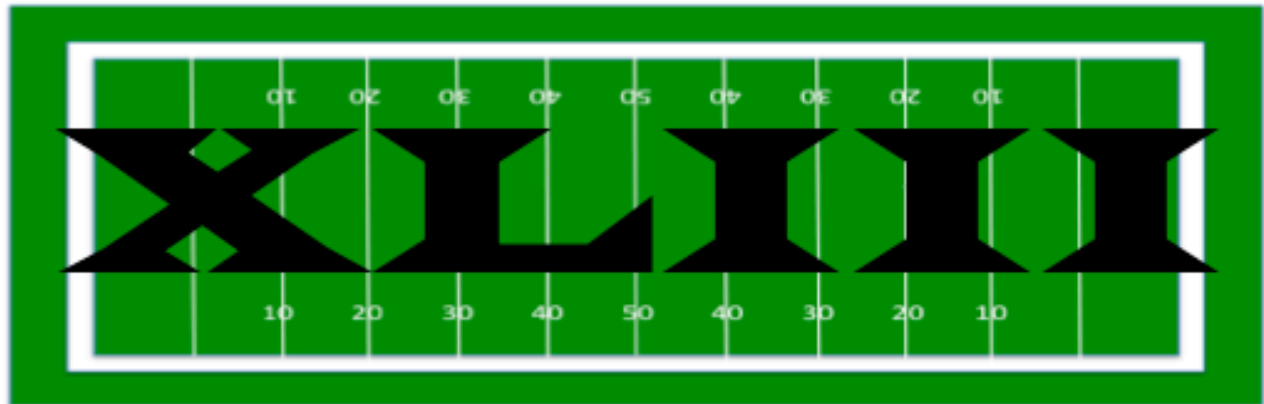


*The Alan Simon Series on
Innovative and Transformational
Business Intelligence*

presents



**SUPER BOWL 43
&
BUSINESS INTELLIGENCE**

**“Replay Challenges”
Must be
Part of Your BI Environment**

This document contains proprietary information, which is protected by copyright. All rights are reserved. No part of this document may be photocopied, reproduced, stored in a retrieval system, transmitted in any form or by any means, or translated into another language without the prior written consent of Precision Business Intelligence, LLC and Alan R. Simon.

This edition of this document published June, 2009.

Copyright © 2009 Precision Business Intelligence, LLC. All Rights Reserved.

This document and its contents are not affiliated with nor endorsed by the Arizona Cardinals or Pittsburgh Steelers organizations, the National Football League, or any player, coach, or team official mentioned herein. No representations are made by Precision Business Intelligence, LLC, the author, or any other party to the contrary.

Author's Introduction

Business intelligence is an incredibly important discipline for any company, governmental agency, or non-profit organization. The basic premise of business intelligence (BI) is not only easy to understand, it's almost a "statement of the obvious." After all, who would argue against making "better" decisions that are driven by real data drawn from real applications, rather than relying solely on one's instincts and experience or other "gut feel" yet often imprecise mechanisms?

But as you attempt to get past the generalized, 30,000-foot level "statements of goodness" about the value of BI to your organization and begin mapping out a strategy and architecture for your own efforts, it's very easy to focus your attentions on perhaps a dozen or so "heavy hitter" best practices and give little or no attention to dozens of other BI best practices that are equally important but...well, they're pretty dry and boring, to be honest.

With this challenge in mind, I'm always on the lookout for some way to convey some of these "not so sexy" precepts and best practices of BI and its companion discipline, *data warehousing*, in a manner that is entertaining...or at least holds the interest of a reader or audience member. And thanks to the Pittsburgh Steelers and Arizona Cardinals, I've found what is probably the best vehicle yet to catch the attention of business and technology executives...especially in Phoenix and Pittsburgh: Super Bowl 43 (in "Super Bowl language" that would be XLIII), played on February 1, 2009 between these two teams.

This isn't the first time I've used NFL football and, in particular, the Pittsburgh Steelers as the foundation for a business intelligence discussion. Lesson (Chapter) #4 of my recently published book ***Blocking and Tackling Your Way to Management Success: 40 Essential Lessons from 40 Years of the Pittsburgh Steelers***¹ used the 1972 National Football League college player draft and the Steelers' surprise selection of Franco Harris as the backdrop for a discussion of several critical but often overlooked BI best practices. I excerpted that chapter as a white paper² and the union between football and business intelligence has been so well received by readers that it seemed a "sequel" was in order.

