

Our Promise: "One Great Story Every Issue!"

World Tennis Gazette

February 2018

Vol 10, No 1

Big Data

**Australian Open's 'Hackathon' Fashions
Algorithm from 10,000 Points;
Rafael Nadal's Computer Chip Shots**

By JOHN MARTIN

MELBOURNE -- In a casual conversation nearly a decade ago, an Australian newspaper editor working on a research fellowship in Washington, DC, expressed his admiration for American baseball.

"It's the data," he confessed, explaining his delight at the reams of details that appeared in American newspapers and online after each game. He said he enjoyed reading the small type that accompanied the inning-by-inning box scores,

As a visiting journalist, scholar (and cricket fan), *The Australian* foreign editor Greg Sheridan found himself easily sifting details, not just of who hit home runs or struck out but the size of the crowd and even the elapsed time of play.

Australians are no strangers to sports data. Australian Rules Football fans revel in obscure rules and opaque scoring rubric, such as keeping track of "stoppage clearances," "disposal efficiency," and "contested possession."

So perhaps it's not surprising that last fall, it was Australian tennis officials who launched a massive survey of tennis match data centered on forced and unforced errors recorded in hundreds of Grand Slam singles and doubles battles

A team of more than a dozen data scientists soon began guiding a "hackathon," a global com-



World Tennis Gazette/John Martin

DATA FACTORY: The new headquarters for Tennis Australia houses team of analysts who study ways to exploit data collected from electronic scoring and court camera images.

petition between some 700 entrants who examined digital clues from roughly 10,000 points played over the equivalent of a single Grand Slam championship.

They used spreadsheets drawn from data generated by both men's and women's singles, according to Dr. Stephanie Kovalchik, a transplanted American data scientist who helped lead the Australian team.

"There aren't many job opportunities for sports statisticians in the US, particularly if you like to do more research," said Dr. Kovalchik, who studied at Cal Tech and

UCLA. "This was the first opportunity that I had come across, and it was a particularly good fit for me because it was all focused on tennis, which is my preferred sport."

During this year's Australian Open, she stood working at a computer console on the eighth floor of Tennis Australia's gleaming new headquarters, which sits at the edge of Melbourne Park's sprawling tennis complex.

Behind her, half a dozen analysts sat at desktop computers, mostly in silence as they

Continued on Page 2

Forced and Unforced Errors Predicted in Data Tongue 'R' or 'Python'

Continued from Page 1

examined and puzzled over data.

"It's an exciting time for tennis, how we use data, and how we use it to predict outcomes," said Craig Tiley, the Australian Open's tournament director.

Tiley is a former tour-level player, successful American college tennis coach (NCAA Men's Team Championship at the University of Illinois), and a longtime student of statistics,

Nevertheless, as the Hackathon proceeded, Tiley conceded that he wasn't certain where the Big Data search will lead.

"Even though I did statistics in college for four years, and I understand, you know, the regression analysis that you need to do, (and) my graduate work and my doctoral work was all in that area as well, so, but I don't know the answer to that," Tiley said.

Asked if his Game Insight Group's work was intended to find ways to replace court statisticians and lines people with machines, Tiley said it was not.

Then he added, "I do think the future is going to involve, you know, line calling that'll be automated, and data collection and analytics that'll be automated, and a lot more. I do think that is the future."

The 2018 Australian Open Hackathon search focused on finding an algorithm that could automate calls of forced or unforced errors.



Stephanie Kovalchik



PONDERING: Aussie data scholars.

A press release said the winning entry was submitted by an American data scientist, Scott Sobel. It did not describe precisely how the winning algorithm works but expressed confidence that it could eventually revolutionize line calling.

Dr. Kovalchik praised the Hackathon's result: "The winning solution is a high-quality tool that could be the first major step toward automating point call in tennis."

The press release reported that winner Sobel "was fascinated by the challenge of developing a computer algorithm that could predict winners or errors just by using the Hawk-eye tracking data.

The release quoted Sobel as saying the data-driven solution points players to ask themselves "what were the top factors that characterize a winner? How can you maximize the chance of an error from your opponent? What do you need to work on as a player?"

On the internet, Sobel was described as an insurance business consultant for a major U.S. firm and longtime data scientist.

Sobel could not be reached for comment by *World Tennis Gazette*. Tennis Australia quoted him describing data's promising future in tennis:

"Not only could it be possible to improve efficiency and consistency of otherwise manually recording point outcomes, but the greater value from using analytics is providing data-driven insights into why."

Sobel called the

Continued on Page 3

World Tennis Gazette
Vol 10, No 1 February 2018

An online magazine for players, coaches, fans, and friends of tennis throughout the world. Each issue contains a story and photographs on a single subject. All are the responsibility of the editor. To subscribe, comment, or make suggestions, email: worldtennisgazette@yahoo.com.



Richard Osborn
Editor: John Martin

Wozniak, Federer Toiled to Victory As 10,000-Point Study Unfolded

Continued from Page 2

Game Insight Group's work as coming at a "pioneering" stage.

Pioneers have been at work for some time in athletics. The world of Big Sports Data has been growing by leaps and bounds for several years.

