and how little she knew about basic areas of security. What is more surprising is that she had grown up with technology and been using it since her earliest days in school. It stunned her to think that no one had ever talked with her about the importance of securing her online life. On the other hand, what troubled her the most was a simple question that would never be definitely answered, "Why me?" There were at least forty other students in the same coffee shop and all early indications are that she was the only one that was attacked. "It just doesn't make any sense," she muttered. Then it slowly dawned on her as she thought of people who might have a vendetta against her. "Ryan Scott's impeachment didn't seem to sit well with him. And I know he's never forgiven me for my article. It MUST be...That rotten...!"

3. CONCLUSION

Kim's senior year was shaping up to be everything she had hoped. She was chosen as the editor for the university's paper, which almost guaranteed her a job with a major news organization after graduation. Her new position went into effect at the end of the spring semester. She had all summer to get ready for the fall semester and the publishing of the first edition under her leadership. She spent most of her summer learning about the operational details of running the paper. She also spent a lot of time talking with the paper's IT staff, which is something she never would have done prior to her hacking experience. In fact, she put into place a new set of policies for securing technology at the student-run paper. She insisted that two-step verification be turned on for all paper staff and asked the IT staff to put on a training session prior to the start of the fall semester on how to use LastPass. In the same predicament that she was once in, the vast majority of her friends on the paper had never heard of two-step verification or LastPass.

One of the perks of Kim's new position was the weekly editor's column that allowed her to sound off on a topic of her choice. She had spent a good part of the summer thinking about her first column and was thrilled when it appeared in print. It was titled "A Digital Native is Hacked: My Story." It began, "In the span of 24 hours my digital life was turned upside-down. This is the story of how I was hacked and I am telling it publicly so it doesn't happen to you."

4. Questions and Student Lab

1) What steps are you taking to protect your digital life? Could you fall victim to a similar attack?
2) If someone were to get access to your email address and password, how many different sites could they log into?
3) How many different sites do you log into regularly? Do you use the same login information for any of these sites?
4) Setup Google two-factor authentication
    4a) If you have a smartphone, install and configure the Google Authenticator App
5) Setup Facebook two-step verification
    5a) If you have a smartphone, configure the Facebook Code Generator
6) Create a LastPass account and generate secure passwords.
7) What is your email recovery password? Is it an email account you still use?
8) Is someone had access to your email, what other online services could they compromise through password resets?
Bonus:

1) Pick one of the services listed by Jenn and research how it could provide additional security: VPN, YubiKey, Whole Disk Encryption, Tor, VMWare and BSD Virtual Machines, HTTPS Everywhere, Bitcoin.
2) Create a walkthrough for your fellow students on how to implement a technology listed in question one.
3) Create an online tutorial demonstrating the technologies featured in the case.

5. REFERENCES

ARP spoofing. (2013, June 14). In *Wikipedia, the free encyclopedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=ARP_spoofing&oldid=558983438

Eder, S., & Valentino-DeVries, J. (n.d.). A Spy-Gear Arms Race Transforms Modern Divorce. *Wall Street Journal*. Retrieved from http://online.wsj.com/article/SB10000872396390443995604578002751421246848.html?mod=WSJ_WhatTheyKnowPrivacy_LeftTopNews

Farhi, P. (2013, June 14). CBS confirms reporter Sharyl Attkisson's computer breached. *Washington Post*. Retrieved from http://articles.washingtonpost.com/2013-06-14/lifestyle/39967790_1_cbs-news-computer-intruder

Fowler, G. A. (2012, October 13). When the Most Personal Secrets Get Outed on Facebook. *Wall Street Journal*. Retrieved from http://online.wsj.com/article/SB10000872396390444165804578008740578200224.html

Green, T. (2013, June 15). CBS News Confirms Investigative Reporter Sharyl Attkisson's Computer Was Hacked. *International Business Times*. Retrieved from http://www.ibtimes.com/cbs-news-confirms-investigative-reporter-sharyl-attkissons-computer-was-hacked-1308429

Honan, M. (2012a, August 6). How Apple and Amazon Security Flaws Led to My Epic Hacking | Gadget Lab | Wired.com. *Wired Magazine*, (20.08). Retrieved from http://www.wired.com/gadgetlab/2012/08/apple-amazon-mat-honan-hacking/

Honan, M. (2012b, August 6). How I Got My Digital Life Back Again After An Epic Hacking | Gadget Lab | Wired.com. *Wired Magazine*. Retrieved from http://www.wired.com/gadgetlab/2012/08/mat-honan-data-recovery/

Johnson, D. (2013, June 7). Turn on 2-step verification to enhance security. *CBS News*. Retrieved from http://www.cbsnews.com/8301-505143_162-57587955/turn-on-2-step-verification-to-enhance-security/

Nicole Perlroth, & Bilton, N. (n.d.). Facebook Says Hackers Breached Its Computers. *New York Times*. Retrieved from http://bits.blogs.nytimes.com/2013/02/15/facebook-admits-it-was-hacked/

Perlroth, N. (2013, February 20). Some Victims of Online Hacking Edge Into the Light. *The New York Times*. Retrieved from http://www.nytimes.com/2013/02/21/technology/hacking-victims-edge-into-light.html

Reputation.com Reviews: Associated Press's Response to Getting Hacked. (n.d.). *Reputation.com*. Retrieved 7/5/2013 http://www.reputation.com/reputationwatch/reputationcom-reviews-associated-presss-response-getting-hacked

Segal, D. (2012, June 9). Hacked on Facebook and Seeking Help - the Haggler. *The New York Times*. Retrieved from http://www.nytimes.com/2012/06/10/your-money/hacked-on-facebook-and-seeking-help-the-haggler.html

Smith, H. A., & McKeen, J. (2011). The Identity Management Challenge. *Communications of the Association for Information Systems*, *28*(1). Retrieved from http://aisel.aisnet.org/cais/vol28/iss1/11

Zetter, K. (2013, January 1). New York Times Hacked Again, This Time Allegedly by Chinese | Threat Level | Wired.com. *Wired*, (21.01). Retrieved from http://www.wired.com/threatlevel/2013/01/new-york-times-hacked/

Editor's Note: *This paper was selected for inclusion in the journal as the ISECON 2013 Best Teaching Case*

Teaching Case

Data Storage Forensics – What is Really Left After I Hit the Delete Button, and How Can I Actually Make Sure It's Gone?

Anthony Serapiglia
Anthony.Serapiglia@stvincent.edu
CIS Department, St. Vincent College
Latrobe, PA 15650

Abstract

The following Teaching Case is designed to expose students to three scenarios related to data stored on hard drives, techniques that could be used to retrieve deleted or corrupted data, and a method for a more thorough deletion of data from a hard drive. These issues are often overlooked in current IT curriculum and in our age of digital clutter this can be a dangerous oversight leading to potential financial loss, exposure to identity theft, and criminal liability. This case study / lab exercise can be utilized in multiple levels of a CS/IS curriculum, adjusted to meet the skill and background levels of introductory courses to specialized capstone courses in hardware or security. It provides talking points to highlight the importance of being aware of the spreading digital footprint, and provides introductory exposure to available tools and techniques for advanced data recovery.

Keywords: Computer Forensics, Data Recovery, Disaster Recovery, Identity Theft, Digital Footprint, Hard Drive, Drive Imaging

1. Introduction

Our lives today are digital. News of Big Data and the amounts of information that are collected and stored abound till the numbers simply make us numb and are inconceivable of what they really mean. One report states that 90% of all of the data in the world has been generated in the past 2 years (Dragland, 2013). To survive today it is necessary to know how to cope with, work with, and become efficient with large amounts of data. It also mandates that just as with real trash, we know how to deal with our digital trash.