On a personal note (and so those of you in the Phoenix business community don't throw this document down in disgust, given my Steelers-themed management book and the outcome of Super Bowl 43!), I should point out that my Arizona roots are almost as deep as my Pittsburgh ones. My family moved from Pittsburgh, where I grew up, to Tucson shortly before my 16th birthday and I graduated from high school, college (ASU), and graduate school

¹ For more information about this book please visit:
<http://www.precisionbusinessintelligence.com/Books.html>

² Also available for download at
<http://www.precisionbusinessintelligence.com/Papers.html>

(University of Arizona) in Arizona. And after being back east since the early 1990s, I moved back to Phoenix in 2008 as part of my recent business expansion initiative and was in Arizona when the Steelers and Cardinals faced off in Super Bowl 43.

Therefore, I decided that this paper should have a “Cardinals slant” to it. Super Bowl 43 was an evening of end-to-end excitement with James Harrison’s 100-yard interception return for a touchdown as the first half expired; the Cardinals’ wham-bam, go-ahead touchdown strike late in the fourth quarter; and finally Pittsburgh’s winning touchdown drive. But as I’ll describe in more detail in the paper that follows, just about totally overlooked amidst all of this excitement was the fact that during the game Arizona coach Ken Whisenhunt *successfully* challenged two on-the-field calls that were reversed in the Cardinals’ favor...which got me to thinking about “instant replay challenges” as an analogy for one of the most problematic aspects of real-world business intelligence in almost any organization.

As you’ll soon read, it’s my belief that *every single business intelligence environment* MUST have immutable, infrastructure-supported capabilities and well-defined “rules of engagement” analogous to those used by the National Football League for game-time coaching challenges and instant replay to support the principle of “making the correct call almost all the time.”

So whether you celebrated into the night after Pittsburgh’s sixth Super Bowl title or if you’re still convinced that Santonio Holmes didn’t get both feet down in the end zone and that the Cardinals were robbed, I hope you score a few “business intelligence touchdowns” of your own by applying the BI best practice described in the pages to follow.

Alan R. Simon

Managing Principal, *Precision Business Intelligence, LLC*

alan.simon@precisionbusinessintelligence.com

Super Bowl 43 and Business Intelligence: “Replay Challenges” *Must* be Part of Your BI Environment

I. The Setting

Super Bowl 43. Tampa, Florida; February 1, 2009. Pittsburgh Steelers versus Arizona Cardinals.

For the second year in a row the Super Bowl was one for the ages, right down to the final seconds. This one...well, this game pretty much had it all:

- ✦ The longest play in Super Bowl history: James Harrison of the Steelers rumbling 100 yards down the sideline with an interception return as the time ran out in the first half. Touchdown, Steelers!
- ✦ The Cardinals refusing to give up, hanging in there...and then suddenly taking the lead late in the game with a lightning strike from Kurt Warner to Larry Fitzgerald for 64 yards. Touchdown, Cardinals!
- ✦ The Steelers shaking off Arizona’s late-game surge with Ben Roethlisberger leading a John Elway-like 78-yard drive that was capped with a “catch for the ages” by Santonio Holmes (if you’re a Steelers fan) or, if you’re a Cards fan, one of the worst on-the-field calls in Super Bowl history because you’re convinced Holmes didn’t have both feet inbounds. But the scoreboard told the tale: Touchdown, Steelers!

- ✦ Kurt Warner attempting to mount an even later-in-the-game drive that ended with a controversial ruling that he fumbled the ball away just as he was preparing for his final passing attempt. Victory, Steelers! Super Bowl Triumph Number Six!

Wow! What a game! But overlooked amidst the spectacular plays and controversy over Holmes’ touchdown and Warner’s fumble (and even some mild grumbling that Harrison stepped out of bounds at one point during his 100-yard jaunt) was that earlier in the game, Arizona head coach Ken Whisenhunt successfully challenged two different on-the-field calls that resulted in 1) canceling out a Pittsburgh touchdown and forcing the Steelers to settle for a field goal, and then 2) overturning a ruling that Cards quarterback Kurt Warner had fumbled, instead allowing Arizona to punt the ball away on the next play rather than turning it over to the Steelers at mid-field.

Whisenhunt was simply doing what a good NFL coach needs to do: *selectively and carefully* use the game-time power he has to challenge an on-the-field call that goes against his team in the hopes that the instant replay system will support his objection and thus overturn the call made by the official on the playing field.



