

DATA COMMITTEE

- Quality
- Protection
- Integrity
- Value

Eduardo Aquino, Data Management Office

December 16th, 2020



BNP PARIBAS
PERSONAL FINANCE



Terms of reference

Frame	Committee Content		
Objective	DATA Roadmap presentation		
Frequency Date	Quarterly	Date: 16/12/2020	Last committee: 23/09/2020
Agenda	Eduardo Aquino presentation: 1,5 h: 1 – Run: Data Comex Dashboard Highlights last 3 months 1.1 - Zoom Inconsistencies on Bank of Portugal Reports CRC/EMF/Consumer Credit 2 - Data Roadmap 2021		2.1. - New Data Warehouse – Implemented 2.2. – Zoom Phase 1 & 2 2.3. – Zoom Phase 3 & 4 3 - Data Resources (Skills/profiles) 3.1 - Zoom IT Data Resources
Executive Summary / Decision to be taken	<p>Executive Summary:</p> <ul style="list-style-type: none"> • GDPR/Protection: Remediation actions of the GDPR Assessments are globally implemented/on tracking; Ongoing subset of the Elimination Project, that will delete 2nd titular information historical when is not necessary. • RaDAR/Quality: Quality principles are strongly present in NDoD and New DW implementation especially due to usage of new Data dictionary 2.0., end-to-end quality checks and golden source identification. Important Data Risk flows are now cover by Data Quality principles. Decommissioning of DWH will be a big step in 2021 for Data Quality principles coverage given the amount of information and flows that will be covered. More difficulties in Central variables coverage: scope was reduced but multidisciplinary meetings are ongoing to remediate the situation. • Data strategy: in order to respond CEO challenge, to bring more ambition to data monetization in parallel of DWH decommissioning, we propose a Data Lake activation anticipation. Actualized 2021 Roadmap will be presented during the committee for validation responding to those ambitions. Data KPIs have been aligned with PF Forward+. Next steps: Pursue technical steps (F1/F2), DMO will have to lead a Data Lake governance workshops in January, with the main involved stakeholders (IT/Data community) and continue to list the current data exploitation tools by the end of January. • Resources: The first phase of Data Skills/ Data Strategic Workforce Planning – Gap Analysis is performed. Up-skill and Roll-out are schedule to 2022/4. <p>Decision to be taken:</p> <ol style="list-style-type: none"> 1. Business ownership of DW Asset: given the transversal usage of the DW the proposal is that CDO or direct manager should be the business owner; 2. Pilotage of Data Monetization: To ensure the transversal use of Data exploitation and Forward+ KPI follow-up, the proposal is that DMO should assume that responsibility in parallel on its current duties. 3. 2021 Data Roadmap validation, slide 7; 4. Collaborative approach, slide12-13. 		

Attendees		Attend	Attendees		Att.	Attendees		Attend	Attendees		Att.
CEO	Chair		CIO	Core member		Dir. Compliance	Core member		Resp. Credit Risk Analytics	Optional member	
CDO	Pilot/Secretary		Dir. Jurídico e R.	Core member		DCIO	Core member		Data Community	Optional member	
Dir. TDPM	Core member		COO	Core member		DPC	Core member				
CRO	Core member		CMSO	Core member		Dir. ADM	Optional member				
CFO	Core member		Dir. RH	Core member		Resp. Backend	Optional member				

Decisions / Actions log

Decisions | September 23rd 2020

Decisions	Status
Data protection processes (ROPA) assessments action plan, allocations areas resources – A first pre-assessment step was performed by DMO as a result 8 ROPA have been detected with High risk. DMO with support of CISO will schedule workshops with those 8 areas in order to perform full DPIA according Plan in slide 8.	✓ Validated
Data Strategy Step1 & Step2 (Current DWH Analysis & Load Source Data) validated. A Mandat will have to be presented Asap. However, CEO asked to have a more aggressive planning for those technical steps.	✓ Validated
Data Strategy Step 3 (DWH Decommission) timeline was not validated. As for step 1&2 CEO asked to have a more aggressive planning. CEO also challenged teams to be able to ensure to develop Business Data projects in parallel of DWH Decommissioning process. Therefore a Data resources skills GAP analysis will have to be performed in order to increase overall Data capacity (IT + Business) and anticipate future ramp-up, upskilling needs asap. As a consequence a New Roadmap will have to be presented before end of year.	■ pending

Actions | Closed (27), On Going (4), On Hold (2), To Start (0)