The term Digital Footprint is often used in relation to a person's online presence. However, it can be extended to include much more than traces of social media posts. Just as dumpster divers can gather much information about a person through their physical trash, so too can digital evidence be gathered through the physical collection of digital trash and physical storage mediums such as hard drives, flash drives, and SIM cards. Simply assuming a corrupted drive is inaccessible or that the drive has been re-formatted so everything must be deleted is a mistake. Possibly, you have left much behind.

A normal hard drive keeps an index of where a file is placed within its storage space. This is the same regardless of operating system (OS), MAC, Windows, or LINUX. The normal deletion process

is to simply remove the pointers from this index that tells the OS where the file is. Thus the OS no longer thinks the file is there. However, the file is, normally, still there on the storage medium and will remain so until it is overwritten by another file. Even when overwritten, a file is often split into various chunks and pieces and scattered across the available space. This can lead to partial files remaining available even if some sections have been overwritten. File carving is a practice of partial recovery of files and can still reveal very valuable material even if the full file is not available. Discarded hard drives and other storage devices have become a digital dumpster diver's dream.

It is important for anyone working in the IT industry to understand the extent of their digital footprint and to manage sensitive data securely. This includes being informed on how to properly delete and destroy data, as well as how to possibly recover accidentally deleted or corrupted data. An entire industry exists of companies that can perform these tasks at a cost, sometimes a very high cost. There does exist, however, a multitude of free and readily available tools that can allow anyone to perform data recover and secure data deletion quickly and easily

Included in this paper is a procedure for a lab exercise that will walk students through the process of collecting an image of a hard drive and investigating that image for any data it may contain. The lab includes the use of several software packages that are free, and can be adapted for several operating systems or other forensic tools. While secure data collection for evidentiary use will be discussed, it is not necessary for the scope of this exercise. An activity can be included to compare various deletion methods to highlight the need for care in data disposal. These exercises have been utilized in a non-major introductory course, a course on computer architecture and operating systems, and an advanced computer security course. The exercises and exploration questions can be adjusted and presented to be applicable for students with a wide range of computer and technological experience and expertise.

2. Three Background Descriptions

The following three scenario outlines have been provided as entry points to stimulate conversation, provide a personal attachment for students to relate to similar experiences in their own lives, and highlight the idea that the amount of places data resides is vast and often not very obvious.

Personal Data Loss

There is a common phenomenon that happens today in regard to the changing mediums of our lives. It centers on family gatherings. These could be happy occasions, such as weddings or reunions, or sadder occasions such as funerals. Whatever the occasion, there is often a presence when generations of a family come together that have not seen each other for extended periods of time – that presence is "the box" or "the albums"… pictures. Pictures in black and white or sepia tone, old Polaroid's, faded prints from the 1960's, 1970's, and 1980's populate the albums and boxes. They are passed around and handled with much care as they are precious artifacts of life. A funny thing happens with those boxes and albums, though. They tend to stop. They all of a sudden drop off around the year 2000. Many people have switched from film and prints to digital cameras and phones with cameras built in. Cost, storage, and convenience have combined to change our habits related to these memories.

A staple of the nightly news, unfortunately, is coverage of a house fire. In the background will be a scene littered with fire trucks, ladders, hoses, and firemen. A reporter will find a resident who is safe from harm, but distraught over the loss. "All of our memories are gone!" they will say. All of those albums and boxes of photographs burned and lost to the fire. With the change in medium today to digital files and storage in hard drives, the same loss can occur in an instant without the fire. The most common point of failure in a computer is the hard drive. The most common drives to fail are large capacity consumer models for home use. Thousands of digital images could be lost in an instant. There are no "film at eleven" news crews to cover this tragedy, but the loss is felt by the owner nonetheless.

Questions

How do you backup your files? Do you utilize cloud services such as Dropbox, Google Drive, or Microsoft SkyDrive? What are some recovery services available that will find data one equipment that has suffered fire/water/physical damage?

Digital Footprints in Unexpected Places

On April 10, 2010 CBS Evening News aired a 5 minute investigative reporting piece that provided just another piece of evidence that our digital

footprints are large and at the same time mostly hidden from our primary sight. The piece by reporter Armen Keteyian came during the 50th anniversary of the ubiquitous steady workhorse of office machinery – the copy machine (Keteyian, 2010). In the fifty intervening years the copy machine had undergone drastic evolution from a more manual piece of machinery such as mimeograph machines, to devices that are as complex and powerful as standalone personal computers. The reality is, since the early 2000's, almost every multifunction printer/scanner/copier (MFP, multi-function printer) is a standalone computer that has a processor, an operating system, and a hard drive as the basis for the machine. Because of their status as a utility device tucked away in a separate room or closet, the MFP is rarely seen by those that use it as a computer – it is just another background object that sometimes needs maintenance, sometimes needs refilling, but falls right in line with the water cooler in how much attention it really gets.

During the course of the investigative reporting piece by Mr. Keteyian, it was shown that a grave vulnerability lies within these multifunction printers – the hard drive. Most office workers do not know that the MFP they use so often operates in such a fashion that each copy is actually a scan that is saved to this hard drive so that the printer side of the device can output multiple hard copies without re-scanning every time. Unfortunately not only do most users not know that the image file is even there, most MFP's also keep this file on the hard drive until space is filled and the need to overwrite existing files is reached.

MFP's have become very expensive pieces of office equipment. Most companies, no matter what size, do not own their own opting rather to lease these machines from office supply companies or even the manufacturer directly. Often they are rotated out after three to five years. The lifespan of these MFP's can be ten to fifteen years or more with proper maintenance and care. The aftermarket for used machines is great with used machines being in demand for their cost savings over new models. Thus it is not uncommon to see a ten to fifteen year old copier that has been in four or more different offices for different companies over time.

The vulnerability that the CBS News investigative report highlighted was that in most cases, being unaware of the existence of the hard drive, MFP's that were recycled to other locations or sold to other office product suppliers rarely, if ever, cleared the hard drive of the accumulated images of documents that had been printed from it. In the report, 4 copiers were purchased from a used equipment warehouse for an average of $300 each. Through freely available forensics software, all four offered up a treasure trove of sensitive documents. One machine was from a sex crimes division of a metropolitan police force and held documents related to criminal cases. A second was found to be from another police department narcotics division that contained documents that included details of suspects in drug raids amongst other information. The third had been used in an architectural firm and it contained structural design plans for buildings in Manhattan, one a block from Ground Zero as well as 95 pages of Human Resource documents for payroll that included social security numbers and pay stubs. The final MFP had been placed with a medical insurance company and produced over 300 pages of private medical records including prescriptions, blood test results, and a cancer diagnosis.

It was revealed during the report that although some manufactures are aware of the issue of saved images on hard drives, many have done nothing about it. Some have made available software updates, at a cost as high as an extra $500. With the life span of older machines being extended in the aftermarket to well over 10 years, this is a problem that will not go away soon.

Questions

How many place does your data live? If you sent an e-mail with a spreadsheet attachment, itemize how many different places that spreadsheet can end up. How much data is stored on your cell phone? If you lost your phone, or if it was stolen – how much of that information is sensitive? Would you have a method of remote deletion?