II. Challenges & Instant Replay in the NFL

It wasn't too long ago that in an NFL game, rulings by officials on the field were as sacrosanct as if they came chiseled into tablets and handed down from a mountain by the spirits of the founding fathers of pro football. In 1986 the NFL instituted a limited instant replay system but it wasn't until 1999 that the capability for a coach to explicitly challenge an on-the-field ruling went into effect.³

As you would expect to happen with any new capability such as this one there were not only explicit rules put into effect to govern the process but also the development of "best practices" to guide coaches in determining when they should – and just as importantly, should *not* – issue a challenge.

First the rules. For purposes of this paper (and in the interest of brevity), rather than recite every little nuance about the NFL's system for using instant replay and challenges as well as minor rule changes along the way I'll briefly highlight ten of the most salient points that we'll revisit later in this paper in the context of business intelligence (and yes, this list may seem long but believe me there's lots more!):⁴

1. "The system" denotes a rigid, finite set of certain on-the-field situations for which challenges can be made (e.g., scoring plays, touching of a kick, etc.)...and excludes all others.
2. Coaches have a finite number of challenges (currently, two) available to them in a game.
3. There is a precisely defined, limited time period (i.e., before the start of the next play) for a coach to issue a challenge.
4. There are also formally designated "inclusive" and "exclusive" situations that govern when challenges can and cannot be made, regardless of whether a coach has used his allotted two challenges. Specifically, a coach cannot issue a challenge within the two-minute warning of either half or when his team doesn't have a timeout left in that half.
5. But wait! There is also a "replay assistant" who takes over after each half's two-minute warning and he (or she) goes by a different set of rules to call for an instant replay review.
6. There is a formally defined mechanism by which a coach definitively indicates that he wishes to issue a challenge, just as there is a related but different mechanism for the replay assistant.
7. There is an always-in-place, always-used "infrastructure" of instant replay video and viewing capabilities by which a challenge is ruled upon.
8. The referee has a finite, limited period of time (60 seconds) to watch the replay and decide if "incontrovertible visual evidence" exists that justifies overturning the on-the-field ruling with minimal effect on the pace and longevity of the game.
9. If a coach's challenge is unsuccessful – that is, the call on the field stands – then his team is penalized by losing a timeout.

³ Source: http://en.wikipedia.org/wiki/Instant_replay_in_American_football#National_Football_League

⁴ *ibid.*

10. Finally, if a coach's challenge or an indication from the replay assistant that something is amiss is upheld due to the aforementioned "incontrovertible visual evidence", the on-the-field call is overturned in favor of "the right call that should have been made in the first place."

As I mentioned earlier, there's lots more and if you're really into learning (or seeing how much you already know) about the littlest nuances of the NFL's instant replay system, check out one of the many sources available on the Internet (as I did!).

Before moving onto the analogous aspects of the NFL's instant replay system with regards to your business intelligence environment and implementation, it's important to note that it's not just the challenge and instant replay *rules* that are important, but also the associated "best practices." For example:

1. Since a coach only has two challenges in a game and will lose a timeout if his challenge is not upheld, he needs to be very careful not to "squander" either in a frivolous manner. Or, stated another way: challenges must be "spent" very carefully.
2. Assistant coaches and even on-the-field players provide time-critical input to a coach to let him know that there's a strong likelihood that the wrong call was made and a challenge would be upheld. Or, stated another way: *select* supporting data is *quickly* made available to the coach so he can form an initial hypothesis about whether or not a challenge might be warranted.
3. The coach must also consider other factors, such as the "value" of the on-the-field call being reversed...i.e., exactly how beneficial will the

reversal (should it occur) be to his team versus leaving the play stand as-is? Or, stated another way: being "right" isn't enough; one has to be "very, very right" to justify using a challenge.

4. Even if everyone in the stadium is convinced that a call should be overturned, does the required "incontrovertible visual evidence" exist to support the challenge? Or, stated another way: even if the coach and his team are right, *can that actually be proven?*

So as you see, it's the combination of well-defined rules along with a set of best practices that has evolved from the earliest days of instant replay and challenges that governs and guides the actions of coaches and officials in the pursuit of "getting the calls right."

Which brings us to business intelligence.



