

PBS Software Vendor Webinar

The PBS Data API data model

Thank you for joining, this webinar will begin at 2.00pm,
24 July 2023



Acknowledgement of Country

We acknowledge the Traditional Custodians of country throughout Australia and their connections to land, sea and community.

We pay our respect to their Elders past and present and extend that respect to all Aboriginal and Torres Strait Islander people on this webinar today.



Australian Government

Department of Health and Aged Care

Before we begin...



Please turn off your webcam



Please mute your microphone unless speaking



Any questions? Email HPP.Support@health.gov.au



Agenda

- PBS Data Distribution Project model background and derivation
- Introduction to the API data model
 - Schedules
 - Items
 - Dispensing and Prescribing Rules
 - Restrictions
 - Summary of Changes



PBS Data Distribution Project

Background

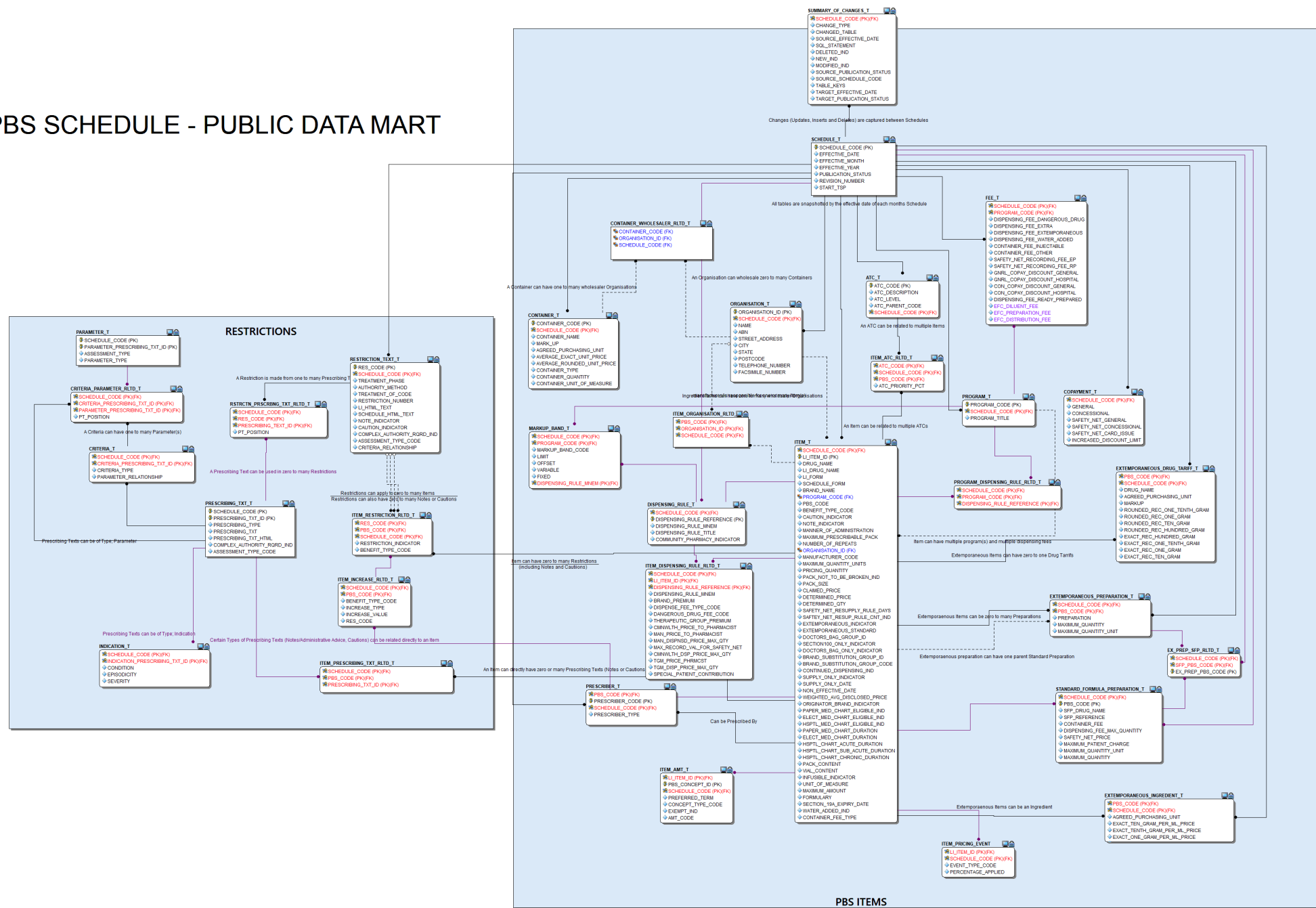
The Department of Health is modernising the consumption and data distribution for the monthly PBS Schedule data.