Evidence

Statistics provided by the 2012 Norton Cybercrime Report shows the rising scale of consumer cybercrime. The numbers are staggering, showing 1.5 million victims daily averaging 18 victims per second. The global price tag of this crime is estimated at $110 billion US dollars annually (Symantec, 2012). Computer crime, in general, is a very vague topic. There are very few set definitions and boundaries within law enforcement. Multiple descriptions exist with multiple standards organizations offering up "best practices" guidelines on evidence collection and handling. The US department of Justice categorizes computer crime in three categories:

Where the computer is the target of the crime, where the computer is used as a weapon in commission of a crime, and where the computer is used as an accessory in commission of a crime (USDOJ, 2013). Considering these broad definitions of computer crime in combination with the broad definition of what a computer can be - one estimate shows 90% of legal cases in 2008 included digital evidence of some sort (Science Daily, 2009).

Issues abound when dealing with digital evidence. While best practice recommendations help, the final decision on admissibility often lays within the decision of an individual judge. Also, once admitted, the value of that evidence can sway greatly with the presentation by an adept lawyer, as well as the perception of a jury who may or may not bring with them expectations of magic based on Hollywood portrayals of the magic of Information Technology. Research has been done on the "CSI-Effect", how the portrayal of technology in television crime shows effects potential jurors (Davis, Paullet, et al, 2010). This and other studies (Overill, 2013; Slaught, 2013; Hayes, Leavett 2013; Cole, 2013) concludes that those who watch a large amount (4 hours a week or more) of crime related shows, often exhibit un-real expectations in regard to the capabilities of technology in relation to crime scene investigation, and are less likely to answer knowledge questions on forensic technology correctly than those who watch less crime related television. Other studies have shown that law enforcement, though, has altered their practices in response to the perceived "CSI-Effect" and expectations of potential jurors in regard to amount and accuracy of collected evidence (Kopaki, 2013).

With the amount of cases involving digital evidence and the expectations of jurors to have that evidence collected and presented perfectly, the need for trained technicians who can present themselves and the evidence in a clear and understandable manner is tremendous.

Questions

Is there a certification process to verify experts in trials have been properly trained in handling digital evidence? What is the percentage of law enforcement personnel who have been trained in digital evidence collection and interpretation? What is a "hash" when imaging a hard drive? What are multiple methods that could be used to write protect a hard drive while imaging? If a hash changes, does it make any evidence inadmissible in a court of law?

3. Jobs

Analysts who work for state or federal law enforcement agencies usually earn a starting salary of between $50,000 and $75,000. Salary can increase with experience, advanced degrees, and security clearance.

Computer Forensics Investigators or Forensic Analysts are in high demand. With abundant opportunities in both public and private sectors, the job outlook is excellent. The US Department of Labor projects a growth rate of over 20% between 2010 and 2020, placing the profession in the projected top 10 percent of growth professions. (BoLS, 2013)

4. Lab Exercises

The following exercises attached (Appendix A) can be performed together, or broken into parts to correspond with ability and level of the course. There are two major tasks involved, taking an image of the hard drive and recovering files from the image. These tasks can be taken separately, or as a series. One possibility for condensing assignments would be to have an instructor take an image of a hard drive ahead of time and have students work individually analyzing the image provided to them.

5. Bibliography

Cole, Simone (2013, June). A surfeit of science: The "CSI effect" and the media appropriation of the public understanding of science. Journal of Public Understanding of Science. Retrieved from http://pus.sagepub.com/content/early/2013/04/09/0963662513481294.abstract

Davis, Gary; Paullet, Karen; Houck, Max; Swan, Tom (2010) Does the Technology Portrayed In Television Crime Shows Have an Effect On Potential Jurors? Issues in Information Systems, *Volume XI, No. 1.* Retrieved from http://iacis.org/iis/2010/154-163_LV2010_1439.pdf

Dragland, Ase (2013, May 22). Big Data, for better or worse: 90% of world's data generated over last two years. *ScienceDaily*. Retrieved from

http://www.sciencedaily.com/releases/2013/05/130522085217.htm

Hayes, Rebecca; Leavett, Lora (2013, June) Community Members' Perception of the CSI Effect. American Journal of Criminal Justice Volume 38, Issue 2, pp 216-235. Retrieved from http://link.springer.com/article/10.1007/s12103-012-9166-2

Kassner,Michael, (2010, June 14). The truth about copier hard drives: Tips for securing your data. The Tech Republic, Retrieved from http://www.techrepublic.com/blog/it-security/the-truth-about-copier-hard-drives-tips-for-securing-your-data/

Ketyean, Armen. (2010, April 19) Copy Machines, a Security Risk? CBS Evening News, Retrieved from http://www.cbsnews.com/video/watch/?id=6412572n

Kopacki, Christopher (2013, August 12). Examing the CSI Effect and the Influence of Forensic Crime Television on Future Jurors. Dissertation Defense Virginia Commonwealth University, retrieved from https://dizzyg.uls.vcu.edu/handle/10156/4465

Overill, Richard (2013). The 'inverse CSI effect': further evidence from e-crime data**.** *Int. J. Electronic Security and Digital Forensics*, 5, 81-89

Norton Security (2012, September 4). The 2012 Norton Cybercrime Report. Norton/Symantec, Retrieved from http://now-static.norton.com/now/en/pu/images/Promotions/2012/cybercrimeReport/2012_Norton_Cybercrime_Report_Master_FINAL_050912.pdf

SIMTEF (2009, January 1) Digital Evidence: Cyber Forensic Researchers Make The Call. Retrieved from http://www.sciencedaily.com/videos/2009/0104-digital_evidence.htm

Slaught, (2013) The Real CSI: A Criticism of Media's Manipulation of Forensic Science. Dissertation Defense. Retrieved from http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1142&context=comssp&sei-redir=1&referer=http%3A%2F%2Fscholar.google.com%2Fscholar%3Fq%3Dcsi%2Beffect%2Bin%2Bthe%2Bcourtroom%26btnG%3D%26hl%3Den%26as_sdt%3D0%252C39%26as_ylo%3D2013%26as_vis%3D1#search=%22csi%20effect%20courtroom%22

US Bureau of Labor Statistics (2013) Occupational Outlook Handbook, Forensic Science Technicians. Retrieved from http://www.bls.gov/ooh/life-physical-and-social-science/forensic-science-technicians.htm

USDOJ (2013) United States Department of Justice, Computer Crime and Intellectual Property Section. Retrieved from http://www.justice.gov/criminal/cybercrime/

Appendix A

The following exercises can be performed together, or broken into parts to correspond with ability and level of the course. There are two major tasks involved, taking an image of the hard drive and recovering files from the image.

Task 1: Taking an image of the hard drive.

Needed: Hard drive, power to hard drive, data connection cable, imaging software.

Hard drive: Any hard drive will do, even failed drives with bad sectors. IDE and SATA drives are most commonly available and most easily connected to a workstation.

Connection: For these exercises, we will assume that we are not completing what would be considered a legal forensics copy of the drive (details can be covered in subsequent follow up research). As such, direct connection to a workstation through internal data cables and internal power supply connectors can be accomplished. However, if such access is unavailable due to security locks or other restrictions, external USB connection cables with a separate power supply can be purchased at a current estimated cost of under $5 each.