No less than five years ago, a cover story in *The Christian Science Monitor Weekly* proclaimed:

"From curbing urban crime to calculating the effectiveness of a tennis player's backhand, people are now gathering and analyzing vast amounts of data to predict human behaviors, solve problems, identify shopping habits, thwart terrorists —everything but foretell which Hollywood scripts might make blockbusters. Actually, there's a company pouring through numbers to do that, too."

The Monitor story also observed that even then, Big Data statisticians kept track of how many first serves Andy Murray hit as he won points on his way to the Wimbledon



World Tennis Gazette/John Martin

DATA DELIVERY: Champions Caroline Wozniaki and Roger Federer played hundreds of points on way to the 2018 Australian Open titles. Raw match data was collected for future study by Game Insight Group.

singles title in 2013.

Not everyone of course — not even every Australian — regards Big Data as an unalloyed gift to the predicting or analysis of information

Judith Mortenson, a subur-

ban Melbourne mayor and research director for a respected arts group, views Big Data with a certain skepticism. Some

Continued on Page 4

Tennis Hackathon Fast Facts

- **750** data scientists and machine learners from **55** countries analyzed **10,000** points of Australian Open tracking data
- **223** participants hailed from India which was the most represented country, followed by the USA (**78**) and Australia (**51**)
- **2,731** solutions were submitted as part of the competition
- **90 percent** of all participants competed as individuals
- The two most popular languages for solutions were written in **R** and **Python**
- The winning model achieved an overall accuracy of **95 percent** – **98 percent** for winners, **89 percent** for forced errors and **95 percent** for unforced errors
- A total prize pool of **\$US8500** was awarded as part of the competition – **\$US5000** for first place, **\$US2500** for second place and **\$US1000** for third place.

— Source: Tennis Australia

How a Computer Chip Recorded Nadal's Strokes — and Mine

Continued from Page 3

scientists use it deftly and productively, she observed. Some others, she said she once heard, lean on data studies “like a drunk on a lamppost, more for support than illumination.”

Still others, including several top players, have embraced data science.

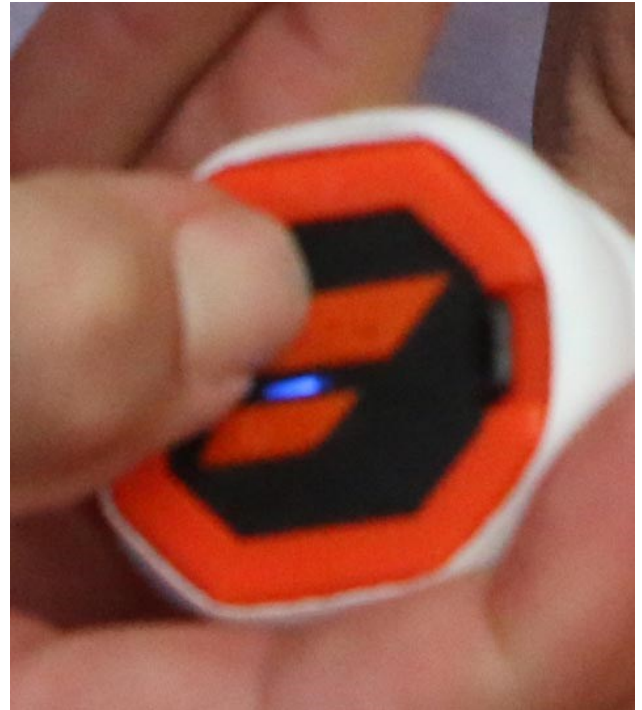
As recently as 2015, big data technology came to Rafael Nadal's racquets at the Australian Open. On the eve of the opening day matches, Nadal and his Uncle Toni invited reporters to examine a new style of Babolat racquet.

Embedded in its handle, a computer chip kept track of where balls met strings on the racquet face for both forehands and backhands. Was the stroke working? The data could be directed to a smartphone, tablet, or computer, then analyzed.

“Not every day,” said Toni Nadal when I asked if his Grand Slam champion nephew consulted the results frequently. But at least once a week, he said, the player and coach looked to see if Rafa was slugging the ball too close to the frame.

Offered a chance to hit a few balls with Rafael Nadal, I quickly grabbed a racquet and began tapping half-volleys, forehands and backhands, as we advanced toward the net.

After four or five strokes, the champion



John Martin/World Tennis Gazette
GLOW OF STROKE DATA: Light shines at Nadal's fingertip showing racquet's computer chip is turned on to operate.

had had enough, so he leaned forward, leveled his racquet parallel to the court and applied a vicious slice to the ball. The sphere landed at my feet but before I could react, it jumped back across the net.

Dazzled by his skill, I never asked whether his racquet chip recorded his stroke as a forehand, backhand, or some other data-based phenomenon.

Nevertheless, I can say without hesitation that I have seen the future of data — and it plays tennis with chip shots.



John Martin/World Tennis Gazette
TENNIS TABLET: Recording strokes and delivering signals to his tablet, Nadal's racquet joined the Digital Age some three years ago.