To achieve this, it was decided to revisit the underlying XML/Text file distribution system and deliver the data in a more relational format via the use of an API.

The data was remodelled into a relational data model, and a series of endpoints designed to deliver the data for each entity/table in the model.

As per standard modelling rules; data lives in its most natural location, with associative tables resolving any many-to-many relationships, and reference tables used to reduce redundant storage.

PBS SCHEDULE - PUBLIC DATA MART



PBS Data Distribution Project

Data Model – Schedule Table

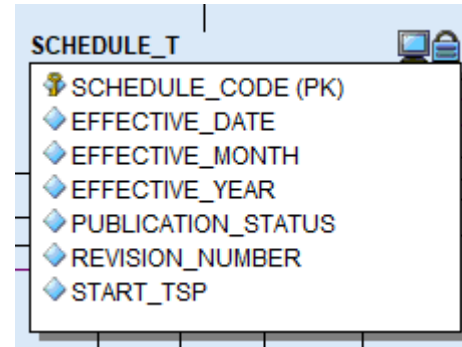
The data model has been designed to act as a repository of data, rather than a transactional or operational database. It is recommended that the endpoint responses be loaded into a local repository for interrogation and use.

The Schedule table is the main controlling entity, and every table can be segmented into data specific to the relevant schedule.

The Schedule Code is then used as part of the Primary Key of all child tables across the model.

The table stores the Schedule Code, with the effective dates and publication status.

The status itself is used to indicate whether the current Schedule is under embargo.



SCHEDULE_T

🔑 SCHEDULE_CODE (PK)
◆ EFFECTIVE_DATE
◆ EFFECTIVE_MONTH
◆ EFFECTIVE_YEAR
◆ PUBLICATION_STATUS
◆ REVISION_NUMBER
◆ START_TSP

PBS Data Distribution Project

Data Model – Schedule Table

```
revision_number,start_tsp,effective_date,effective_month,effective_year,publication_status,schedule_code
"1","2023-03-01T00:00:00.000+11:00","2023-03-01","MARCH","2023","PUBLISHED","2890"
"5","2023-04-01T00:00:00.000+11:00","2023-04-01","APRIL","2023","PUBLISHED","3616"
"3","2022-09-01T00:00:00.000+10:00","2022-09-01","SEPTEMBER","2022","PUBLISHED","3411"
"4","2022-11-01T00:00:00.000+11:00","2022-11-01","NOVEMBER","2022","PUBLISHED","3455"
"1","2023-01-01T00:00:00.000+11:00","2023-01-01","JANUARY","2023","PUBLISHED","2830"
"6","2022-12-01T00:00:00.000+11:00","2022-12-01","DECEMBER","2022","PUBLISHED","3477"
"3","2023-06-01T00:00:00.000+10:00","2023-06-01","JUNE","2023","PUBLISHED","3675"
"3","2023-02-01T00:00:00.000+11:00","2023-02-01","FEBRUARY","2023","PUBLISHED","3532"
"3","2022-08-01T00:00:00.000+10:00","2022-08-01","AUGUST","2022","PUBLISHED","3353"
"6","2022-10-01T00:00:00.000+10:00","2022-10-01","OCTOBER","2022","PUBLISHED","3436"
"3","2023-05-01T00:00:00.000+10:00","2023-05-01","MAY","2023","PUBLISHED","3651"
"4","2023-07-01T00:00:00.000+10:00","2023-07-01","JULY","2023","PUBLISHED","3751"
"1","2023-08-01T00:00:00.000+10:00","2023-08-01","AUGUST","2023","EMBARGO","3071"
```



