RETAIL ENERGY CODE

tThis document is classified as **WHITE**.
Information is for public, unrestricted dissemination and publication.

REFERENCE	**RECPAB_26_0224_11**
STATUS	**FOR DISCUSSION**
PAGES	**9**

DEEP DIVE ON DATA CLEANSE

BY ANTON MODEN, PERFORMANCE ASSURANCE LEAD

REVIEWED BY WALTER CARLTON, LEAD PERFORMANCE ASSURANCE PARTNER, REC CODE MANAGER

FOR DISCUSSION	This paper sets out the responsibilities of PAB relating to data cleanse within the REC. This includes data that is within REC systems, or updated by REC processes. Typically this includes information updated by Suppliers and their metering agents (for example metering data, address data and meter point status). In some cases DNOs or GTs are also responsible, for example in relation to some address data. It sets out enhancements to the approach to enable estimates of costs and benefits. It includes in the appendix estimates of the benefits of data cleansing, as well as the process by which we intend on validating these estimates and seeking estimates of the cost of cleansing data from Parties. We are seeking input from the PAB on the approach at this stage and wish to **DISCUSS** the content.

1 INTRODUCTION

The REC contains a range of requirements in relation to maintaining data, and is also impacted by data management responsibilities within other Codes. Despite these requirements it has proven necessary to operate a Data Cleanse process (and similar processes have also been included in most recent industry change programmes) as in reality data is not always maintained as required. Making sure that data cleanse activities are performed consistently is a key aspect of ensuring that actions taken by Parties do not conflict and that Parties that do cleanse data do not suffer a 'compliance penalty'. This is particularly important where data quality issues are transferred from one business to another when a consumer switches.

Data cleanse is therefore a critical part of the Performance Assurance approach. There are currently four data cleansing activities under the REC:

1. Performance Assurance data cleanse, with the focus and targets set by the PAB. The Code Manager intervenes with Performance Assurance Techniques (PATs) where Parties do not progress against targets, overseen by PAB.
2. Address data cleanse, currently overseen by DCC, however from 27 February 2026 R0283 will be implemented which creates a role for the PAB to establish Code level targets, prioritising them and overseeing any appropriate application of PATs.
3. In response to technology issues, we may ask Parties to resolve data issues. This follows the data issue resolution process set out in *RECPAB_25_0826_05 – Data Error Resolution Process.* These issues are inherently unpredictable, and the level of effort to resolve them can vary substantially.
4. During the Supplier of Last Resort process there is often significant data cleanse needed before the accounts of customers can be transferred to the new Supplier, and billed so that the failing Supplier's administration can be completed. Most Suppliers also switch consumers to their existing businesses, which can identify data issues.

1.1 ROLE OF THE PAB

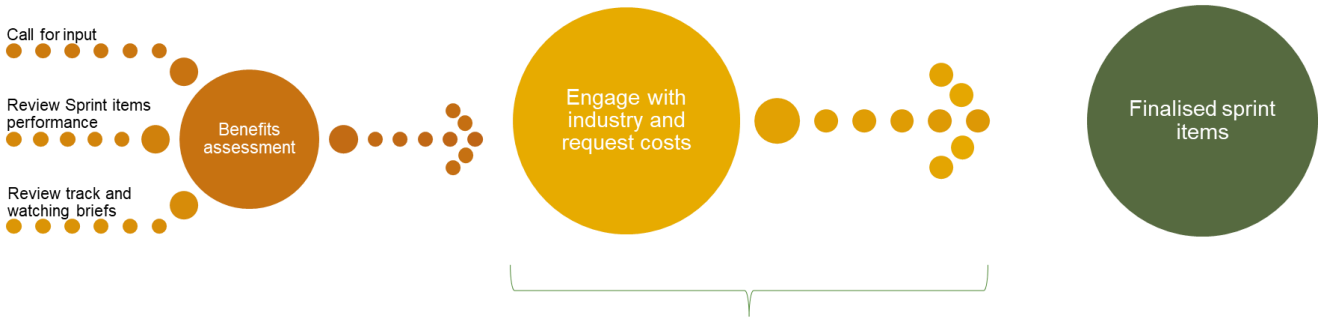
Whilst a vital activity, PAB needs to balance the benefits of cleansing data alongside the impacts on Parties. In particular, the pace of data cleanse, amongst other demands on Parties, is a critical consideration. PAB's role is to prioritise and set appropriate targets for cleansing. In doing so it should consider:

1. The benefits of cleansing on REC Parties, consumers and other impacted stakeholders.
2. The aggregate effort required by Parties to cleanse the data.
3. Impacts on other parts of the industry (e.g. some recent REC data cleanse has benefited MHHS).
4. If one Party alone can resolve an issue, or if several must cooperate to solve the issue.
5. How best to structure targets (e.g. the successful 6 month sprint approach).
6. If cleansing is the right approach – poor data quality often indicates the need to change Code processes/requirements to prevent data misalignment from occurring in the first place.