#	Actions	Date	Owner	Status
33	Transversally identify Data roles and key resources, map them (location and numbers, seniority and skills) and determine the gaps with current needs in order to accelerate the Data Strategy implementation 11/12/2020: Data skills mapping done, please see detailed information in slide 11.	31/12/2020	D. TDPM + D.ADM + HR	■ Closed
32	Data Strategy Data Vision : Schedule a meeting with CFO and COO to go deeper on the Strategy and architecture explanation 4/11/2020: meeting performed and new DW Architecture clarified.	31/10/200	CDO + Resp. Backend	■ Closed
31	Action in “Data Breach on personal data of recovery portfolio (July)” incident: send ASAP the letter to the client via direct courier. 13/10/2020: Letter delivered directly to the Client.	ASAP	DMO	■ Closed
30	Define new Data Quality KPI: Cross Data incidents route cause with projects with Data assessments. 9/12/2020: New KPI was put in place, showing the #Data incidents in assets covered by implemented recommendations.	31/12/2020	DMO	■ Closed

Actions log

#	Actions	Date	Owner	Status
29	In line with Action #6: Present Shadow IT activity results to Comex in ad-hoc meeting to be scheduled. 30/11/2020: Done By CISO in November IT Risk Committee.	31/10/2020	CDO+CISO	■ Closed
28	Define and implement Modus Operandi to enforce the DMO added value with deeper analysis capacity. 11/12/2020 Data incidents resolution with more DMO intervention; given DMO restructuration, on hold operational trainings (RH) covering the main systems and operational processes.	31/12/2020	DMO	■ On hold
26	Data incident BdP: misalignment between CRC/EMF/Crédito Consumidores . Incident remediation lacks of efficiency. DMO will analyze how the incident management process could be improved, and support current teams 11/12/2020: please see overview update slide 6.	30/6/2020 31/12/2020	CDO/head of Statistics/ head of CRA	■ On going
23	Plan B for Data Management Platform strategy has to be deepened analysis with IT and Data community to propose new solution proposal to Comex 11/12/2020: Please see slide 7.	31/7/2020 16/12/2020	CDO	■ On going
21	Data Incidents follow up operationalization - Data Incidents follow up operationalization will be challenged with a Kaizen Ace/BPM – DMO – ORC in order to find the best incident report and follow up process. 27/5/2020 ACE/BPM resources are not available to do the kaizen right now. Postponed due to Covid crisis.	31/03/2020 1Q 2021	CDO/DPC	■ On hold
18	Data incidents KPIs follow up – to identify patterns in data incidents in order to prevent them 11/12/2020: Data incidents pattern are now started organized in root causes; on going analysis of the link between this root causes and data assessments recommendations.	31/7/2020 1Q 2021	CDO	■ On going
8	KPI : Identify a main DATA indicator at the enterprise scale 11/12/2020: Please see KPI's Data dashboard covering Protection, Quality and monetization.	23/5/2019 30/06/2020	CDO + DCIO + CISO	■ Closed
6	Prepare and diffuse an awareness campaign to the Top Management and launch a new campaign for Shadow IT/EUC identification 27/8/2020: Shadow IT inquiry with 85% answers. On going: work with areas in order to accept risk or identify remediation plans. By now identified 32 Extreme and major risks and 26 action plans.	30/6/2019 30/06/2020	CISO + CDO	■ On going

1 - Data Dashboard

Data KPI's

GDPR Consent (Black list => -8%).....40%
Use Cases.....38

Use Case Value.....TBC
Incidents quality/process covered.....TBC
% Data Sustainable Investment.....TBC
Stars Strategic Workforce Planning.....1/3

DATA General Information

- ▶ **Cybersecurity Varonis project and Data classification** – (CISO / DMO) implementation ongoing, and production is scheduled to January. CISO/ITRO/CDO are working in the data classification rules in order to publish it in an Internal Note and support Varonis implementation.
- ▶ **Data Directive** (DMO) was published, updating Quality frame work and including several formalizations at Data Protection as Registration Of Protection Activities (ROPA) and Data Protraction Agreements (DPA).
- ▶ **Data Dictionary** (DMO) and **Reporting Inventory** (new procedure - DMO) are published.
- ▶ **Shadow IT campaign** (CISO) 98% entity covered by now; 103 risks identified, all rated; 33 high risk with remediation plan.