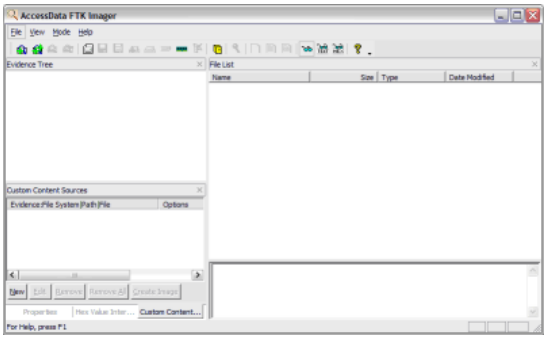
Imaging Software: There are many tools available to achieve this task. This paper will provide a step by step procedure utilizing FTK Imager from Access Data. This application is available in many other compilation packages of forensics tools, including the Forensic Tool Kit also available from Access Data. It is also available as a standalone program:
http://www.accessdata.com/support/product-downloads#Utilities

Assumption: The following procedure describes the creation of a drive image utilizing FTK Imager 3.1, a Windows 7 host machine, an external USB connection cable, and a 20GB IDE hard drive.

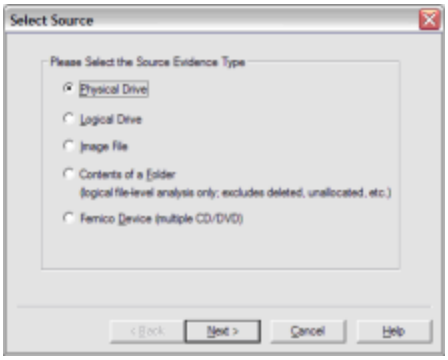
1 – Prepare the hard drive for connection. The drive may or may not have a jumper that manually sets the behavior of the drive when connected to act as a master or slave drive. Most often, removing the jumper will cause the drive to act in the mode of "cable select" and leave the device open for connection without interfering with any existing drive configurations of the host machine. This is of much greater importance if connecting the drive internally.

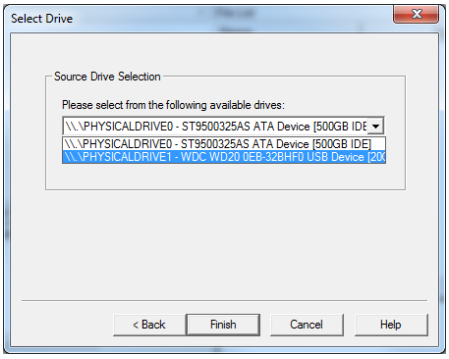
2 – Connect the drive. Connect the data cable and then the power cable. Power up the machine if it is not already on. The hard drive should be automatically found by a windows based machine.

3 – Run **FTK Imager.exe** to start the tool.

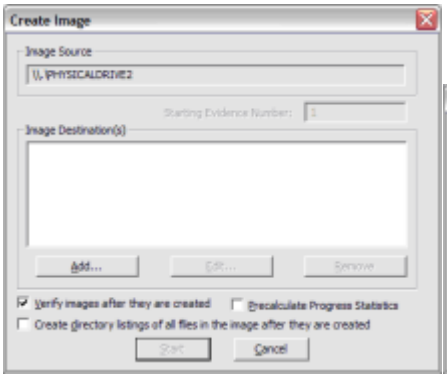


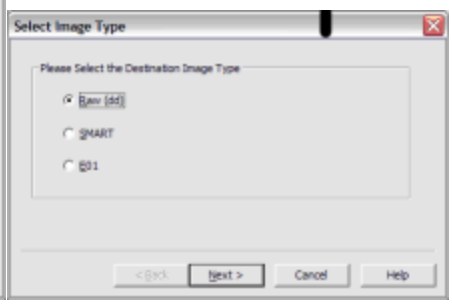
4 - From the **File** menu, select **Create a Disk Image** and choose the source of your image – **Physical Drive**. After clicking NEXT, you should see at least two options in the drop down menu, the system drive of the host machine, as well as the attached external drive. In this case, "PHYSICALDRIVE1" is also tagged as being a "USB device" for easy identification. Click **Finish.**



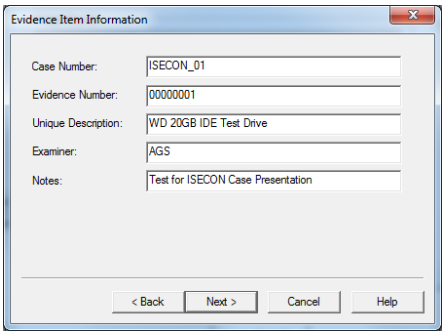


5 - To direct the image destination in the next screen, click **Add**. For greater flexibility in recovery software later, choose the **Raw (dd)** format and **Next**. Check **Verify images after they are created** so FTK Imager will calculate MD5 and SHA1 hashes of the acquired image.

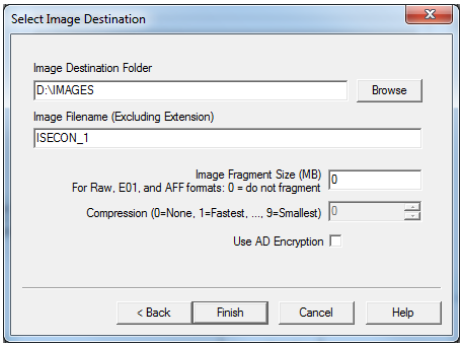




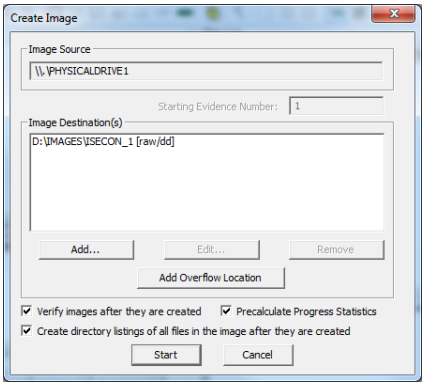
6 – Provide some information to identify the drive and your imaging session. If you select raw (dd) format, the image meta data will *not* be stored in the image file itself. A text file log will be created at the end of the process. Select **Next**.



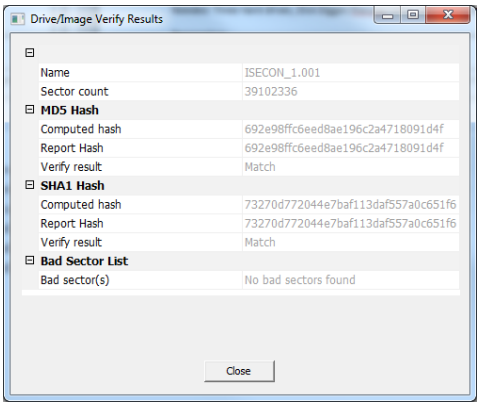
7 - Select the Image Destination folder and file name. You can also set the maximum fragment size of image split files for large capacity hard drives. For this example, enter "0" to create one image file. Click Finish to complete the wizard.



8 - Click **START** to begin the acquisition:



A progress window will appear. For this example, a 20GB IDE drive was imaged in approximately 45 minutes. Once the acquisition is complete, you can view an image summary and the drive will appear in the evidence list in the left hand side of the main FTK Imager window. You can right-click on the drive name to Verify the Image:



FTK Imager also creates a log of the acquisition process and places it in the same directory as the image, ***image-name*.txt**. The file will list the evidence information, details of the drive, check sums, and times the image acquisition started and finished:

Created By AccessData® FTK® Imager 3.1.3.2
Case Information:
Acquired using: ADI3.1.3.2
Case Number: ISECON_01
Evidence Number: 00000001
Unique description: WD 20GB IDE Test Drive
Examiner: AGS
Notes: Test for ISECON Case Presentation

Information for D:\IMAGES\ISECON_1:

Physical Evidentiary Item (Source) Information:
[Device Info]
Source Type: Physical
[Drive Geometry]
Cylinders: 2,434
Tracks per Cylinder: 255
Sectors per Track: 63
Bytes per Sector: 512
Sector Count: 39,102,336
[Physical Drive Information]
Drive Model: WDC WD20 0EB-32BHF0 USB Device
Drive Serial Number: 152D20337A0C
Drive Interface Type: USB
Removable drive: False
Source data size: 19092 MB
Sector count: 39102336

ATTENTION:
The following sector(s) on the source drive could not be read:
1091244
The contents of these sectors were replaced with zeros in the image.