To do so the Code Manager has previously focused on industry engagement of multiple data items, asking for Party input on prioritisation. We propose to enhance this by estimating the costs and the benefits of cleansing items.

1.2 PROPOSED ENHANCED APPROACH

To further develop our approach for the next data cleanse sprint (Sprint 6, expected to start in May) we propose to make estimates of costs and benefits of any proposed data cleanse activity. We have set out the proposed approach below:



Key new step

1.3 APPROACH TO BENEFITS IDENTIFICATION

To quantify the benefits of cleansing data, we have pinpointed issues that could arise if the specific data item is inaccurate, and estimated associated costs. The total cost of resolving all identified issues becomes the estimated total benefit to cleanse this data item.

There may be multiple problems caused by inaccurate data, even with multiple organisations impacted. As an example, we have set out below some of the issues that having a site recorded as energised that does not have data on the meter in industry systems.

Unbilled consumption	• Simply, if the meter point is energised but not metered, a consumer could be using energy but not being billed for it.
Switch blocking	• A consumer may have a meter, which is read and accurately billed, but the data is missing in industry systems. This could mean when they attempt to switch they can't, as central data is used to price and validate switches.
Asset management challenges	• If there isn't a record of the meter, there may be issues on exhange, fault or removal. These could be exacerbated if the asset is confused for a nearby meter, causing issues for others.
Potential asset loss	• Meter Asset Providers (MAP) finance meters, then bill Suppliers for the use of this asset. The loss of this asset data may mean the MAP loses this rental, effectively losing the asset value. This can also involve not replacing a meter when needed, as the meter data is not known.
Additional administrative costs	• Meter reads, may involve additional steps. It may be neccesary to visit the premesis to identify the meter details, or confirm the site is de-energised.
Customer billed based on estimates	• Customers could be billed based on estimates such as EACs and AQs, irrespective of if they are consuming energy or not.

The impact of many of these issues will vary significantly on a case by case basis. For example, the costs of unbilled consumption are highly related to the length of time it was unbilled.

We have made the following simplifying assumptions:

- We have made estimates using sources where possible, but in cases where we have no information on the costs these issues cause, we will assume a notional rate of £100 per data item.
- We have not estimated settlement impacts of data issues. We would welcome input from the BSC and UNC if relevant, as part of our consultation process.
- Some estimated costs for identified issues cover a period of time. For these items that span a period we have estimated the annualised cost. Other costs arise when a specific event takes place, for example if there are data issues and a switch request happens, there may be specific costs that then arise. These costs are not annualised, but we attempt to estimate how many times this would occur during a year so that the costs are comparable.
- Benefits of resolution are rounded, to avoid false precision.

1.4 IDENTIFIED BENEFIT AVAILABLE FROM CLEANSE

We have included as Appendix 1 estimates of the benefits available from the Sprint 5 data cleanse items. The total identified benefit is set out below.

Data item	Party Responsible	Total benefit (£m)
Energised no meter	Supplier	6
Invalid topline	Supplier & DNO	0.4
Missing ESME ID	Supplier (both current and installing)	4
Missing Meter Location	Supplier	0.4
Related MPAN with inconsistent Domestic Premises Indicator (DPI)	Supplier	0.1
Unmeasured MPAN with Domestic DPI	Supplier	0
Gas CDSP Reconciliation	MAM	26.2
MEM-EES Reconciliation Differences	MOA	TBC

This indicates that the most benefit arises from the Energised no Meter, Missing ESME ID and the Gas CDSP reconciliation, although two of these are the reports with the largest number of items to cleanse, and will have correspondingly higher costs to resolve.

Several of the reports have a very low benefit, as these have been cleansed effectively during the period, significantly reducing the population with identified data quality issues.

1.5 VALIDATING THIS INFORMATION WITH INDUSTRY

Parties will have better information on the benefits of cleansing data and in the associated costs of doing so. We therefore intend on consulting with Parties on these items.

In the coming month we plan on undertaking the following activities:

Activity	Due Date
Issue a consultation, to improve the accuracy of the benefits identified and identify the associated costs	27 Feb
Consultation close	17 Mar
Feedback consultation findings to Parties at the PA webinar, providing a further opportunity to provide input	19 Mar
Analyse the results, alongside progress on the sprint, preparing a recommendation for PAB	24 Mar

1.6 ADDRESS DATA CLEANSE DEVELOPMENTS

Following the change R0283 we anticipate that DCC will provide information to PAB to enable PAB and the Code Manager to set targets for industry.

Our expectation is that the 2026-27 Address Quality Plan will have include elements data which should be prioritised, but that this will not be in the form of targets that can be formally adopted. We expect that over the coming year these will be enhanced to enable the population of the Address Prioritisation Matrix described in the Code. We anticipate this will be an input into future Sprints, but that this will not be ready for Sprint 6.