III. The Quest for "A Single Version of the Truth" (SVOT)

One of the earliest motivations for an organization embarking on the twin journeys of *business intelligence* and *data warehousing* was the absence of a "single version of the truth" within almost every enterprise. Since modern computer systems and applications share data among themselves with reckless abandon and often continue to manipulate their own copies of shared data, it was and is common for two or three or more systems to produce different results and tallies from what should be the exact same sets of data. And to complicate

matters further, these “should-be-the-same-but-aren’t” sets of data are often only part of the overall information that is fed into reports and analytics, and therefore tend to “pollute” the results that are derived from data consolidated from various sources, even if other data is of a much higher, more uniform quality.

The solution that surfaced circa 1990 was to bring together data from various sources into a *data warehouse* as the single location from which the organization’s most critical analytics would be run. Without going into too much detail about data warehousing, and also sidestepping many of the long-standing controversies over centralized versus decentralized architectures and the like, suffice it to say that one of the primary objectives for spending all that time and money building and maintaining a data warehouse was to provide a definitive, authoritative, and *incontrovertible* (remember that word from the previous section?) source of data content for critical reports and analytics and the subsequent decision-making.

What still happens far too often, though, is that the data warehouse provides “a single version of ‘*an answer*’...which may or may not be the truth!” Reports and analytics sourced from organizational data warehouses are frequently “challenged” by executives and managers who contend that those “official” reports and analytics are flawed, erroneous, untrustworthy, pollute the planet...well, maybe not the last point but you get the idea.

And what about these “challenges?” Here is where Super Bowl 43 and the NFL’s system for challenges and instant replay review dovetails into this critical shortcoming in most organizations’

respective BI environments. Specifically: unlike the system that Cardinals coach Ken Whisenhunt so expertly used during Super Bowl 43 (game results notwithstanding), **the act of challenging report and analysis results in corporate and governmental settings is rarely 1) governed by formal rules; 2) supported by a solid technology infrastructure; or 3) accomplished in accordance with use-vs.-don’t-use best practices adhered to by those managers.**

The frequent result: repetitive, time-consuming, resource-draining efforts to check and double-check and triple-check data warehouse results versus those from the source data as challenge after challenge after challenge is issued.

The consequence: lengthy delays in making critical decisions and taking decisive, information-driven action...basically, the result becomes the *exact opposite* of having readily available timely, accurate, high-value, actionable insights that the BI/data warehouse environment was built to provide.



IV. Magnifying the Problem: “Spreadmarts” and Departmental Data Marts

What’s behind the multitude of challenges to the results produced from the entity (the data warehouse) that’s supposed to contain “the single version of the truth?” Are managers just acting in a contrary manner because reports and analytics from the data warehouse indicate that their

organization isn't meeting their objectives? Are they challenging the "official" results on a basis of unsubstantiated hubris and bravado, without any supporting data at all?

Perhaps sometimes, but more often than not the challenges come from managers and executives armed with *their own reports and analytics* produced not from the enterprise's official data warehouse(s) but rather from their own departmental data marts into which they themselves feed data. Or, despite an organization's stated objectives at the outset of a data warehousing initiative to banish for all time spreadsheets serving as official stores of critical analytical and reporting data – i.e., the dreaded "spreadmarts⁵" – the reality is that across the typical enterprise spreadmarts continue to thrive and replicate like something out of a 1950s horror movie.

Thus, perhaps even more so than with database-hosted data marts, managers come to meetings armed with spreadsheet printouts indicating that "their version of the truth" clashes with what the data warehouse says.

What happens then? Are there rules akin to the NFL's that govern these kinds of managerial challenges? No way! What usually occurs is that several members of an organization's IT staff are sent away to dig, dig, DIG at all kinds of data to figure out whose "version of the truth" is really the truth.

And all the while, the organization's objective for rapid decision-making and follow-on action based on data-driven insights

is sorely compromised...and the return on their investment in business intelligence and data warehousing technology continues to diminish...

Before moving from problem to solution, it's important to note that it's my belief that the approach organizations traditionally have pursued to address this problem – trying to gradually banish "unofficial" data marts and spreadmarts – is *absolutely, positively unachievable*. From a technology-driven perspective, having all reports and analytics coming only from officially sanctioned enterprise-supported data warehouses seems to make sense. But the culture of *every* organization for which I've ever done BI strategy and architecture work is the absolute antithesis of an environment where this monolithic, autocratic approach could actually work for a prolonged period of time.