[Computed Hashes]
MD5 checksum: 692e98ffc6eed8ae196c2a4718091d4f
SHA1 checksum: 73270d772044e7baf113daf557a0c651f6155602

Image Information:
Acquisition started: Mon Aug 19 11:00:23 2013
Acquisition finished: Mon Aug 19 11:45:13 2013
Segment list:
D:\IMAGES\ISECON_1.001

Questions:

1 – What steps need to be performed to ensure a sample hard drive does not have any data written to it during the image taking process?

2 – Can any drive be imaged? What steps may be taken to image a drive that may be damaged or is not recognized by Windows?

3 – What can be used to verify that a hard drive has not been tampered with, or that anything has changed after an initial image has been taken?

4 – If there is confirmation that the drive contents have changed between an initial imaging and later testing, would this automatically preclude the drive from being used as evidence? Explain.

Task 2: Recovery Comparisons

Needed: Drive image, Disk Digger/Autopsy/ ReclaiMe /other data recovery tool

http://diskdigger.org/download

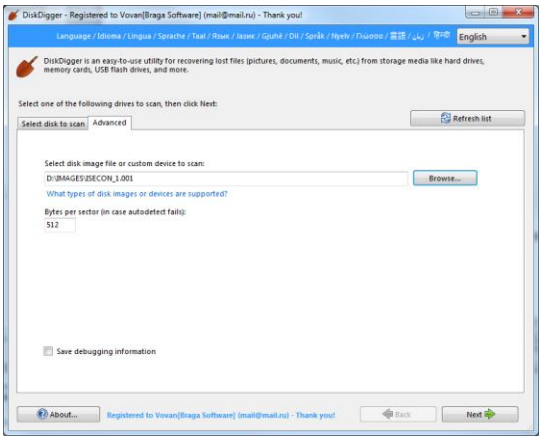
http://sourceforge.net/projects/autopsy/files/autopsy/3.0.6/

http://www.reclaime.com

http://killdisk.com/downloadfree.htm

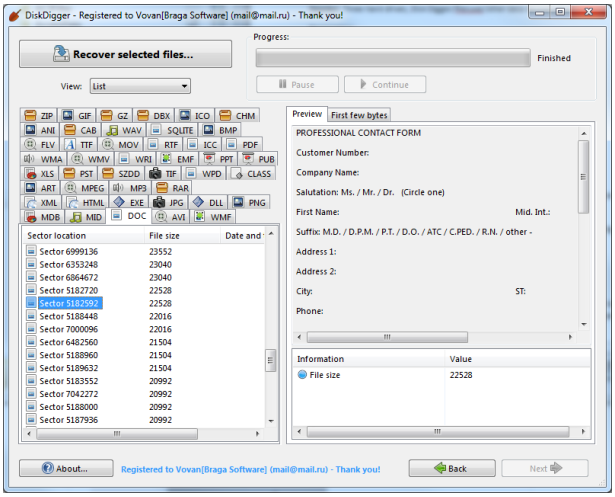
Recovering files with Disk Digger –

1 – Run DiskDigger.exe to start the tool. Click on the Advanced tab to load the drive image file. Browse for the file and select Next.



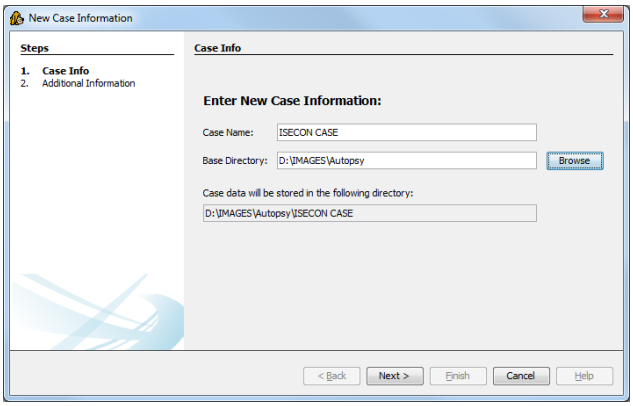
2 – For this exercise choose Dig Deeper. Continue with default selections. Selecting next will start the scan. This example scan completed in approximately 10 minutes.

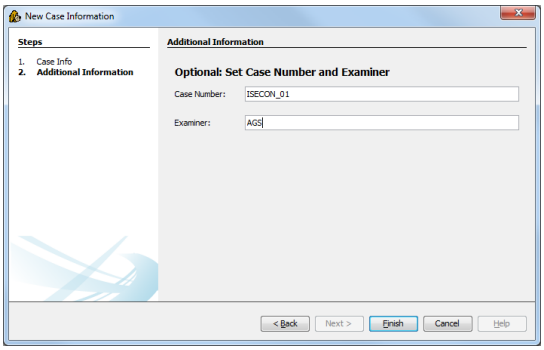
3 – During the scan and once completed, the Disk Digger interface allows for a preview of found files. Selected files can also be exported and saved to a collection directory of your choosing.



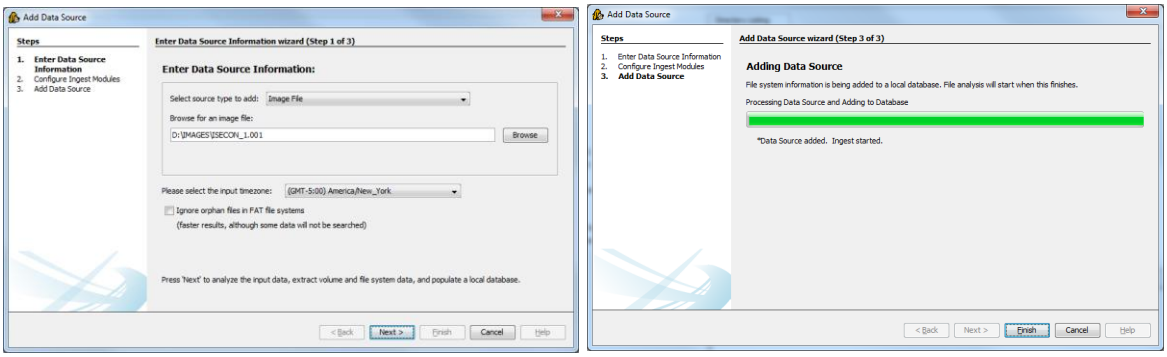
Recovering files with Autopsy

1 – Open Autopsy from the start menu and create a new case. Provide details for case name and directory to store files. You will also be asked for a case name and name of the examiner.