DATA Culture



- ▶ **Data Protection Training Sessions** designed with DPO.
 - **Progress** = 100% sessions delivered (14), all employers covered; - Participant's Level Global of Satisfaction - 4.7 (de 5); - KPI Data Goal – 98% trained Employees Covered.
- ▶ **Training on the new 2.0 Data Dictionary** - To be included in Data Training Plan 2021
- ▶ **Data AI Passport trainings** for 2021, on-going with HR.
- ▶ **Data Skills:**
 - Global Data skills identification (Dir. Transformation/IT/RH), including data exploitation and technical data treatment roles – Done: Strong business skills in data business analysts; Weaker technical data skills

DATA Opportunities



- ▶ **Megreen Data use cases:** 9 Data 2019 use cases under review; 17 new Data uses cases identified and under viability analysis (Operations, Risk and Marketing) with correspondent Central areas.
- ▶ **Global entity Data Use Cases** follow up: under construction, prepared by DMO, including MeGreen and other data use cases not yet analysed in MeGreen scope. By now total of 38 Data use cases identified.
- ▶ **Data Roadmap under review** a co-construction with Data Community, that represents the main stakeholders Data of the entity, It will be present in the December 16Th Data Committee.

DATA Quality Coverage: End-to-end & Projects

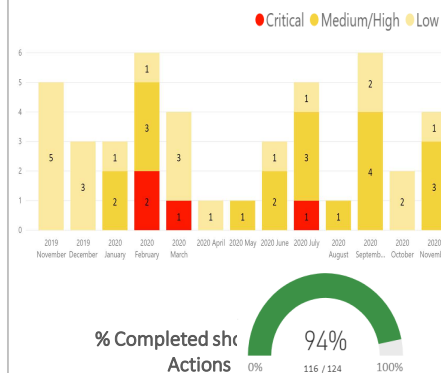
Data Assessments 2019-2020 (from Jan 2019 until Nov 2020)



- ▶ **Complete data quality coverage:** 125 NDoD variables in production since July with quality coverage in place; Data Assessment in all projects.
- ▶ **Partial data quality coverage:** 325 additional variables in Data Dictionary, controls identification and activation waiting by DW decommissioning.
- ▶ **Central variables Quality coverage:** (update and new variables in data dictionary): 85/114 ownerships confirmed. Resources mobilization continues to be a constrain when validating variables (62 validated). A narrower scope was defined to accelerate validation and decrease stakeholders workload.

1st Line of Defence DATA

Data Incidents Nov 2019-Nov 2020



- ▶ **Data Risk Cartography:** New risk identified "Lack of Data Acculturation" and included in data risk cartography.
- ▶ **DPO assessment results:** Remediation plan on-going with 80% finished actions; several remaining actions expected to be their closed confirmation during December, regarding Data Directive publication.
- ▶ **ROPA Privacy Assessment** Identified the need for 8 DPIAs, with high risk processes. Process Owners are performing the assessment (with IT). 60% performed. Deadline: 31.Dec.2020
- ▶ **Banco de Portugal reports incoherence –CRC/EMF/Crédito consumidores,** Total 12 (inconsistencies): 50% closed. Open inconsistencies: 4 IT dependent (1 is almost finished); 3 on-going by the task force.

1.1 - Zoom Inconsistences on Bank of Portugal Reports

CRC/EMF/Consumer Credit

Context

Bank of Portugal (BdP) announces that several inconsistencies, with significant volume, are being detected between CRC, EMF and consumer credit reports.

Such inconsistencies are due to differences detected in variables that should be the same, according to BdP.

The origin of these inconsistencies lies both on different interpretations in variables definitions provided by BdP, and the lack of a golden source.

A transversal task force has been created to analyse and permanent mitigate all the detected inconsistencies (March), with the active participation of stakeholders from DFIN, DMO, DRISK e IT.

Zoom “Date of Celebration” inconsistency

- Critical point, due to large number of contracts involved.
- Significant part already solved (waiting for BdP feedback on this point).
- Next actions, correction technical or operational process, could be a solution, but waiting BdP guidelines about definition.
- On going technical analysis in order to estimate developments.

