18 FTI Methodology Report

Independent Review of Integrity in Tennis

APPENDIX

16 APRIL 2018

DATA ACQUISITION AND ANALYSIS IN SUPPORT OF THE INDEPENDENT REVIEW OF INTEGRITY IN TENNIS

A SUMMARY OF RELEVANT FTI COUNSULTING ACTIVITIES AND ANALYTIC METHODOLOGIES

EXPERTS WITH **IMPACT**™

Table of Contents

Introduction

In July 2017, FTI Consulting LLP ("FTI") was retained on behalf of the Independent Review Panel ("Panel") conducting an Independent Review of Integrity in Tennis (the "Review") by the Panel's Solicitors and Secretariat ("Secretariat"). FTI's role has been to support the Panel by rapidly ingesting, validating, analysing and reporting and visualising information critical to the Panel's understanding of the issues surrounding integrity in tennis.

The purpose of this document is to set out, in plain terms, the methodologies deployed by FTI in delivering on these requirements. It does not attempt to provide full context for the Panel's Independent Review of Integrity in Tennis (the "Review"), and should therefore be read in light of the Panel's Interim Report and accompanying Record of Evidence and Analysis ("REA").

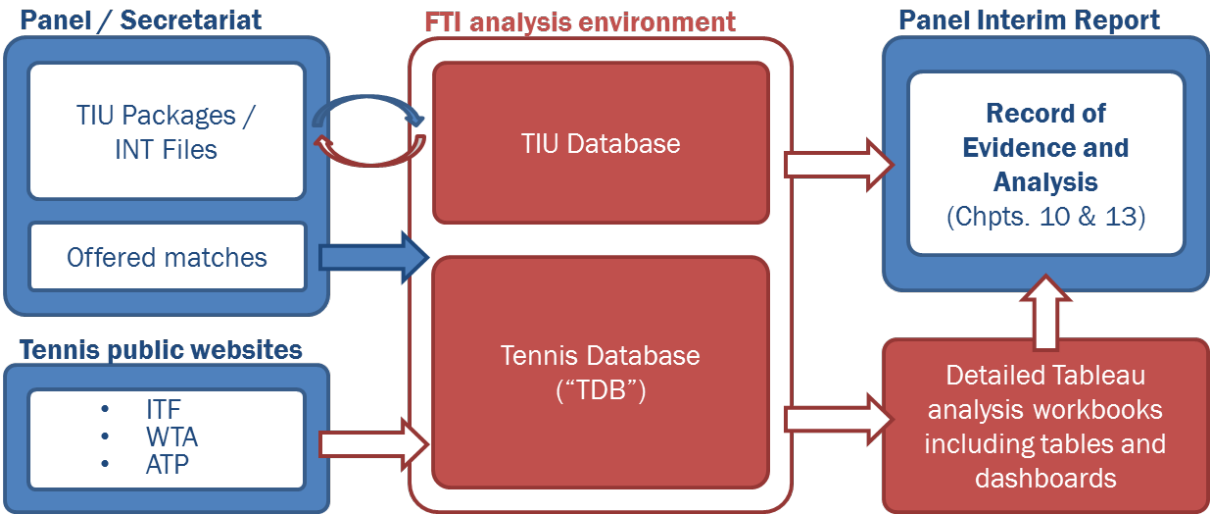
All analyses relate to at least one of two information sources:

- Data from the TIU's records collated by the Secretariat reflecting TIU Packages and INT files; and
- Publicly available tennis data for the period 2009 through 2017.

FTI's involvement in compiling this information into uniform sets in a single relational database is described below, for each information source.

The further work performed to enrich and relate information sources is detailed in the analytics section, which sets out assumptions, methods and calculations from which the tabular and visual analysis results are derived.

FTI's work to produce a series of interactive dashboards presenting an integrated view of the information sources is described in the Presentation section, along with an overview of the tabular analyses produced and prominently featured in REA, Chapter 13.



Data acquisition and validation

Effective analysis requires that information exists and is reliable; analysis at scale requires data be usefully structured and internally consistent. With these principles in mind, FTI created a single relational database repository of relevant information.

TIU alert data

Overview

The Secretariat provided FTI with a collated set of relevant alert data from the TIU's records, distributed across 16 spreadsheets generally differentiated by year and source type. Each alert entry contained an information-rich primary row (header row) potentially followed by one or more sparsely populated subsequent rows in respect of the same alert. The header row, in different columns, listed key facts related to each alert such as date, geographic and match information, the nature of the activity being alerted and its source, and name and role information for persons. Where appropriate, the information included the Panel's assessment of whether a player (or official) was a "subject" of the alert, as described further in REA Chapter 13. Rather than repeating this same information on related following rows, only differentiated new individual/role (e.g., the opponent, or in the case of doubles, the playing partner of the person listed in the primary row) and subject person flag information was recorded. In total, this information was spread across approximately 5,700 spreadsheet rows.

Data preparation and validation

FTI undertook a data transformation effort in order to standardise, clean and validate the raw alert data provided by the Secretariat. This involved transforming each multi-row spreadsheet alert entry to a single record in a dedicated database table (the "TIU Database"). Concurrent with the transformation, FTI worked closely and in an iterative manner with the Secretariat to standardise referencing, populate deficiencies (e.g. missing data such as forenames), and resolve discrepancies. Included in this initial data quality review were the "subject player" identifications made by the Panel for 2013 – 2017 alerts.

Further critical post-transformation validation processes were performed within the database by rigorously linking alert records to the publicly available tennis data. These are described in the Analytics section at page 9 below.

Publicly available tennis data

Overview

Similar to the TIU Database, FTI consolidated publicly available tennis data originating from the websites of the ATP, WTA and ITF into a single repository covering 2009 – 2017 ("TDB"). This rich information set includes data as to players (including age and ranking at any given time); events (including draw details); and match outcomes (including set splits, set scores, and retirements).