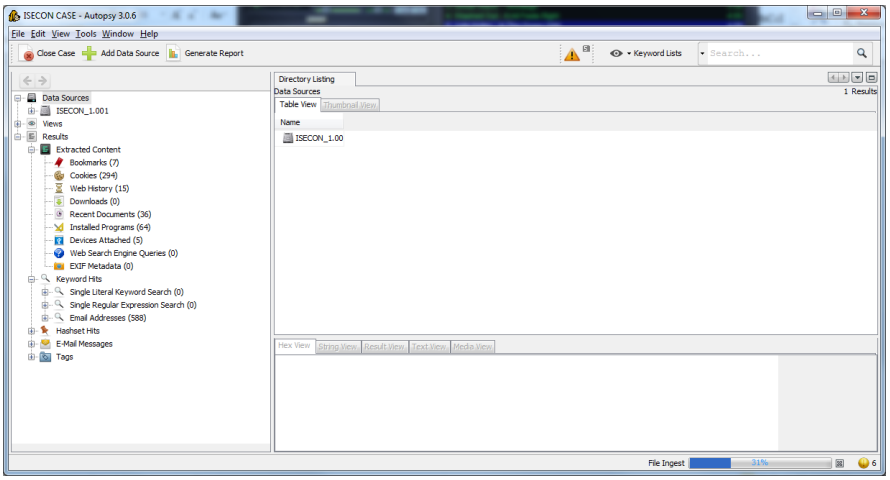




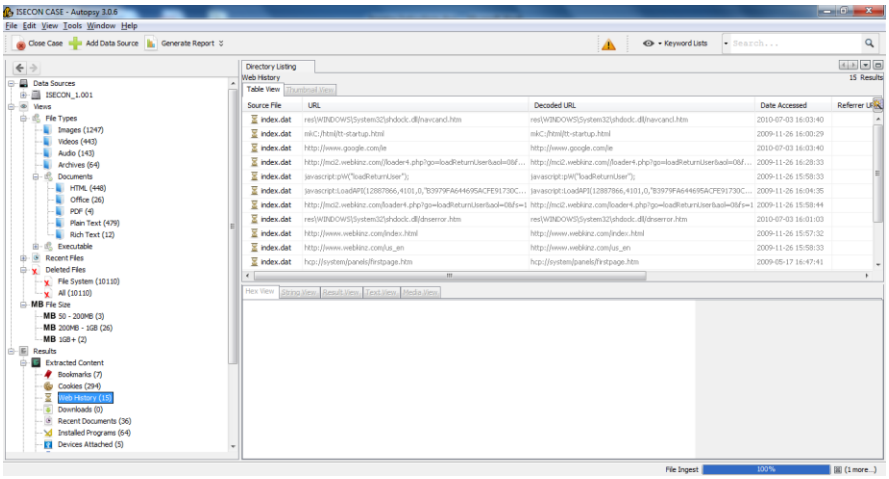
2 – Enter the path to the image file. Take default selections for ingest models. A pause will occur as the image file is added. Scanning of the image file will commence immeadiately.



3 – A progress bar can be seen in the lower right of the program. Scanning can take a considerable amount of time.



4 – The final report from Autopsy is arranged differently than Disk Digger as the purpose for the program is different. Results are arranged in more of a report fashion with categories laid out for more investigative purposes rather than file recovery

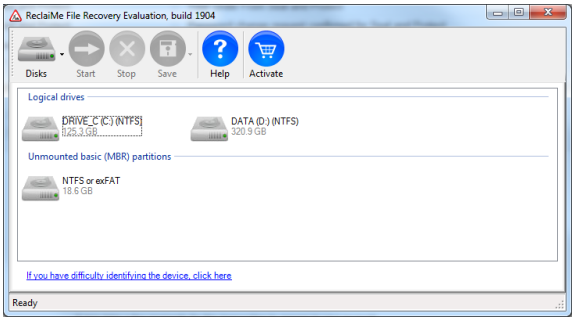


Recover with ReclaiMe:

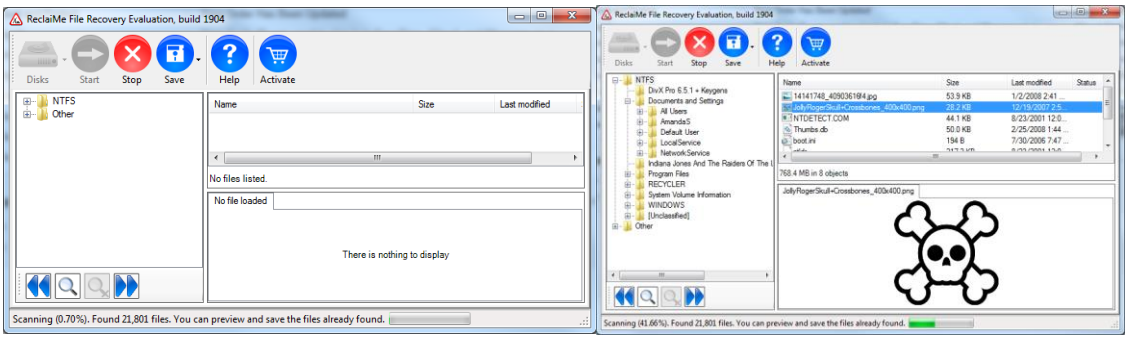
1 - Start program

2 – The program will automatically find the existing hard drives. Add your image file by clicking on the "Disks" icon in the top left tool bar. Browse for the folder the image is in. You may need to set th file type to "All Files" as ReclaiMe will work with dd images, but does not automatically find the find from FTK Imager tagged with a "001" file extension.

3 – To start the scan, double click on the newly shown "un-mounted drive" or highlight and choose "Start".



4 – A new window will appear and show a progress bar at the bottom left corner. A preview of the found files can be seen while the scan is still running.



Task 3 - Deletion Comparison:

Needed: 3 hard drives, preferably of similar size and previous use

Task: Compare recovery possibilities from three different hard drives

Setup: Prepare the three drives as A) simply removed from a computer with no deletion or formatting; B) perform a "quick format" process on the drive; and C) run the drive through a third party tool such as Active Kill Disk

Activity: Ask students to image all three drives and attempt to recover any files from the drives.

Questions:

1 – What was the approximate number of files found for each drive?

2 – Have students find and evaluate three other hard drive erasing products.

3 – Have students write up a lab report in a proper lab report format. The report can be for one, or a comparison of multiple methods of retrieval and deletion. Proper formatting, section identification, and conclusions should be evident.

Teaching Case

The Power of an MIS Degree: Inspiring students by connecting with innovators

Cameron Lawrence
Cameron.Lawrence@business.umt.edu

Shawn Clouse
Shawn.Clouse@business.umt.edu

David Firth
David.Firth@business.umt.edu

Gerald Evans
Jerry.Evans@business.umt.edu

School of Business Administration
The University of Montana

Nathan Stephens
nathan.stephens@groundswellmediaproductions.com
Groundswell Media

Abstract

Recently the ISEDJ introduced a new area within its case study category that seeks to capture and disseminate successful classroom practices and teaching strategies. The primary motivation behind this enterprise is the construction of a repository that will help new academics draw from successful classroom strategies and techniques that have been successfully employed by peers around the world. This paper is a contribution toward this inaugural effort. The aim of this paper is to share a website we developed that helps our department tackle several challenging problems including the active engagement of students throughout the MIS curriculum, recruiting students to the MIS major and connecting our students and program to innovators throughout our region. Through this site we are effectively showing our students the importance of gaining a deeper understanding of technology and how studying MIS can help one pursue almost any career imaginable. This paper describes the site we developed and clearly shows how others can implement this format within their programs. Finally, this is applicable to all courses within the MIS Model Curriculum.