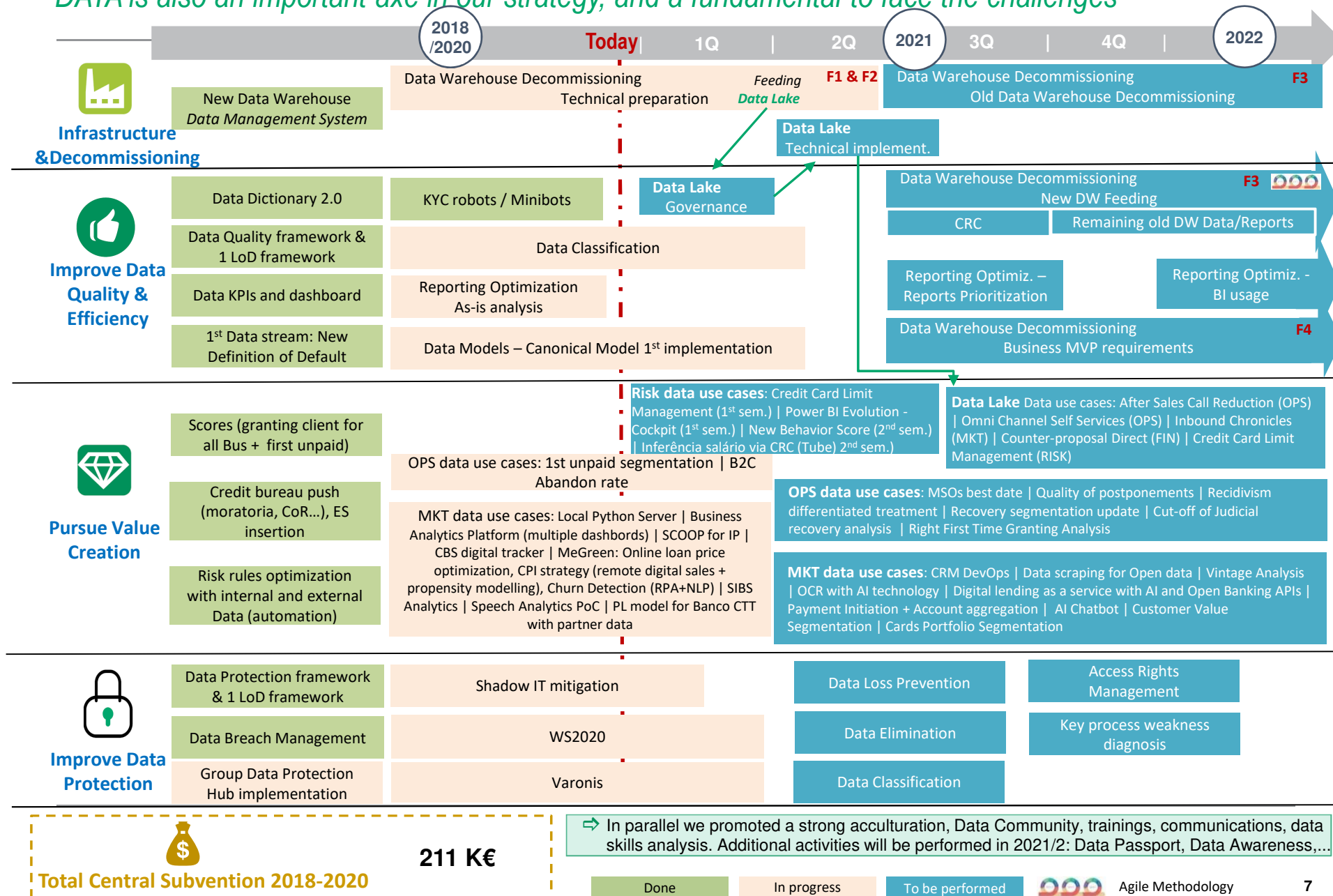
Global Current Situation

- 12 inconsistencies detected
- 50% closed:
 - 4 inconsistencies with IT dependency
 - 1 is being finalized.
 - 3 On going within the task force