Upon being engaged in July 2017, the Secretariat provided FTI with a consolidation of this information for the period 2009 – 2016. Before placing any reliance upon the data, FTI performed rigorous testing and validation in the manner described below. Separately, FTI undertook a comprehensive web scraping and validation exercise to extend the data through 2017. In the course of obtaining 2017 data, FTI was also able to supplement information covering the earlier periods, particularly related to player rankings, age and nationality.

The specific sources are the official websites of:

- International Tennis Federation (the "ITF") - www.itftennis.com;
- Women's Tennis Association (the "WTA") - www.wtatennis.com; and
- Association of Tennis Professionals (the "ATP") - www.atpworldtour.com.

These sites present data for Grand Slams so it is not necessary to separately scrape data from the websites of AELTC, FFT, USTA or Tennis Australia.

Data scraping and validation

Web scraping exercises fail and return incomplete or erroneous results for a multitude of reasons ranging from inconsistent site structures and insufficient scraping algorithms to server time-outs and mid-scrape site changes. Scraped data sets must be evaluated for completeness and accuracy.

The combination of ample publicly available reference/control sources coupled with the structured order of play in professional tennis enables logical and rigorous completeness validation. FTI performed this review of the TDB in an orderly three-stage manner, considering: tournament, match and field completeness.

Tournament completeness

By comparing tournaments in the TDB to the separately-published official tournament calendars, FTI was able to develop confidence that relevant tournament populations from each of the ITF, ATP and WTA websites had been captured. The tables below summarises the FTI's findings for 2016 (from the provided data period) and 2017 (the independently scraped data).

Table 1. 2016 Tournament completeness validation[1]

Tournament Class	Tournament Calendar count	TDB count	Number of additional tournaments in TDB	TDB Tournament count matching Tournament Calendar	% TDB Tournament count matching Tournament Calendar
ATP	67	67	0	67	**100%**
ITF	1,215	1,220	5	1,215	**100%**
WTA	61	61	0	61	**100%**

The 2016 ITF Women's tournament calendar contained 572 tournaments; however four of these were subsequently cancelled. Further, four additional tournaments not included in the calendar were present in the TDB (see table 2). ITF M's Chile F8 appeared on the ITF website but was not present in the 2016 ITF M's tournament calendar.

Table 2. 2016 Women – ITF Tournament count discrepancy

In tournament calendar but not in TDB		
Tournament Country	Tournament Date	Draw Size
Turkey	12 Sep	-
Turkey	19 Sep	-
Turkey	26 Sep	-
Turkey	03 Oct	-

In TDB but not in tournament calendar		
Tournament Country	Tournament Date	Draw Size
Hungary	04 Jul	32M/32Q/16D
Poland	14 Nov	32M/32Q/16D
Bolivia	05 Dec	32M/64Q/16D
Bolivia	12 Dec	32M/64Q/16D

[1] Certain 2016 tournament calendars were provided to FTI for the data completeness review but WTA W 125k's and ATP Challenger official calendars were unavailable at the time of validation.

Table 3. 2017 Tournament completeness validation

Tournament Class	Tournament Calendar count	TDB count	Number of additional tournaments in TDB	TDB Tournament count matching Tournament Calendar	% TDB Tournament count matching Tournament Calendar
ATP	223	223	0	223	**100%**
ITF[2]	1,154	1,154	0	1,154	**100%**
WTA	67	67	0	67	**100%**

Match completeness

Having validated the tournament populations, FTI compared the actual number of TDB matches per tournament to calculated expected match counts based on the draw size of each tournament. The rulebooks of international governing bodies provide the bases. Broadly, the expected number of matches for any given tournament is calculated as follows:

- Main draw:

Draw size is used to find the number of rounds required to result in a single winner. For a tournament knock-out structure, the number of players participating in the first round (the draw size) has to be rounded up to the nearest power of two (to arrive at r_1 below). These additional participants are bye match "opponents". For a given draw size, the number of rounds in the main draw of a tournament, *n*, is calculated as:

$$\frac{r_1}{2^n} = 1 \, , n = \log_2 r_1$$

For example, a WTA Premier Mandatory tournament such as the Miami Open, with a main round singles draw size of 96 will have r_1 value of 128, meaning 32 bye matches, and hence a number of rounds, $n = 7$.

From this value of *n* the number of expected main round matches, M_{main}, is found by summing half the round one participant, with half again, and half again, *n* times (making sure to subtract the number of bye matches).

$$\sum_{i=1}^{n} \frac{r_1}{2i} = M_{main} = 127$$

$$127 - 32 = 95$$

- Qualifying draw:[3]

In a qualifying tournament, players compete in rounds until a tournament-specific number of qualifiers remain, to progress to the first round of the main tournament.

To find the number of expected matches, the draw size is used in conjunction with the number of qualifying players who will advance to the main draw, as stated in the rulebook for each type of tournament.

[2] The 2017 ITF calendars contained 98 tournaments that were subsequently cancelled or rescheduled (excluded from the figures presented). 2 out of the 1,154 ITF tournaments were cancelled during the course of each tournament and as such contributed played matches to the database; they are thus accounted for in both the calendar and TDB counts.

[3] In tournaments where doubles qualifying is taking place, one qualifying team advances to the main round in all cases. Therefore, the number of doubles qualifying matches will always be expected to be three where doubles qualifying is taking place, with bye matches included.

The number of qualifying rounds can be inferred from the number of times the draw size must be divided by two to get to the number of advancing players. For example, as prescribed by the rule book, an ATP Indian Wells tournament with a main draw size of 96 would have a qualifying draw size (*d*) of 48, with 12 qualifying players (*q*) advancing. $\frac{d}{2^n} = q$, where n is the number of qualifying rounds. For the ATP Indian Wells example, using the above equation, $\frac{48}{2^n} = 12,\ 2^n = \frac{48}{12} = 4$, resulting in the number of qualifying rounds *n* = 2.