PBS Data Distribution Project

Data Model – Domains

Even though it still looks complex, the model has been designed to deliver data in two main domains:

- Items – the data related to the PBS Items themselves
- Restrictions – the data related to the rules used for the prescribing of PBS Items

PBS Data Distribution Project

Data Model – Item Domain

The Item table holds data at the Trade Product Pack (TPP) AMT level:

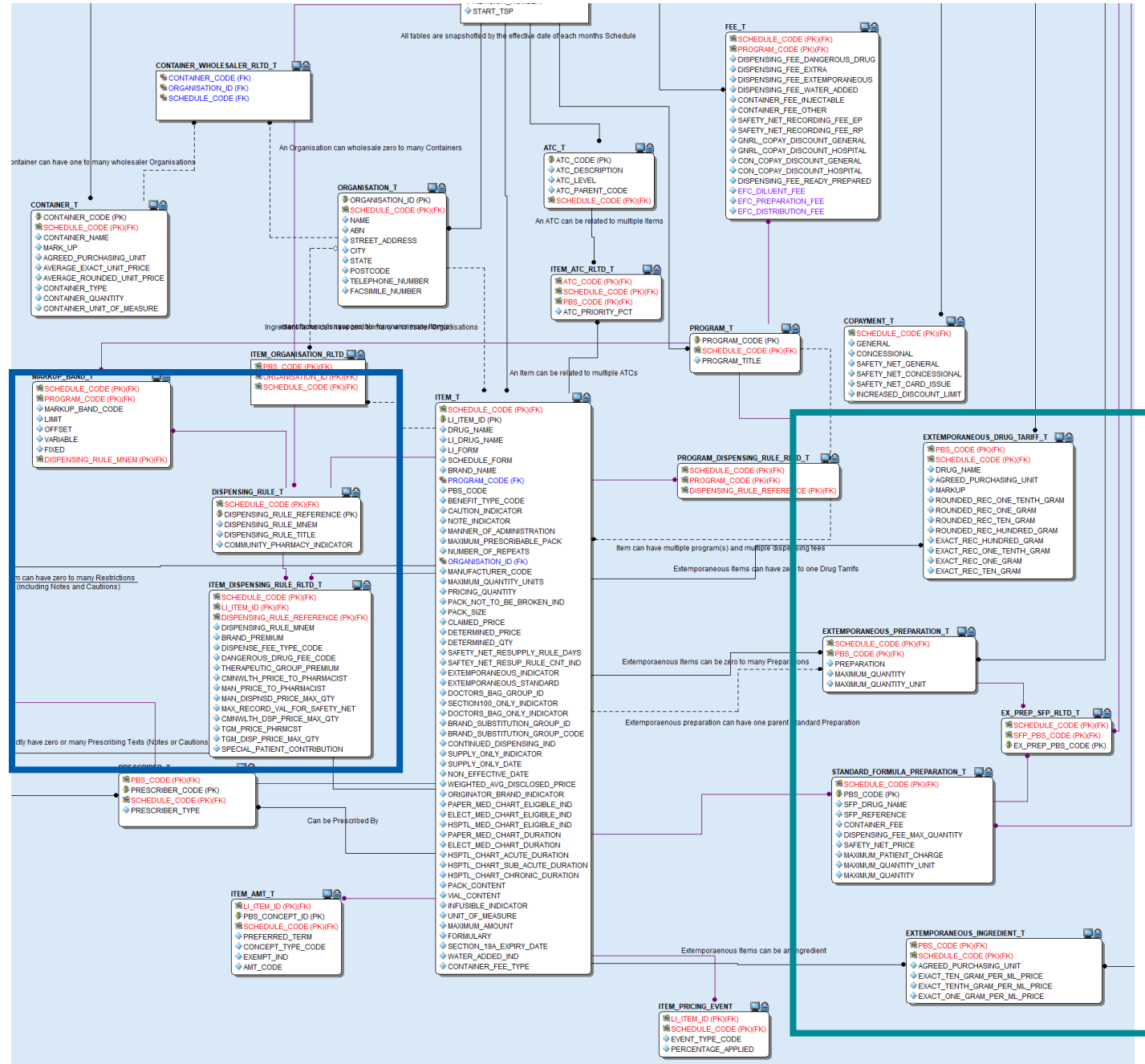
- One record per Drug, Form and Brand. Both Ready Prepared and Extemporaneous items are stored in this table
- An identifier (LI_ITEM_ID) has been added to the table to force a single unique identifier onto the records. LI_ITEM_ID is a concatenation of the PBS_CODE; PIG_ID (Pharmaceutical Item Group); MP_ID; MPP_ID and TPP_ID
- Much of the data then relates to the Item records using established relationships, i.e. Dispensing Rules; Extemporaneous Items; Prescribers; etc.
- Various reference tables are also included such as Fees; Program; Markup Band; Container; Organisation; etc.