So: as we segue into the next section, we need to take it as a given that in all but the rarest situations (and I've yet to see one), a decentralized topology of spreadmarts, departmental data marts, and at least one enterprise data warehouse *will* exist in any given company, governmental agency, or non-profit organization... and must be dealt with as that organization continues its laudable journey towards a single version of the truth for reports and analytics. Difficulties and complications aside, an organization *must* pursue the ideal of a single version of the truth to survive and thrive, in both challenging economic climates and then as recovery and re-growth takes hold.

How, then, can an organization move past today's version of a "challenges-driven" model that puts

⁵ Wayne Eckerson of *The Data Warehousing Institute* (TDWI) is usually credited with coining the term "spreadmart."

managers and executives at odds with one another to the detriment of the organization as a whole?

Simple.

Your organization needs to support – in fact, *encourage* – the concept of challenges, but in a new era based upon:

- Well-defined rules
- An infrastructure to support the enforcement of, and adherence to, those rules
- A universally accepted “code of best practices” with regards to those rules and making use of that infrastructure

Basically, you should look to those who have made this model work – the National Football League – and follow their example.



V. Addressing the Problem, Part I: Conceptual Model

Part of the reason organizations tend to shy away from forcefully addressing the “challenges” issue when they plan and architect their BI environments – or rearchitect and “revitalize” existing ones – is that it’s too easy to get “lost in the weeds” of technology components, their known data integrity issues, and many other lower-level details. So let’s first look at a conceptual model abstracted away from technological matters as we build a framework for what an organization might do.

To do so let’s “repackage” those key concepts from how the NFL institutes an effective, efficient

mechanism for challenges and instant replays into a set of critical success factors (CSFs) for the prototypical BI environment.

1. *Strive for a manageable number of challenges.* You’ll never be able to do what the NFL does – limit the number of challenges to two (or three or some other very low number)...nor should you try. You need to allow for the possibility that there may be significant merit in the objections of departmental managers due to serious deficiencies in the quality of the data warehouse’s contents. So rather than try to institute some sort of cap to how many challenges a given manager can make, you should:
 - a. Preemptively and proactively devote significant effort to identify and eliminate discrepancies between “official” sources of data (e.g., your data warehouses) and the aforementioned departmental data marts and spreadmarts that are the sources of the objections. Basically, you don’t wait until a whole lot of “on-the-field bad calls” occur; you try to prevent the bad calls from occurring in the first place. Again, this needs to be done in a “grounded in reality” manner, recognizing that while you may be able to eliminate some departmental, “unofficial” sources of reports and analytics, you’ll never be able to eradicate all of them. In the next section, we’ll look at the technological foundation that you should strongly consider to help eliminate discrepancies in the first place.
 - b. But given that discrepancies *will* continue to occur and will almost certainly always be present within your enterprise, give managers the tools to self-analyze the reasons for those discrepancies *before* they feel

- the need to “issue a challenge.” Hold that thought for the next section, we’ll look at how you can do so.
- c. Foster a culture where challenges must be issued *very carefully*, otherwise some sort of penalty is the consequence. The NFL’s taking away one timeout for a challenge that isn’t upheld acts as a deterrent against a coach issuing a challenge to, for example, a touchdown by the opposing team that clearly should not be questioned. The most flagrant violators of the “challenge protocol” within your organization need to face some sort of penalty. Perhaps it’s something as simple as an offending manager’s challenges defaulting to a category of “suspicious and probably without merit” until he or she starts “behaving better.” On the more extreme side, particularly flagrant violators could see their budgets or even personal compensation docked slightly (sort of an “Unsportsman-like Conduct” or “Personal Foul” penalty). But whatever is decided, the penalties need to be as tangible and as much of deterrent as those used by the NFL.
 - d. Managers should be *required* to issue their formal challenges accompanied by actual supporting data – even if that data comes from their own departmental sources – rather than simply on a “hunch” that the official numbers from the data warehouse “have to be wrong, because they don’t make sense.”
2. *Have a formal, immutable mechanism by which challenges are issued.* In the NFL, coaches currently throw a red flag onto the field before the start of the following play. You need to have in place something similar:
 - a. No “in the middle of a meeting” surprises that derail a meeting’s agenda and cause chaos to ensue. Perhaps it’s at least a day’s advance notice to the meeting organizer that a challenge needs to be put on the agenda, or maybe it’s some sort of web-based application that allows *authorized* managers (more on that in a moment) to upload their specific objections along with the supporting data.
 - b. As with the NFL’s allowing only the head coach to issue the challenge, your organization needs to have a similar role. I’ve been in far too many meetings where the “challenge surprise” comes from a data analyst or another lower-level employee – sometimes without his or her manager being forewarned! Such employees need to escalate their own challenges, along with their supporting data, to their managers for “pre-qualification” of the challenge.
 - c. You need to have some sort of formally designated time limit for challenging *specific* reports or analytics...which first means that challenges do need to be against specific results, not just general “the data warehouse is wrong!” complaints. (Requiring supporting data helps enforce this rule.) Of course you need to allow more than a minute or so (i.e., the NFL’s by-the-start-of-the-next-play limit), so maybe it’s within two business days after a report or analytic is produced. This has the side benefit of actually requiring managers to *use* (or at least look at) reports and analytics coming from the data warehouse; too often managers with a propensity for managing by “their own numbers” act as if the data warehouse and the enterprise’s official business