The number of matches, M_{qual} in a qualifying tournament will simply be half of the draw size plus half again for n = 2, and plus half again for n = 3. More generally,

$$\sum_{i=1}^{n} \frac{d}{2i} = M_{qual} = 36$$

2009 - 2017 Match Completeness Summary

The TDB coverage period of 2009 – 2017 includes 1,084,679 matches scraped from the websites of the international governing bodies. This differs by 3,031 or 0.3% from the expected total, more than half of which are attributable to scraping limitations or data quality issues in the pre-2017 dataset. As described further below regarding the linking of TIU Database and TDB information, in addition to a 99.7% match completeness figure, FTI has established that none of these discrepancies relate to alerted matches.

Table 4. 2009 - 2017 Match completeness summary

Tournament Categorisation[4]	Matches expected based on draw size	Matches Observed	Difference
Exact match between draw size and TDB	1,062,738	1,062,738	0
Round not played (e.g. early termination of tournament)	384	211	173
Unusual qualifying round size	8,279	7,216	1,063
Exclusion due to structure (e.g. RR)	560	154	406
Scrape or DQ issue	15,749	14,360	1,389
Total	**1,087,710**	**1,084,679**	**3,031**

Key field completeness

The most granular and final level of completeness validation was performed at the individual field level (e.g., a player's name). This validation specifically considered fields containing information critical to the tabular and visual FTI analyses relied upon by the Panel.

The results presented for FTI's general field-level completeness validation cover only the years 2013 – 2017 because of the prominence of this period in the Panel's review. FTI observed improved completeness in its

[4] "Unusual qualifying round size" tournaments have qualifying draw sizes not in line with the official rule books, however, the TDB match counts correspond to the matches available on websites; "Exclusions due to structure" are related to tournaments including round robin components or unparseable round structures; "Scrape or DQ issues" relate to data quality issues or incomplete HTML data extraction

2017 scrape, and generally in more recent periods when compared with pre-2013, and expect this relates in part to evolution of the ATP, ITF and WTA websites.

Completeness results by field for Singles (33 fields) and Doubles (51 fields) TDB records are presented both by level of sport and source website in Appendix A. Overall, field completeness is greater than 99.9% in the substantial majority of cases. Ranking data for the Lowest and Mid Levels of sport are notably less complete, attributable to: 1) inaccessibility of ranking data at the time of data acquisition; and 2) the occurrence of a significant number of unranked players playing at the lowest levels in the early stages of their tennis careers.

In instances where TDB records relate to alerts from the TIU Database, incomplete or missing information was evaluated and supplemented if available.

Accuracy Testing

In addition to the rigorous completeness testing, FTI performed structured accuracy testing before relying upon the consolidated 2009 – 2016 publicly available tennis data.

To establish a statistically robust reliance basis, FTI applied Bayes success-run theorem,

$$R = (1 - C)^{\frac{1}{n}}$$

which states that from a given number of samples tested and a desired confidence level, the reliability of a population can be asserted. To achieve a 99% confidence that the data as analysed is 99.5% reliable, FTI randomly selected and reviewed 919 records from the TDB population, examining 27 key fields.

The rigor of this sampling approach originates in FTI's need to adopt a previously collated data set. In the case of each match alert identified in the TIU Database, FTI performed further validation as a critical part of establishing the linkage to TDB. This validation identified differences between the existing database information and the publicly available websites. Null or unpopulated values were not considered, as these were addressed separately and systematically where required (e.g., when a player's age at the time of the match was found to be unavailable, it is disclosed as such in the REA). The role of the TDB data for non-alerted matches then is to provide a context for the alerted ones and a canvas on which FTI alert-related analyses can be presented.

For 21 of these fields, FTI found 0 discrepancies in the 919 samples analysed. Considered individually, four of the fields had one difference (0.1%) in the sample, and a fifth had three (0.3%). The sixth field, the losing player's ranking as stored in the TDB, showed 11 differences (1.2%) out of 919. The existence of a single failure limits the appropriateness of referring to reliability in the context of Bayes success-run theorem. It would be possible to provide a subjective significance weighting to each of the 27 fields, and provide further quantification. Another way to think about this data is that out of 24,813 (27 multiplied by 919) data points evaluated, an error rate of 18, or 0.07%, was observed. FTI however does not believe further precision is required, and that the reliability of the data is well established.

Analytics

This section sets out the additional work performed subsequent to the acquisition and validation of source information, specifically:

- The relating and consequent enrichment of TIU Database alerts by TDB public information;
- The adjustment of "matches played" figures to reflect just that population on which betting could occur; and
- The assumptions, exclusions/filters and aggregations upon which analytics are based.

Relating and enriching Information

In order to present alert incidence in the context of match activity, it is essential to map the TIU Database to the TDB by deploying relational database techniques that rigorously link key attributes shared by the two information sources, including event- and player-level characteristics. Beyond the intended benefits of enriching individual alert data with publicly available TIU Database information, and broad context-setting for the population of alerts reviewed by the Panel, the mapping achieves a cross validation and corresponding data confidence augmentation because the TIU Database and TDB are two independent sources that have been made to agree.

Key TIU Database fields used in the mapping and in subsequent analysis include:

- Tournament and match-level identifiers: event name, date, location, type (level of sport), singles / doubles
- Alert activity attributes: type of alert, TIU/INT reference, alert source
- Player-level attributes: name, role (winner / loser), subject player flag (Y/N)

TIU alert mapping

Match Specific alerts

Alert entries containing both match- and player-level identifying information (henceforth "Match Specific Alerts", inclusive of "Betting Match Alerts"[5] and "Other Match Alerts"[6]) were mapped to records in the TDB using logic incorporating event date, name, location and player name criteria. Once identified in the TDB, an alerted matched was linked to the alert data via a unique identifier.