Should progress exceed expectations, we can adapt our approach to consider the extent of address data cleanse alongside other data cleanse activities, to make better prioritisation decisions.

1.7 PAB ACTION

PAB ACTION

We are seeking input from the PAB on the approach at this stage and wish to **DISCUSS** the content.

1.8 APPENDIX 1 ESTIMATES OF BENEFITS OF CURRENT DATA CLEANSE ITEMS

We set out below the estimates of the benefits of Sprint 5 data cleanse items. Where possible we have used the same estimate for similar instances. In aggregating the data we have rounded the benefits, to avoid false levels of precision.

The Code Manager has estimated both the cost of an issue, and its incidence within the data set.

Data item	Party Responsible	Beneficiary	Issue avoided	Cost per issue	Source of estimate	Estimated incidence	Current number	Benefit to industry to address (£m)	Total benefit (£m)
Energised no meter	Supplier	Supplier	Unbilled Consumption	£879	Price cap average bill[1]	15%	18,435	2.4	6
		Consumer	Switch Blocking	£88	5% saving on Price cap annual bill[2]	15%	18,435	0.2	
		Supplier	Asset management challenges	£100	Default Estimate	50%	18,435	0.9	
		MAP	Potential asset loss	£200	Code Manager Estimate	15%	18,435	0.6	
		Supplier	Additional administrative costs	£100	Default Estimate	100%	18,435	1.8	

[1] Energy price cap explained | Ofgem. Here we assume that only one of the meters is not registered, and take half the annual value of £1758.

[2] Assuming a 5% price below the price cap is achievable. This varies significantly over time, with smaller suppliers offering up to 13% discounts currently, with larger Supplier discounts currently closer to the price cap.

Data item	Party Responsible	Beneficiary	Issue avoided	Cost per issue	Source of estimate	Estimated incidence	Current number	Benefit to industry to address (£m)	Total benefit (£m)
		Consumer	The customer could be billed based on estimates, when no consumption exists. We have assumed this would be rectified, but create administrative issues	£100	Default estimate	5%	18,435	0.1	
Invalid topline	Supplier & DNO	Consumer	Billing error due to incorrect line loss factor	£44[3]	Price cap and Code Manager estimate	50%	9,168	0.2	0.4
		Consumer	Loss of access to a time based tariff (for non-smart meters) through inaccurate standard settlement configuration data	£50	Code Manager Estimate	50%	9,168	0.2	
Missing ESME ID	Supplier (both current and installing)	Supplier	Cannot communicate with the meter	£100	Default Estimate	5%	539,147	2.7	4
		Consumer	Loss of access to a time based tariff	£50	Code Manager Estimate	5%	539,147	1.3	

[3] We are assuming the error is 5% of the bill. For simplicity we have taken the electricity bill being half the total price cap bill. Published rates are slightly higher for electricity than gas, but not significantly). We'd assume some would underpay, some overpay, but both create issues for competition.

Data item	Party Responsible	Beneficiary	Issue avoided	Cost per issue	Source of estimate	Estimated incidence	Current number	Benefit to industry to address (£m)	Total benefit (£m)
Missing Meter Location	Supplier	Consumer	Missed appointment, as meter cannot be located.	£100	Default Estimate	5%	87,207	0.4	0.4
Related MPAN with inconsistent Domestic Premises Indicator (DPI)	Supplier	Supplier	Unbilled Consumption	£440	Price cap average bill[4]	5%	959	0	0.1
		Consumer	Switch Blocking	£88	5% saving on price cap annual bill	100%	959	0.1	
Unmeasured MPAN with Domestic DPI	Supplier	Supplier	Unbilled Consumption	£440	Price cap average bill	5%	8	0	0
Gas CDSP Reconciliation	MAM	Consumer	Switch Blocking	£88	5% saving on price cap annual	2%	1,474,745	2.6	26.2
		Consumer	Billing error due to incorrect meter details	£44[5]	Code Manager Estimate	25%	1,474,745	16.2	
		Supplier	Additional administrative costs	£100	Default Estimate	5%	1,474,745	7.4	

[4] We've assumed for a related meter that one of the related meters is not being accurately billed.
[5] We have assumed meter technical detail issues in gas have the same impact as electricity topline issues, using the same value here.

Data item	Party Responsible	Beneficiary	Issue avoided	Cost per issue	Source of estimate	Estimated incidence	Current number	Benefit to industry to address (£m)	Total benefit (£m)
MEM-EES Reconciliation Differences	MOA	Consumer	Switch Blocking	£88	5% saving on price cap annual	50%	TBC[6]	TBC	TBC
		Consumer	Billing error due to incorrect meter details	£44[7]	Code Manager Estimate	50%	TBC	TBC	
		Supplier	Additional administrative costs	£100	Default Estimate	25%	TBC	TBC	

[6] Our current understanding of the population is incomplete, as this is dependent on data submission from MOAs, however not all MOAs have submitted data. We intend on updating this estimate as our understanding of the volume of issues improves.