intelligence environment doesn't even exist.

- d. Whereas the NFL has the replay assistant to issue challenges in addition to coaches, you should have the equivalent...specifically, the BI environment itself must support *data reconnaissance and reconciliation* "under the covers and behind the scenes" as we'll discuss in the next section.
3. *You need to have a formally designated timeframe to **resolve** challenges.* In the NFL, an official has all of 60 seconds to make the call whether a challenge should be upheld and an on-the-field call overturned. You definitely need to allow a longer timeframe than that (!) but you also need to prevent what so often occurs: challenges are issued and the investigation and resolution goes on and on and on... So:
 - a. The specific resolution timeframe you select for your organization needs to vary according to how mature and robust your infrastructure is (see the next section), and will likely evolve over time. At the outset you may allow, say, four or five business days (don't go longer than that!) as new data digging and synthesis has to occur, followed by the necessary analysis. But as your infrastructure matures you should be able to get to a point where challenges can be resolved *in a single business day, two at the most*. Yes, I know this sounds optimistic, perhaps even unachievable...but trust me, you can do it!
 - b. You need to have the equivalent of an NFL referee – that is, an objective individual with sufficient "standing" and authority within the organization, as well as someone who doesn't have a vested interest in the outcome of the challenge – be the individual

who "makes the call" whether or not to uphold a challenge and what should occur next. Organizations who have a formally designated Chief Data Office (CDO) should strongly consider bestowing this responsibility on that individual (especially since the CDO has overall responsibility for the quality of the enterprise's data and the reports and analytics that result).

4. *A challenge must be formally upheld or not...no middle ground!* In an NFL game, it's impossible for a coach's challenge and the subsequent instant replay review to be left unsettled. The on-the-field call is upheld, or it's overturned; anything else is impossible and violates the physical laws of the universe (or something like that...) In too many business organizations, though, there is a sort of middle ground – Purgatory, perhaps? – in which *any given manager's challenge* can go on and on and on... For your organization, though:
 - a. Recognize that the longer you leave a given challenge "fester" (i.e., your resolution timeframe isn't met for whatever reason), 1) the more the distrust of your overall BI environment and the underlying data warehouse contents increases, and 2) the more your organization's data-driven decision-making processes will be compromised.
 - b. If "the call" is to support a manager's challenge, meaning that his or her departmentally-produced data is correct and the official data is not, then you need to officially – i.e., publicly – support the challenge and overturn the results of what was produced by the data warehouse.
 - c. Going a step further than the NFL needs to, you need to "regression-correct" previous

erroneous results that had slipped through.

- d. Sometimes in the NFL, a *particularly* bad on-the-field call occurs...often it's overturned as a result of an instant replay review, though sometimes it isn't. But regardless, the league office goes to work and issues memos and communiqués to its officials to help prevent such errors from reoccurring. Or, stated another way: *rapid corrective action becomes paramount*. For your BI environment and underlying data warehousing, the analogy should be an obvious one: you must quickly fix erroneous data, ETL code, business rules, embedded report template logic...whatever has been found to be incorrect.

And there you have it: a nice, straight-forward conceptual model for your BI environment based on tried-and-true precepts courtesy of professional football.

But conceptual models are one thing; making those concepts actually work is another. Which brings us to our next section.