In cases of no exact correspondence between the TDB and alert match- and player-level characteristics, database searches and manual reconciliation were performed, ensuring completeness of the mapping exercise. Instances of incorrectly recorded, misreported or duplicated alert entries were flagged and excluded from subsequent analyses. Approximately 40 such exclusions were made, with a remaining count of approximately 1,400 alerts over the period 2009 – 2017 that were subject to reporting and analysis.

Non- Match Specific alerts

The TIU Database contains multiple alert types that are not specific to matches or lack match-specific identifiers. Of these alerts, "Integrity Concerns – Player" is the only alert type analysed and summarised by FTI alongside Match Specific Alerts. Approximately 600 instances of this alert type exist in the relevant period (inclusive of non-Match Specific alerts alternatively classified by the TIU as "Match-Fixing Allegations", which FTI were instructed to consider as being synonymous with Integrity Concerns – Player where the Match-Fixing Allegation did not contain match-specific identifiers).

This player-specific type of alert is presented exclusively in two visuals:

[5] "Betting Match Alerts" were formerly referred to as "Betting Alerts" in the TIU's records provided to FTI by the Secretariat.
[6] "Other Match Alerts" were formerly referred to as either "Suspicious Match" or "Match-Fixing Allegations" (in respect of Match-Fixing Allegations that contained match-specific identifiers only) in the TIU's records provided to FTI by the Secretariat.

- "Overview of All Other Alert Types" which summarises counts of non-Match Specific Alerts by type (e.g. "Abusive emails", "Courtsider Issues" etc.), and
- "Recurrence" which considers players who, on the Panel's assessment, have been the "subject" of one or more alerts of the Match Specific Alert or Integrity Concern - Player type in the period 2013-2017. In order to facilitate inference of this nature, FTI combined subject player alerts emanating from Match Specific Alerts with the non-Match Specific Integrity Concern - Player alerts in a separate database table that was used to produce a dynamic, interactive view of counts of players based on groupings drawing from players' total alerts to which they were subject, based on alert type selection.

TIU subject player alert mapping

Match Specific Alerts in the TIU Database for the 2013 – 2017 period reflect the "subject player analysis" performed by the Panel (described in REA Chapter 13) so that alerts are only attributed to specific player(s) and not every player involved in the match.

In order to derive match- and player-level characteristics from the TDB and blend these with the product of the Panel's subject player analysis, FTI mapped subject players to the TDB using (a) a unique alert ID assigned at the Match Specific Alert mapping step described above; (b) the player name; and (c) the role (winner / loser) of the alerted player.

To facilitate 100% mapping of the subject player alerts, extensive validation was performed ensuring player name discrepancies within the alert data and between the alert database and the TDB were accounted for. To do so, FTI assigned to all players in the TDB randomised unique identifiers (preventing the discernment of player identity from ID). These identifiers were later utilised for the purposes of player anonymisation in the Recurrence dashboard summarising alert activity by player for highly alerted players.

Bettable match data

Fundamental to the Panel's analysis of alert incidence is the identification of circumstances that enable the behaviour. As betting is not available on all matches, the Panel asked FTI to segregate the populations to identify alert incidence in the context of 'Bettable Matches' as contrasted with the TDB 'Played Matches' in their original form.

To facilitate the analysis of alert incidence in the context of Bettable Matches, FTI undertook an enrichment of the TDB with data on tournament booking levels. By cleaning and transforming this data, FTI was able to calculate a bet availability index (BAI) identifying, at specific levels of granularity, the proportion of played matches that were available for betting.

The two booking data sources and their date coverage and structure are:

(1) **ATP**, 2012[7] – 2017. Tabular rows with each row providing the number of matches offered per

- Tournament
- Level of sport (and gender)
- Singles / Doubles
- Draw (Main / Qualifying)

for ATP Challenger, ATP Tour, WTA 125s, WTA Tour

(2) **Sportradar**, 2013 – 2016. Tabular rows with each row representing a single match offered for betting for ITF 10s, 15s, 25s, ITF W 100-50s

[7] Given the primary utility of Bettable Match figures in terms of Referral Ratio calculation, the 2012 Bettable Match figures were not presented in the summary tables and visuals shared with the Panel due to the incompleteness of the target player analysis in this year.

For reasons of consistency in level of granularity between the two sources, FTI calculated the BAI by mapping tournaments in the data provided by ATP and Sportradar directly to TDB tournaments, differentiating along Women's game / Men's game, Singles / Doubles, and Main / Qualifying draw lines. Further, bye matches were excluded at the stages of calculation of the BAI and its application to TDB as they are not played and thus cannot be offered for betting. FTI thus:

- Counted all the matches listed in TDB for each tournament for each Women's game / Men's game, Singles / Doubles, and Main / Qualifying draw combination;
- Counted all the matches listed in (and therefore deemed to have been "on offer" to betting operators") in the data provided by ATP and Sportradar for each Women's game / Men's game, Singles / Doubles, and Main / Qualifying draw combination; and
- Divided non-bye listed matches offered by non-bye TDB matches to obtain a BAI for each category.

The BAIs derived from these sources were used to update all match-level records in TDB so that when the BAI of all matches (excluding byes) are summed, each match was counted as a fraction representing the tournament-level BAI ratio for that combination.

Further, due to 2017 data availability limitations, FTI applied the below methodology to calculate the 2017 lowest level of sport and ITF W 100s-50s bettable matches:

- 2016 Sportradar data was provided at the tournament name level, from which FTI inferred
    - Gender
    - Main / Qualifying draw
    - Singles / Doubles
    - Granular level of sport[8]
- FTI aggregated tournament names, creating a matrix for all of the above attributes and calculating a BAI for each.