Keywords: MIS Enrollment, MIS Curriculum, Student Recruitment, Innovation, Entrepreneurship

1. Introduction

"So, what can I do with an MIS degree?" This is a question that prospective students ask almost every MIS faculty member. In fact, in our discipline's recent history, it appears that many of us have not had a satisfactory answer to this simple question as evidenced by the enrollment declines observed globally. This realization and the subsequent conversation and ideas on how to combat enrollment declines has been documented widely and discussed in many of the major MIS journals (Koch, Slyke, et al 2010; Gefen, Ragowsky, et al 2012). For those of us who have committed our professional lives to the advancement of our discipline, we are somewhat puzzled by the question because from our perspective the answer is self-evident: "What can't you do with an MIS degree?"

The transformative role of technology and its importance are well documented and a crucial element of modern business and society. In fact, this is one of the dominant narratives of our time, and its importance can't be overstated (Friedman and Mandelbaum 2012; Schmidt and Cohen 2013). Furthermore, many argue that a critical understanding of technology is also necessary for the cultivation of an informed citizenry, which underpins dynamic and vibrant nations (Rushkoff and Purvis 2010). This is why it is particularly surprising that many of our students don't see the connection between studying MIS and a successful career. The faculty in our program have wrestled with this issue and have made substantial contributions to the discipline's exploration of the enrollment crisis (Firth, Lawrence, et al 2008, Eoin and Firth 2013). The aim of this paper is to share a successful strategy that has far surpassed our expectations. Beginning in October of 2012 we began publishing weekly profiles that highlight entrepreneurs, technologists, and innovators within our geographic area. As you will see, the profiles we publish are unique in that they focus on the nature of modern work and how technology fits into their successes.

The remainder of this paper is influenced by the structure the Journal of Information Systems Education recommends for its Teaching Tips submissions (See Lending and Chelley 2012). Following this introduction you will find a more detailed description of our suggested practice and a brief discussion of how we manage the process. We will then discuss how we employ the profiles in classroom discussions and suggest a few alternative uses. We then provide evidence in the form of a Google Analytics report that shows the number of page views during the 2012-2013 academic year.

2. Site Description: Meet the Innovators

During the late summer of 2012, we conceived of an idea that we believed would help our students see the value of an MIS degree and to enrich our classroom discussions with real-world examples of how people use technology to accomplish exciting work. Accordingly, we put together a website that features innovative technologists and entrepreneurs and we asked them to discuss topics that complement the MIS mission. It also encourages students to explore technologies with which they are unfamiliar. The site became operational on October 1, 2012, and has far exceeded our expectations. Initially, we simply wanted to inspire our students by introducing them to people doing impressive work in our area; however, it has turned into a resource that has captured the imagination of not only our students but also the larger technology and entrepreneurial communities in our region. This has been particularly helpful because our MIS program is increasingly drawing students interested in entrepreneurship.

Each profile consists of six questions, which are:
1) Who are you and what are you doing?
2) What hardware are you using?
3) What software and web services do you use?
4) Describe the system you use to manage your time and resources to make sure the right things are getting done.
5) What books, ideas, and people have influenced your thinking and might be of interest to others?
6) What can our state do to increase its creative and entrepreneurial cultures?
*** See http://www.mtusesthis.com to view completed profiles.

In choosing participants we seek out people we believe our students can identify with and ask them to write their profiles in an approachable and engaging way. When visiting the site the reader will immediately notice that we strive to feature a diverse group of occupations and individuals. The profiles range from coders working in their basements to successful entrepreneurs who have sold their companies for hundreds of millions of dollars. Perhaps the most useful feature is that we link to all of the technologies and resources our featured guests reference. For example, if a developer uses

Github to manage software development, we link to the technology. We also ask the participants to share two pictures with our audience. The first image a participant provides is a profile image, while the other is a picture of his or her workspace. In the future we are considering having participants share screenshots of their mobile device's home screen as well.

The management of this project takes some planning and organization, particularly as this project has grown. For example, we currently have over forty-five people working on profiles. We have streamlined our processes and rely heavily on Google Docs. We simply create and share a document with the person completing the profile. That person then completes the profile and subsequent revisions in Google Docs prior to being published on our site. In addition, we rely on two Google Spreadsheets to track profiles that are actively being worked on as well as setting and managing the publishing schedule. The simple tools we use allow us to easily manage this project and are widely available to colleagues around the world. We choose to host our site off of the university's servers because this allows us more freedom in the process. The hosting requirements are modest, and all modern higher education institutions can easily provide the hosting capabilities for a project such as this.

We introduced the site to our students in October 2012 by taking a little time to share the project with students in many different courses ranging from our Intro to MIS course to graduate students within our MBA program. We shared with them our goal for the project and demonstrated how the site worked. We then placed links to the site in our learning management system and announced new profiles as they were published. Our early intention was to get students to begin to explore and learn about people and technology on their own and without the motivation of an exam. We chose not to introduce exam questions related to the profiles in our courses, and we will probably keep that policy in place during the next academic year. Almost immediately we began to receive positive comments from students across our program. As new profiles were published, we would often begin class by discussing the profiles and the technologies our featured participants were using.

Throughout the last academic year, two dominant themes emerged in our students in relation to this project. The first major theme was the surprise our students experienced when they discovered local, enterprising professionals immersed in work our students had not yet begun to consider. We often feature people who have chosen paths outside of the typical corporate environment to pursue more entrepreneurial and creative opportunities, a concept which is consistent with the direction of our school and program. Slowly our students began to look at our region and its business environment in a different and more positive way. In addition, we are able to show how technology is often a critical component that is infused throughout the modern work environment. We then began to invite some of the people we profiled into the classroom as guest speakers, which has been an extremely successful practice.

The next major theme we identified will shock some outside of our discipline, but many MIS faculty will not be the least bit taken aback. This theme concerns the limited exposure of our students to widely available mainstream technologies. The authors of this paper collectively have over 50 years of experience teaching undergraduate and graduate MIS courses, and we believe one of the greatest myths regarding the "millennial" generation is that cohort's technical sophistication. While today's students have grown up with digital technology embedded in their lives, it is our experience that a significant proportion of this group only uses these incredible digital tools in very superficial ways; therefore, discussing the various technologies in class turns out to be a tremendous learning opportunity because it exposes our students to new tools and applications while simultaneously encouraging them to explore in a low-pressure environment. Effectively, we are introducing our students to successful people with whom they may often identify and then showing them that these individuals use technology in efficacious and innovative ways.

The exposure and exploration of these themes in class discussions helps our students see the power of an MIS degree. We emphasize that studying MIS helps prepare people to take advantage of the technology-intensive nature of modern business regardless of the industry or area within which they wish to work. In addition, we show our students that it is incumbent upon them to go beyond the surface of technology and to dive deeper. Once we discuss the profiles, it is evident to our students that their own use of technology often pales in comparison to how these successful people use the same tools. Furthermore, we encourage our students to look

critically at their own use of technology and to realize that their use often favors the simple consumption of information rather than higher level creative and management applications of technology.

While we are thrilled with our students' interest, we are equally happy—and surprised—by the interest in this project by outside constituencies, including the tech and entrepreneurship communities. We hear regularly from members of these groups and their interest in this project. It turns out that our site serves as an important resource for some members of these communities to learn about people, technical tools, and influences of those we profile. This also helps our program connect with these individuals, which allows us to share what we are working on and to get them involved with our program and students. This has also turned out to be a good mechanism for connecting students with employers for internships and job opportunities. In addition, some of those featured have been asked to join our MIS advisory board.