Item Domain

As an example:

- Additional data relating to Extemporaneous Items can be found in the tables marked in Green;
- Data relating to Dispensing Rules is marked in Blue.



PBS Data Distribution Project

Data Model – Restrictions domain

Anybody who has attempted to piece together the Restrictions using the XML would understand the complexity. With this in mind, the Restrictions domain has been designed to deliver the data in the most effective way possible. Restriction data is delivered in two ways:

- The fully formed Restriction text is found in the Restriction endpoint;
- With the fragmented text, as found in the XML currently, located in the Prescribing Text endpoint.

A document has been included on the HPP Website that delves deeper into the Restrictions and explains how to use the data model to piece together a complete Restriction using the fragments as delivered through the endpoints.

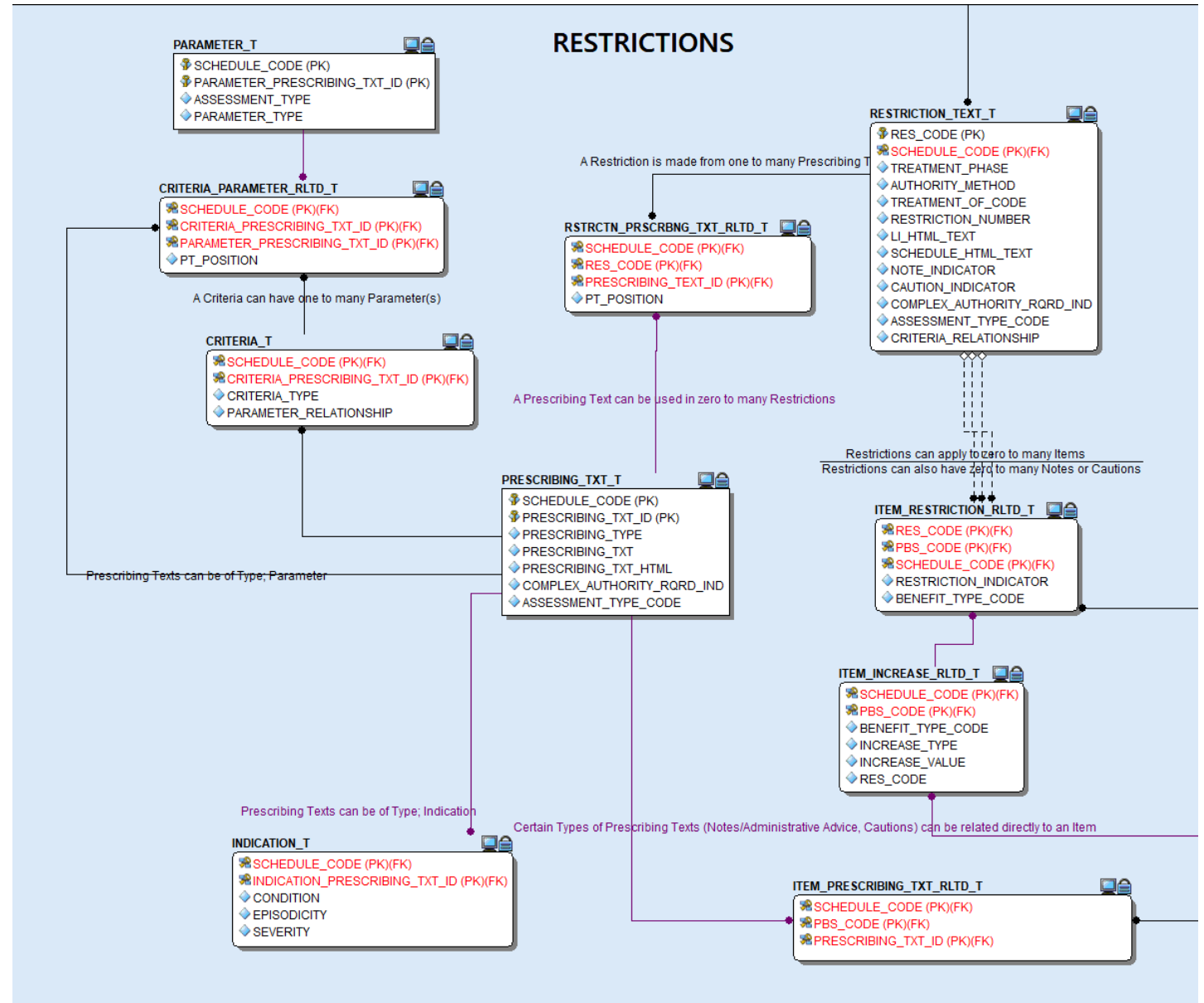


Restrictions domain

The Restriction text table holds both the full text for the Schedule version and Legal Instrument version of a restriction in marked-up format ready for display through rich text field.

Using the Restriction_Prescribing_Rltd table, the fragments can be identified. These are available in raw text and markup versions.

Further granular info is available through the Indication and Criteria tables.



PBS Data Distribution Project

Data Model – Summary of Changes

To assist Users with consuming the data available in the API, a Summary of Changes endpoint has been included which contains a catalogue of all Inserts, Deletes, and Updates to the previous schedule.

- Each record includes the source Schedule Code; the affected Entity; the type of change (i.e. Insert, Update, Delete), the Keys used to form the “Where” clause; and the actual SQL Statement;