The 2016 BAI matrix was applied to 2017 TDB data to create 2017 bettable matches. In cases when the 2016 BAI was greater or equal to 99%, FTI applied a BAI of 100% to 2017 data. This included main draw singles and doubles for both genders.

Assumptions, exclusions and aggregations

In producing the summary tabular and visual analyses, FTI made the following assumptions and exclusions:

- Grand Slam bet availability: A BAI of 100% was assumed for all Grand Slam events.
- Due to issues with the 2009 – 2016 data scraping of ranking data for women's doubles matches, the presentation of rank-related information has been limited to singles only.
- Bye matches: The TDB includes both played and bye matches, based on draw sizes. Since bye matches are not played, and thus cannot be alerted or offered for betting, they were excluded.
- International level of sport: The TDB includes a small number of Olympics tournaments and matches that were excluded from analysis.
- Alert exclusions: Approximately 200 alerts pre-dating the establishment of the Tennis Anti-Corruption Program in 2009 were excluded from the visual and tabular analyses.

[8] Level of sport is derived from event name for Women's game, and from prize money in the Men's game by mapping to TDB.

- The Age Profile visual excludes figures for players whose age was unavailable at the time the information was gathered. FTI has established that none of these exclusions relate to alerted matches.
- In the ITF Confederation Nationality and ITF Confederation Location visuals, country representations were limited to nationalities and locations with player-level match counts of 500 or greater.
- Aggregation by year: Since the TDB does not contain match dates, tournament start date was relied upon when summarising data along temporal dimensions. Further, in some instances tournaments commence late in the year but are included in the following year's tennis calendar. FTI used the calendar year of the tournament start date in these cases.

In producing the summary analyses, FTI aggregated raw data along multiple dimensions, grouping certain values together using insight drawn from sources shared by the Secretariat and the Panel.

- Level of Sport
    - Lowest Level: ITF 10s, 15s and 25s
    - Mid Level: ATP Challenger, ITF W 100s-50s, WTA W125ks
    - Tour Level: ATP Tour, WTA Tour
    - Grand Slam
- Ranking: For the purposes of summarising alert levels based on the subject player analysis by ranking, FTI grouped bettable matches in 10 ranking bands for each gender based on the number of Bettable Matches for Singles, so that each grouping represents 10% of all Bettable Matches for the relevant grouping. Bettable matches of unranked players or players with a missing ranking (due to data limitations) were grouped separately.
- Regional Association (see Appendix B)
- Referral Ratio: FTI has calculated and presented in the summary visual and tabular analyses figures representing the ratio of alerts and matches for each combination of alert and match type:

    (1) Match Specific Alerts, Betting Match Alerts, Other Match Alerts: and

    (2) Played Matches, Bettable Matches.

    In the visual analyses, the Referral Ratio is presented as a gold-red colour gradient across ratios from 0% to 1% in eight colour steps.

Presentation of analyses and contribution to REA Chapter 13

To support the analysis presented in REA Chapter 13 and draw insight from the blended TDB, bettable match and alert data sources, FTI produced a series of Tableau dashboards and summary tables. Many of the analyses—whether represented in a visual or tabular from—represent alerts in the context of some match or player attribute. To facilitate analysis of match- and player-level alert activity, FTI created two views of the TDB:

(1) Match-level TDB (row count of 1 per singles or doubles match); and

(2) Player-level TDB (row count of 2 per singles match and 4 per doubles match).[9]

As a result, figures summarising played and bettable matches and alerts in the visual and tabular analysis do not reconcile across the match- and player-level views.

Visual analytics

FTI produced a series of dashboards presenting the analysis in a way that would enable users to view and analyse the underlying data sources by considering a variety of relevant dimensions, across time and at different levels of granularity. This approach ensured that the visuals could drive insight into overall trends, as well as more granular levels of detail.

Given the emphasis on alert activity relative to matches offered for betting, all but one of the dashboards considering matches or alerts in the context of match activity, were presented against the backdrop of bettable matches. In view of interest in Betting Match Alerts, Other Match Alerts, as well as Match Specific Alerts as a whole, ability to toggle between the three was embedded in visuals, enabling users to switch between alert variables of interest in presenting counts of alerts and Referral Ratio calculations.

Further and in line with insight presented in REA Chapter 13, emphasis was on allowing inference to be drawn across level of sport, gender, singles / doubles, and event geography lines, as well as certain player-level characteristics such as age, ranking, nationality and recurrence of alert activity.

Thematically, FTI produced visuals that are:

- Alert-centric: trends overtime by type and level of sport and recurrence of subject player alerts
- Match-specific: matches, alerts and referral ratios viewed in conjunction
- Player-specific: matches, subject player alerts and referral ratios viewed in conjunction

Tabular analysis

In conjunction with the visual analysis, FTI produced a series of summary figures presenting match, alert and referral ratio data at different levels of granularity and across time. A total of 13 different views were presented summarising the key attributes of interest – Played Matches, Bettable Matches, Match Specific Alerts as a whole, Betting Match Alerts, Other Match Alerts, as well as the six corresponding Referral Ratios over the 2009 – 2017 period and broken down by:

- Level of Sport (at various levels of granularity)
- Gender
- Event Location
- Age

[9] Player-level analyses exclude approximately 100 player-matches between 2009 and 2017 whose player identifying information (player website identifier, name, date of birth, age, ranking and nationality) was unavailable at the time the information was gathered from the websites. FTI has established that the vast majority of these occurred at the Lowest Level of sport and that none of the excluded Player-matches, or the corresponding opponent Player-matches of the affected players, relate to alerted matches.

- Ranking
- Nationality (geography)
- ITF Confederation Nationality

Specialised analyses

In addition to its contributions to REA Chapter 13 outlined in the previous sections, FTI supported the Secretariat and Panel with the production of tabular and visual analyses of the match activity of the 34 players who have been the subject of successful disciplinary prosecutions (with full and final determination by an AHO) since 2009. That match activity is stated as at 31 December 2017 (being the last full date of the data scrape exercise described above). One successful prosecution was in respect of a Covered Person who was not a player and therefore has no match activity.