3. Google Analytics Report

We are surprised at the reach and the number of viewers our site received from October 1, 2012, through May 31, 2013. Appendix A reflects metrics captured through Google Analytics. The data presented are descriptive measures that simply illuminate the general exposure the site received. As the Google Analytics report indicates, we received over seven thousand unique visitors, which implies that our reach is well beyond our student population. Interestingly, 67.7% of our site visitors are new viewers, while 32.3% are returning viewers, meaning that we have over two thousand viewers that have returned over multiple sessions. In the future we plan to develop a more sophisticated set of measures to specifically capture our students' use of the site and its effect on their choices of majors. We hope to report our findings to colleagues through this journal.

4. CONCLUSIONS

This paper represents a contribution toward a new category within the case study area of the ISEDJ. The project outlined in this paper has far exceeded our expectations, and we believe colleagues from around the world will find our experience useful and easily see ways they can implement this concept. As the field of MIS continues to mature and evolve, we must strive to find creative and novel ways to engage our students and show them the value and power behind the MIS discipline. We submit the effectiveness of this project is due to the fact that we bring in credible, engaging voices from outside the academy to convey the value of integrating business and technical knowledge, which, at the end of the day, is the story of MIS.

5. REFERENCES

Diane Lending, & Vician, C. (2012). Writing IS Teaching Tips: Guidelines for JISE Submission. *Journal of Information Systems Education*, *23*(1).

Eoin Whelan, & Firth, D. (2012). Changing the Introductory IS Course to Improve Future Enrollments: An Irish Perspective. *Journal of Information Systems Education,*, *23*(4), 395–406.

Eric Schmidt, & Cohen, J. (2013). *New digital age.* [S.l.]: Random House.

Firth, D., King, J., Koch, H., Looney, C., Pavlou, P., & Trauth, E. (2011). Addressing the Credibility Crisis in IS. *Communications of the Association for Information Systems*, *28*(1). Retrieved from http://aisel.aisnet.org/cais/vol28/iss1/13

Firth, D., Lawrence, C., & Looney, C. A. (2008). Addressing the IS Enrollment Crisis: A 12-step Program to Bring about Change through the Introductory IS Course. *Communications of the Association for Information Systems*, *23*(1).

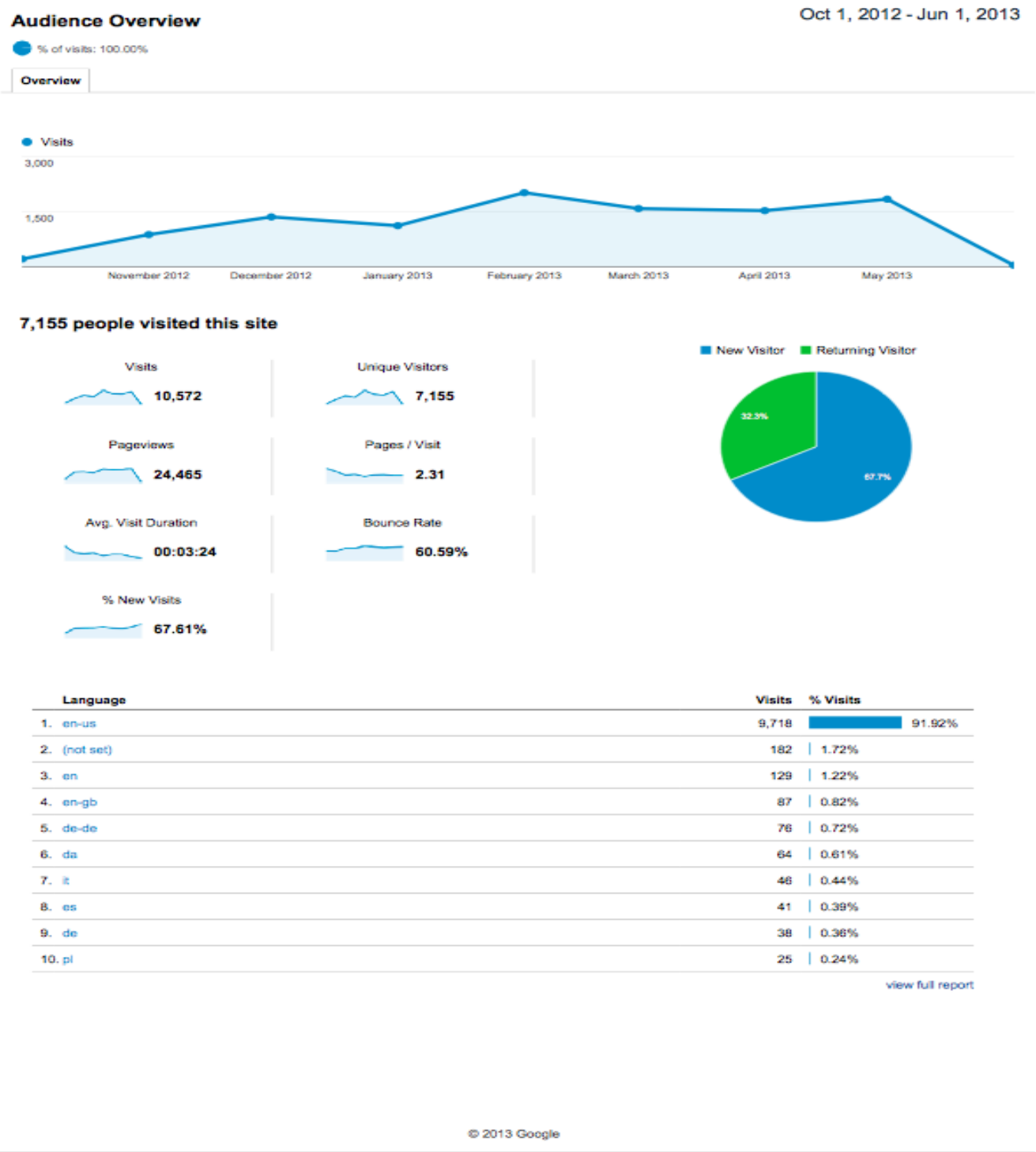
Friedman, T. L., & Mandelbaum, M. (2012). *That used to be us: how America fell behind in the world it invented and how we can come back*. New York: Picador/Farrar, Straus and Giroux.

Gefen, D., Ragowsky, A., McLean, E., Markus, M., Rivard, S., & Rossi, M. (2012). ICIS 2011 Panel Report: Are We on the Wrong Track and Do MIS Curricula Need to Be Reengineered? *Communications of the Association for Information Systems*, *30*(1).

Koch, H., Slyke, C. V., Watson, R., Wells, J., & Wilson, R. (2010). Best Practices for Increasing IS Enrollment: A Program Perspective. *Communications of the Association for Information Systems*, *26*(1).

Rushkoff, D., & Purvis, L. (2010). *Program or Be Programmed*. OR Books.

Appendix A


Audience Overview
Oct 1, 2012 - Jun 1, 2013
% of visits: 100.00%
Overview
Visits
3,000
1,500
November 2012
December 2012
January 2013
February 2013
March 2013
April 2013
May 2013
7,155 people visited this site
Visits
10,572
Unique Visitors
7,155
Pageviews
24,465
Pages / Visit
2.31
Avg. Visit Duration
00:03:24
Bounce Rate
60.59%
% New Visits
67.61%
New Visitor
Returning Visitor
32.3%
67.7%
Language
Visits
% Visits
1. en-us 9,718 91.92%
2. (not set) 182 1.72%
3. en 129 1.22%
4. en-gb 87 0.82%
5. de-de 76 0.72%
6. da 64 0.61%
7. it 46 0.44%
8. es 41 0.39%
9. de 38 0.36%
10. pl 25 0.24%
view full report
© 2013 Google