- The SQL Statements themselves can be used to apply all changes against the previous Schedule data to create the new Schedule.

SOURCE SCHEDULE CODE	CHANGED TABLE	CHANGE TYPE	TABLE_KEYS	SCHEDULE CODE
3651	ITEM_T	UPDATE	"SCHEDULE_CODE":"3675","LI_ITEM_ID":"10620Y_20518_375_39453_39459"	3675
3651	ITEM_T	UPDATE	"SCHEDULE_CODE":"3675","LI_ITEM_ID":"9356K_11492_1310_5739_6966"	3675
3651	ITEM_AMT_T	DELETE	"SCHEDULE_CODE":"3675","PBS_CONCEPT_ID":"38398","AMT_CODE":"","LI_ITEM_ID":"10693T_17636_143_38396_38398"	3675
3651	ITEM_AMT_T	DELETE	"SCHEDULE_CODE":"3675","PBS_CONCEPT_ID":"38400","AMT_CODE":"","LI_ITEM_ID":"10683G_17643_38400_38403_38405"	3675



PBS Data Distribution Project

Data Model – Summary of Changes – Example

Below is the SQL Statement from an Item insert that occurred for Schedule 3651, and the Table_Keys entry, which in this case shows the Schedule Code, and the LI_ITEM_ID for the new record.

```
INSERT INTO ITEM_T VALUES ('3675' , '12979D_9505_1344_5821_56349' , 'lenalidomide' ,  
'Lenalidomide' , 'Capsule 25 mg' , 'lenalidomide 25 mg capsule, 21' , 'Lenalidomide Viatrix' ,  
'HB' , '12979D' , 'A' , 'Y' , 'Y' , 'ORAL' , '1' , '21' , '2' , '142' , 'AF' , '21' , '14' , 'N' , '2397.21' ,  
'1598.14' , 'N' , '20' , 'N' , 'N' , 'null' , 'null' , 'Y' , 'N' , 'null' , 'null' , 'N' , 'N' , 'null' , 'null' , 'null' ,  
'N' , 'N' , 'N' , 'Y' , 'null' , 'null' , '12' , '12' , '12' , 'null' , 'null' , 'N' , 'null' , 'null' , 'F2' , 'N' , 'null'  
, 'other');
```

```
TABLE_KEYS: "SCHEDULE_CODE":"3675","LI_ITEM_ID":"12979D_9505_1344_5821_56349"
```



PBS Data Distribution Project

Data Model – Summary of Changes – Embargo

Prior to publishing, the Schedule data may be revised several times whilst still under Embargo. This results in new Schedule Codes being assigned to the revised Schedules, and a new Revision number being assigned to the record.