◆

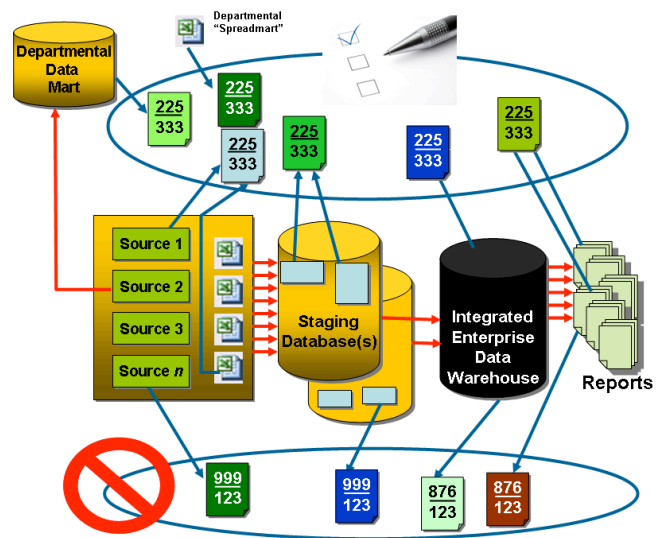
VI. Addressing the Problem, Part II: Architecture, Technology, and Implementation

As noted earlier, the National Football League's system for challenging and attempting to overturn on-the-field calls requires an immutable, always-available (during games) infrastructure to support an authoritative ruling on

whether or not a challenge should be upheld: specifically, the instant replay recording and reviewing system.

Your BI environment needs to have the equivalent: an always-there, ready-to-use technology infrastructure that unfailingly supports the challenge process and subsequent rulings. The problem: far too many BI implementations lack this fundamental capability when deployed because the always-too-limited development resources and funding was focused elsewhere.

The following figure illustrates the basic precepts of what I introduced earlier as an inherent, pervasive model of *data reconnaissance and reconciliation*. As in most any business intelligence environment, data flows exist among the various source systems, staging databases, data warehouse(s), and reports. You'll also find data feeds from data sources over to departmental data marts that are outside the architecture of the data warehousing environment.



The principle behind this architecture is a simple one. The

reconnaissance occurs in an “always-on” manner, regularly (i.e., daily) pulling a large number of data samples from various data stores along the information chain into several “buckets.” Within each “bucket” each and every data sample should present exactly the same results according to the data itself and the business rules that are applied for each of these behind-the-scenes reconnaissance extracts.

As illustrated in the diagram, the top extract checks out correctly, while in the bottom one, the data warehouse and its reports produce different results than the source systems and the staging database(s) do. And so, the “replay assistant” – the system itself – challenges this portion of the data and flags it for “instant replay.”

This bucket-based, multiple-data-sample model helps those charged with isolating and fixing the problem to determine where they need to focus their investigative efforts: in this particular case, it appears that somewhere between the staging databases and the data warehouse errors are being introduced.

Now let’s look back at the CSFs presented in the previous section and see how the data reconnaissance and reconciliation capabilities help you meet them:

1. *Striving for a manageable number of challenges* – The system itself will act as a sort of always-on “first alert” instant replay official, scanning and comparing the check figures produced from the various components and at the first sign of discrepancy, triggering investigation and repair. However, the system won’t be able to verify every possible “check figure package” that could conceivably be produced from every single component, so managers will

inevitably still feel the need to challenge certain results. But before doing so their respective technical teams will be able to access the reconnaissance infrastructure and *much more rapidly than if such an infrastructure didn’t exist* do the necessary “pre-investigation” work by modifying, manipulating, and combining the code used to produce existing check figures to advise the manager if a challenge is warranted...unlike today’s all-too-common model in which the investigative process more often than not needs to start from scratch (i.e., raw, lowest-grain, business rule-void data) at the outset of a given challenge.

2. *Having a formal, immutable mechanism for issuing challenges* – Having the reconnaissance and reconciliation infrastructure in place gives organizations the justification to demand that all challenges be accompanied by supporting data; no more excuses along the lines of “it will take too long to gather the data...but I just *know* the data warehouse is wrong anyway.”
3. *Having a formally designated timeframe for resolving challenges* – Some amount of manual data gathering and synthesis will still have to occur to resolve a challenge, especially during the early stages of your infrastructure being in place as capabilities are being built out. But the more robust your environment is, the quicker its “replay officials” (i.e., the technologists) can splice together different existing check figures and their respective code and come up with a definitive ruling on the merits of any particular challenge. Essentially, you will have provided the equivalent of the video cameras and the instant replay viewer that has already “recorded” most, maybe even all, of the inputs necessary to rule on a challenge.
4. *Upholding or rejecting a challenge rather than leaving things in a state of uncertainty* – The preceding three

CSFs and the conclusive evidence produced from the environment make it very difficult to let politics, overbearing personalities, or other similar factors play into the merits of a challenge and the subsequent ruling...or lack of a ruling, i.e., data/report/analytics controversy continues to fester indefinitely. Either A equals B, or A doesn't equal B; it's that simple. Thus a challenge is upheld – proven to be correct – or it isn't. You still need to have the supporting business processes in place, i.e., a senior executive with sufficient authority such as a Chief Data Officer being the one to rule on a challenge and not be swayed by managerial bluster and bravado over actual evidence.