Specifically, REA Chapter 10 includes FTI's interrogation of the TDB setting out activity across the following dates:

(1) Start of TIU investigation
(2) Date of referral to PTIOs
(3) Date of Notice
(4) Date of AHO Decision

For the periods punctuated by these dates, the analysis presents:

- Number of matches played
- Duration (in days) of match activity
- Duration (in days) of play hiatus
- Monthly play averages

Match counts in this analysis were inclusive of byes as indicative of the inclusion of players in tournament draws and their progression to a following round.

FTI prepared a series of visual representations of the timeline of match activity of the above players in relation to the corresponding Referral, Notice and Decision dates for each player. These visuals mirror the tabular analysis and present the key milestones of successful disciplinary prosecutions of each player against a backdrop of the player's match activity in scatter plot format and counting numbers of matches per tournament. The clusters and surrounding gaps (if present) in the visuals facilitate ready and intuitive identification of match activity which might be incongruous with investigative, referral, notice, and decision dates.

FTI CONSULTING™

EXPERTS WITH IMPACT™

About FTI Consulting

FTI Consulting is an independent global business advisory firm dedicated to helping organisations manage change, mitigate risk, and resolve disputes: financial, legal, operational, political and regulatory, reputational, and transactional. FTI Consulting professionals, located in all major business centres throughout the world, work closely with clients to anticipate, illuminate and overcome complex business challenges.

www.fticonsulting.com

APPENDIX A: Field Completeness

Singles Matches 2013-2017 (by Association, non-byes)

Field	ATP	WTA	ITF-M	ITF-W
gender	100.00%	100.00%	100.00%	100.00%
itf confederation	100.00%	100.00%	100.00%	100.00%
level_of_sport	100.00%	100.00%	100.00%	100.00%
l1_dob_range	99.62%	99.84%	88.23%	99.79%
loser_1_age_at_jan_1	99.62%	99.84%	88.23%	99.79%
loser_1_dob	99.62%	99.84%	88.23%	99.79%
loser_1_confed_nat	99.98%	100.00%	99.95%	99.99%
loser_1_nationality	99.98%	100.00%	99.95%	99.99%
loser_1_id	100.00%	100.00%	100.00%	100.00%
loser_1_name	100.00%	100.00%	100.00%	100.00%
loser_1_Ranking	99.76%	99.98%	35.87%	32.48%
loser_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%
prizemoney_usd	82.10%	100.00%	100.00%	100.00%
round	100.00%	100.00%	100.00%	100.00%
round_flag	100.00%	100.00%	100.00%	100.00%
round_number	100.00%	100.00%	100.00%	100.00%
set_count	98.83%	99.99%	100.00%	100.00%
sgl_draw	100.00%	100.00%	100.00%	99.89%
surface	100.00%	100.00%	100.00%	99.89%
tournament_association	100.00%	100.00%	100.00%	100.00%
tournament_country	100.00%	100.00%	100.00%	99.89%
tournamentlocation	100.00%	100.00%	100.00%	100.00%
tournamentname	100.00%	100.00%	100.00%	100.00%
tournamentstartdate	100.00%	100.00%	100.00%	100.00%
w1_dob_range	99.95%	99.92%	96.11%	99.83%
winner_1_age_at_jan_1	99.95%	99.92%	96.11%	99.83%
winner_1_dob	99.95%	99.92%	96.11%	99.83%
winner_1_id	100.00%	100.00%	100.00%	100.00%
winner_1_name	100.00%	100.00%	100.00%	100.00%
winner_1_confed_nat	100.00%	100.00%	99.96%	100.00%
winner_1_nationality	100.00%	100.00%	99.96%	100.00%
winner_1_Ranking	99.95%	99.99%	58.39%	38.61%
winner_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%

Doubles Matches 2013-2017 (by Association, non-byes)

Field	ATP	WTA	ITF-M	ITF-W
dbl_draw	100.00%	100.00%	100.00%	100.00%
gender	100.00%	100.00%	100.00%	100.00%
itf confederation	100.00%	100.00%	100.00%	100.00%
l1_dob_range	99.75%	99.77%	99.99%	99.99%
loser_1_dob	99.75%	99.77%	99.99%	99.99%
loser_1_age_at_jan_1	99.75%	99.77%	99.99%	99.99%
loser_2_dob	99.77%	99.81%	99.99%	99.99%
l2_dob_range	99.77%	99.81%	99.99%	99.99%
loser_2_age_at_jan_1	99.77%	99.81%	99.99%	99.99%
level_of_sport	100.00%	100.00%	100.00%	100.00%
loser_1_nationality	99.97%	99.99%	99.99%	99.99%
loser_1_confed_nat	99.97%	99.99%	99.99%	99.99%
loser_1_id	99.98%	99.99%	99.99%	99.99%
loser_1_name	99.98%	99.99%	99.99%	99.99%
loser_1_Ranking	99.86%	99.98%	91.73%	84.65%
loser_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%
loser_2_nationality	99.93%	99.99%	99.99%	99.99%
loser_2_confed_nat	99.93%	99.99%	99.99%	99.99%
loser_2_id	99.98%	99.99%	99.99%	99.99%
loser_2_name	99.98%	99.99%	99.99%	99.99%
loser_2_Ranking	99.90%	99.98%	91.55%	83.57%
loser_2_rankings_bucket	100.00%	100.00%	100.00%	100.00%
prizemoney_usd	96.37%	100.00%	100.00%	100.00%
round	100.00%	100.00%	100.00%	100.00%
round_flag	100.00%	100.00%	100.00%	100.00%
round_number	100.00%	100.00%	100.00%	100.00%
set_count	95.56%	99.99%	100.00%	100.00%
surface	100.00%	100.00%	100.00%	100.00%
tournament_association	100.00%	100.00%	100.00%	100.00%
tournament_country	100.00%	100.00%	100.00%	100.00%
tournamentlocation	100.00%	100.00%	100.00%	100.00%
tournamentname	100.00%	100.00%	100.00%	100.00%
tournamentstartdate	100.00%	100.00%	100.00%	100.00%
w1_dob_range	99.95%	99.92%	100.00%	100.00%
winner_1_dob	99.95%	99.92%	100.00%	100.00%
winner_1_age_at_jan_1	99.95%	99.92%	100.00%	100.00%
winner_1_confed_nat	100.00%	100.00%	100.00%	100.00%
winner_1_nationality	100.00%	100.00%	100.00%	100.00%
winner_1_id	100.00%	100.00%	100.00%	100.00%
winner_1_name	100.00%	100.00%	100.00%	100.00%
winner_1_Ranking	99.96%	99.99%	95.03%	92.08%