Revised Schedules may appear in the data during the Embargo period, but previous revisions will be removed once the Schedule is published.

The Summary of Changes will also reflect this via the Source Schedule Code and Schedule Code fields indicating which set of data the changes should be applied to.

PBS Data Distribution Project

Data Model – Summary of Changes – Embargo Schedules

```
revision_number,start_tsp,effective_date,effective_month,effective_year,publication_status,schedule_code
"1","2023-03-01T00:00:00.000+11:00","2023-03-01","MARCH","2023","PUBLISHED","2890"
"5","2023-04-01T00:00:00.000+11:00","2023-04-01","APRIL","2023","PUBLISHED","3616"
"3","2022-09-01T00:00:00.000+10:00","2022-09-01","SEPTEMBER","2022","PUBLISHED","3411"
"4","2022-11-01T00:00:00.000+11:00","2022-11-01","NOVEMBER","2022","PUBLISHED","3455"
"1","2023-01-01T00:00:00.000+11:00","2023-01-01","JANUARY","2023","PUBLISHED","2830"
"6","2022-12-01T00:00:00.000+11:00","2022-12-01","DECEMBER","2022","PUBLISHED","3477"
"3","2023-06-01T00:00:00.000+10:00","2023-06-01","JUNE","2023","PUBLISHED","3675"
"3","2023-02-01T00:00:00.000+11:00","2023-02-01","FEBRUARY","2023","PUBLISHED","3532"
"3","2022-08-01T00:00:00.000+10:00","2022-08-01","AUGUST","2022","PUBLISHED","3353"
"6","2022-10-01T00:00:00.000+10:00","2022-10-01","OCTOBER","2022","PUBLISHED","3436"
"3","2023-05-01T00:00:00.000+10:00","2023-05-01","MAY","2023","PUBLISHED","3651"
"4","2023-07-01T00:00:00.000+10:00","2023-07-01","JULY","2023","PUBLISHED","3751"
"1","2023-08-01T00:00:00.000+10:00","2023-08-01","AUGUST","2023","EMBARGO","3071"
```



PBS Data Distribution Project

Data Model – Summary of Changes – Embargo Schedules

Example: during embargo period (before 1 August 2023)

Schedule code	Revision number	Effective Date	Status	SOC records	SOC relative to
3001	1	1-Jul-2023	Published		
3005	1	1-Aug-2023	Embargo	1000	3001
3008	2	1-Aug-2023	Embargo	20	3005
3010	3	1-Aug-2023	Embargo	5	3008

Example: after embargo period (from 1 August 2023)

Schedule code	Revision number	Effective Date	Status	SOC records	SOC relative to
3001	1	1-Jul-2023	Published		
3010	3	1-Aug-2023	Published	1025	3001



PBS Data Distribution Project

Data Model – Documentation

To assist Users with downloading and using the API endpoint data, several documents can be accessed from the HPP website.

API (data dictionary and database diagram): [HPP Article Details](#) · [API Data Dictionary](#)

Restrictions: [Piecing together Restrictions fragments](#)

PBS Data Distribution Project

Comparisons to the existing download methods

The data held in the API has been aligned closely with the XML and an internal reporting system checks the veracity of the API data against the expected XML outputs to maintain this.

The text files are based on the XML and due to the number of Users still accessing the text files, a task is underway to map the API data to the text file data and provide guidance on how to recreate the text files from the API endpoint data.

Resources

- Project page: [Data Distribution Project | PBS Data](#)
- API (data dictionary and database diagram): [HPP Article Details · API Data Dictionary](#)
- Restrictions: [Piecing together Restrictions fragments](#)

Questions?



Australian Government
Department of Health and Aged Care

Any feedback during the public beta can be provided
via email to HPP.Support@health.gov.au