VII. Beware the “Metadata Talisman”

But what about metadata? Doesn't metadata cure all data-related ills as well as stop global warming? (Okay, I haven't actually heard an over-zealous software salesperson make the “stop global warming” claim about his or her company's product line but I'm not that far off...)

A robust metadata capability certainly adds understanding to questions of the type “where did this number come from?” and “how did this piece of data get transformed at each hop on its way into the data warehouse?” But whereas metadata by itself provides enlightenment and increases the pace of investigation and resolution, it (metadata) still requires the rest of the infrastructure described in the previous section, *along with the rules and best practices*, to be an effective

part of a challenge-replay-decision model for your enterprise.

Watch for a future *Precision Business Intelligence, LLC* white paper that will present a best practices-driven architecture and topology for metadata and master/reference data in support of not only “challenges” but your BI environment as a whole.



VIII. Conclusions

For many years it's been almost mandatory to declare that solutions and systems must have three equally important aspects to be successful: technology, process, and human factors. The “challenges” environment and supporting infrastructure presented in this document is no exception. The technology is actually the most straightforward of all, and regardless of its potential usage to support managers' challenges should have been part of any BI environment in the first place...specifically, a robust, highly functional, end-to-end data quality assurance (QA) capability. Alas, as discussed earlier, far too many BI/data warehousing environments possess only limited QA capabilities along the chain of data flows and are totally or near-totally void in end-to-end reconnaissance and reconciliation capabilities.

But beyond technology, the processes and human factors described in this paper are essential to facilitating the infrastructure being able to successfully support the processing and resolution of challenges. Absent the rules and best practices managers across your

organization will continue to claim the existence a whole lot of data discrepancies but you'll find it extremely difficult to turn those claims into a definitive resolution and, if valid, the appropriate corrective action.

So even if you're not Arizona Cardinals coach Ken Whisenhunt, and you ply your trade in the technology and business world rather than on the playing fields of the NFL, I strongly urge you to consider the principles presented in this paper as you plan your new BI environment or embark on a revitalization effort for one that already exists. You'll have to expend time and effort and budget but will likely see extremely high returns on this investment in the effectiveness of your business intelligence and performance management missions.

About the Author



Alan Simon is the author or co-author of 28 books including *Data Warehousing and Business Intelligence for E-Commerce*, *90 Days to the Data Mart*, and *Data Warehousing For Dummies*. He is currently completing a new book about applying innovative best practices to a new generation of business intelligence solutions to enable enhanced performance management.

In his most recent book, ***Blocking and Tackling Your Way to Management Success: 40 Essential Lessons from 40 Years of the Pittsburgh Steelers***, Alan – a native Pittsburgher and long-time Steelers fan – presents a year-by-year journey through recent Steelers history (1969 through 2008) as backdrop for critical lessons in management and business success. For more information or to purchase the book, or to learn about related book-based training for your company please visit:

<http://www.precisionbusinessintelligence.com/Books.html>.

Alan is the Managing Principal of his own consultancy, *Precision Business Intelligence LLC*. He held vice president-level global and national BI practice lead positions at several professional service organizations before forming his own firm. His client work emphasizes BI and performance management strategy, architecture, and roadmap development along with techniques for revitalizing existing underperforming BI environments. He also regularly conducts “inspector general” style reviews of client projects.

Alan’s BI clients have included PNC Bank, US Steel, Highmark, Coca-Cola, McDonald’s, Quaker Oats, Pfizer, SmithKline Beecham, CoreStates Bank, and many other global companies, plus governmental agencies such as the Pennsylvania Department of Transportation (PENNDOT), Pennsylvania Department of Health, Wisconsin Department of Administration, and the United States Department of Defense.

He holds a Bachelor’s Degree from Arizona State University’s College of Business Administration with a major in Computer Information Systems (1980) and a Master’s Degree from the University of Arizona’s Management Information Systems program (1982), and also pursued post-graduate work in medical information systems and artificial intelligence at the University of Denver.