Field	ATP	WTA	ITF-M	ITF-W
winner_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%
w2_dob_range	99.96%	99.92%	100.00%	100.00%
winner_2_dob	99.96%	99.92%	100.00%	100.00%
winner_2_age_at_jan_1	99.96%	99.92%	100.00%	100.00%
winner_2_nationality	99.98%	100.00%	100.00%	100.00%
winner_2_confed_nat	99.98%	100.00%	100.00%	100.00%
winner_2_id	100.00%	100.00%	100.00%	100.00%
winner_2_name	100.00%	100.00%	100.00%	100.00%
winner_2_Ranking	99.96%	99.96%	96.12%	93.82%
winner_2_rankings_bucket	100.00%	100.00%	100.00%	100.00%

Singles Matches 2013-2017 (by Level of Sport, non-byes)

Field	Lowest	Mid	Tour	Grand Slam
gender	100.00%	100.00%	100.00%	100.00%
itf confederation	100.00%	100.00%	100.00%	100.00%
level_of_sport	100.00%	100.00%	100.00%	100.00%
l1_dob_range	92.42%	98.94%	99.93%	100.00%
loser_1_age_at_jan_1	99.78%	99.91%	100.00%	100.00%
loser_1_dob	99.78%	99.91%	100.00%	100.00%
loser_1_confed_nat	99.96%	99.97%	100.00%	100.00%
loser_1_nationality	99.96%	99.97%	100.00%	100.00%
loser_1_id	100.00%	100.00%	100.00%	100.00%
loser_1_name	100.00%	100.00%	100.00%	100.00%
loser_1_Ranking	61.64%	74.88%	99.97%	99.82%
loser_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%
prizemoney	100.00%	100.00%	100.00%	100.00%
round	100.00%	100.00%	100.00%	100.00%
round_flag	100.00%	100.00%	100.00%	100.00%
round_number	100.00%	100.00%	100.00%	100.00%
set_count	99.53%	98.95%	99.05%	99.27%
sgl_draw	99.93%	100.00%	100.00%	100.00%
surface	99.93%	100.00%	100.00%	100.00%
tournament_association	100.00%	100.00%	100.00%	100.00%
tournament_country	99.93%	100.00%	100.00%	100.00%
tournamentlocation	100.00%	100.00%	100.00%	100.00%
tournamentname	100.00%	100.00%	100.00%	100.00%
tournamentstartdate	100.00%	100.00%	100.00%	100.00%
w1_dob_range	97.50%	99.84%	100.00%	100.00%
winner_1_age_at_jan_1	99.85%	99.90%	100.00%	100.00%
winner_1_dob	99.85%	99.90%	100.00%	100.00%
winner_1_id	100.00%	100.00%	100.00%	100.00%
winner_1_name	100.00%	100.00%	100.00%	100.00%
winner_1_confed_nat	99.98%	100.00%	100.00%	100.00%
winner_1_nationality	99.98%	100.00%	100.00%	100.00%
winner_1_Ranking	65.12%	77.24%	100.00%	99.79%
winner_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%

Doubles Matches 2013-2017 (by Level of Sport, non-byes)

Field	Lowest	Mid	Tour	Grand Slam
dbl_draw	100.00%	100.00%	100.00%	100.00%
gender	100.00%	100.00%	100.00%	100.00%
itf confederation	100.00%	100.00%	100.00%	100.00%
l1_dob_range	99.77%	99.90%	100.00%	100.00%
loser_1_dob	99.81%	99.62%	99.92%	100.00%
loser_1_age_at_jan_1	99.81%	99.62%	99.92%	100.00%
loser_2_dob	99.81%	99.67%	99.90%	100.00%
l2_dob_range	99.81%	99.67%	99.90%	100.00%
loser_2_age_at_jan_1	99.81%	99.67%	99.90%	100.00%
level_of_sport	100.00%	100.00%	100.00%	100.00%
loser_1_nationality	99.99%	100.00%	100.00%	100.00%
loser_1_confed_nat	99.99%	100.00%	100.00%	100.00%
loser_1_id	99.99%	100.00%	100.00%	100.00%
loser_1_name	99.98%	100.00%	99.96%	100.00%
loser_1_Ranking	98.40%	99.40%	99.90%	100.00%
loser_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%
loser_2_nationality	99.99%	100.00%	100.00%	100.00%
loser_2_confed_nat	99.99%	100.00%	100.00%	100.00%
loser_2_id	99.99%	100.00%	100.00%	100.00%
loser_2_name	99.98%	100.00%	99.96%	100.00%
loser_2_Ranking	98.37%	99.56%	99.92%	100.00%
loser_2_rankings_bucket	100.00%	100.00%	100.00%	100.00%
prizemoney_usd	100.00%	100.00%	100.00%	100.00%
round	100.00%	100.00%	100.00%	100.00%
round_flag	100.00%	100.00%	100.00%	100.00%
round_number	100.00%	100.00%	100.00%	100.00%
set_count	99.98%	100.00%	100.00%	100.00%
surface	100.00%	100.00%	100.00%	100.00%
tournament_association	100.00%	100.00%	100.00%	100.00%
tournament_country	100.00%	100.00%	100.00%	100.00%
tournamentlocation	100.00%	100.00%	100.00%	100.00%
tournamentname	100.00%	100.00%	100.00%	100.00%
tournamentstartdate	100.00%	100.00%	100.00%	100.00%
w1_dob_range	99.92%	99.98%	100.00%	100.00%
winner_1_dob	99.95%	99.96%	100.00%	100.00%
winner_1_age_at_jan_1	99.95%	99.96%	100.00%	100.00%
winner_1_confed_nat	100.00%	100.00%	100.00%	100.00%
winner_1_nationality	100.00%	100.00%	100.00%	100.00%
winner_1_id	100.00%	100.00%	100.00%	100.00%
winner_1_name	100.00%	100.00%	100.00%	100.00%
winner_1_Ranking	99.04%	99.83%	99.96%	100.00%

Field	Lowest	Mid	Tour	Grand Slam
winner_1_rankings_bucket	100.00%	100.00%	100.00%	100.00%
w2_dob_range	99.96%	99.97%	99.98%	100.00%
winner_2_dob	99.96%	99.97%	99.98%	100.00%
winner_2_age_at_jan_1	99.96%	99.97%	99.98%	100.00%
winner_2_nationality	100.00%	100.00%	100.00%	100.00%
winner_2_confed_nat	100.00%	100.00%	100.00%	100.00%
winner_2_id	100.00%	100.00%	100.00%	100.00%
winner_2_name	100.00%	100.00%	100.00%	100.00%
winner_2_Ranking	99.25%	99.83%	100.00%	100.00%
winner_2_rankings_bucket	100.00%	100.00%	100.00%	100.00%

APPENDIX B: Regional Association and Country Mappings

Regional Association	Country
Asian Tennis Federation (ATF)	Afghanistan
	Bahrain
	Bangladesh
	Brunei
	Cambodia
	China P.R.
	Hong Kong
	India
	Indonesia
	Iran
	Iraq
	Japan
	Jordan
	Kazakhstan
	Republic of Korea
	Kuwait
	Kyrgyzstan
	Laos
	Lebanon
	Malaysia
	Mongolia
	Myanmar
	Nepal
	Oman
	Pakistan
	Philippines
	Qatar
	Saudi Arabia
	Singapore
	Sri Lanka
	Syria
	Taiwan
	Tajikistan
	Thailand
	Turkmenistan
	United Arab Emirates
	Uzbekistan
	Vietnam
	Yemen

Regional Association	Country
Confederation of African Tennis (CAT)	Algeria
	Angola
	Benin
	Botswana
	Burkina Faso
	Burundi
	Cameroon
	Cape Verde
	Central African Republic
	Congo
	Cote D'Ivoire
	Djibouti
	Egypt
	Eritrea
	Ethiopia
	Gabon
	Ghana
	Kenya
	Lesotho
	Liberia
	Libya
	Madagascar
	Malawi
	Mali
	Mauritania
	Mauritius
	Morocco
	Mozambique
	Namibia
	Niger
	Nigeria
	Reunion
	Rwanda
	Senegal
	Sierra Leone
	South Africa
	Sudan
	Swaziland
	Tanzania
	Togo
	Tunisia
	Uganda
	Zambia
	Zimbabwe

Regional Association	Country
Confederacion Sudamericana de Tenis (COSAT)	Argentina
	Bolivia
	Brazil
	Chile
	Colombia
	Ecuador
	Paraguay
	Peru
	Uruguay
	Venezuela

Regional Association	Country
Confederacion de Tenis de Centroamerica Caribe (COTECC)	Antigua & Barbuda
	Aruba
	Bahamas
	Barbados
	Belize
	Bermuda
	Costa Rica
	Cuba
	Curacao
	Dominican Republic
	El Salvador
	Grenada
	Guadeloupe
	Guatemala
	Guyana
	Haiti
	Honduras
	Jamaica
	Mexico
	Netherlands Antilles
	Nicaragua
	Panama
	Puerto Rico
	Saint Vincent and the Grenadines
	Suriname
	Trinidad & Tobago
	Virgin Islands

Regional Association	Country
Oceania Tennis Federation (OTF)	American Samoa
	Australia
	Fiji
	French Polynesia
	Guam
	New Caledonia
	New Zealand
	Northern Mariana Islands
	Papua New Guinea
	Samoa
	Solomon Islands
	Vanuatu
	American Samoa

Regional Association	Country
Tennis Europe (TE)	Albania
	Andorra
	Armenia
	Austria
	Azerbaijan
	Belarus
	Belgium
	Bosnia & Herzegovina
	Bulgaria
	Croatia
	Cyprus
	Czech Republic
	Denmark
	Estonia
	Finland
	France
	Georgia
	Germany
	Great Britain
	Greece
	Hungary
	Iceland
	Ireland
	Israel
	Italy
	Latvia
	Liechtenstein
	Lithuania
	Luxembourg
	Macedonia
	Malta
	Moldova
	Monaco
	Montenegro
	Netherlands
	Netherlands Antilles
	Norway
	Poland
	Portugal
	Romania
	Russia
	San Marino
	Serbia and Montenegro
	Slovakia
	Slovenia
	Spain
	Sweden
	Switzerland
	Turkey
	Ukraine

Grouping	Country
North America*	Canada
	U.S.A.

** Canada and U.S.A. are not members of an association, but have been mapped to 'North America' for the purposes of the analysis.*

Unaffiliated	Country
Other	French Guyana
	Martinique