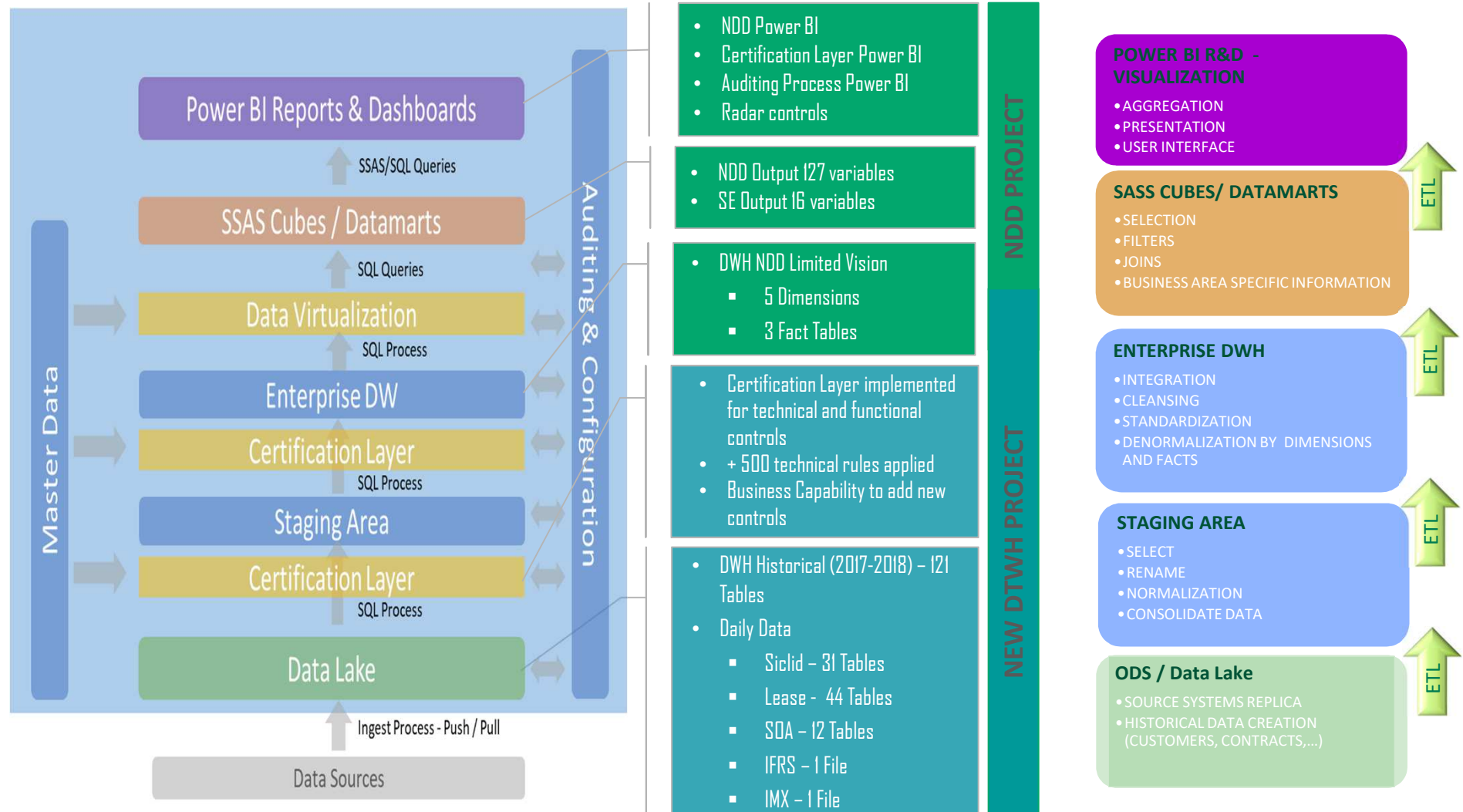
2 - Data Roadmap 2021

Decision to be taken

DATA is also an important axe in our strategy, and a fundamental to face the challenges



2.1 - New Data Warehouse - Implemented



- NDD and SE Daily Calculation/Outputs

- Different Certifications Layers between architecture levels

- Historical data based on source system tables

Highlights

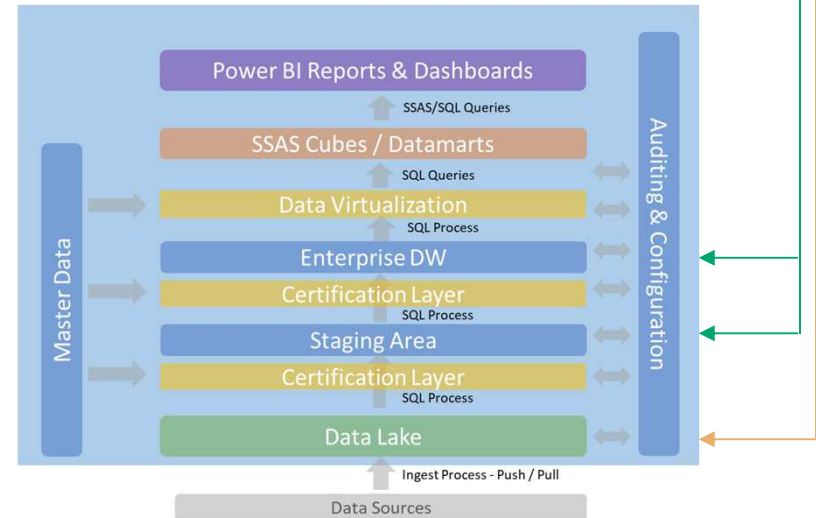
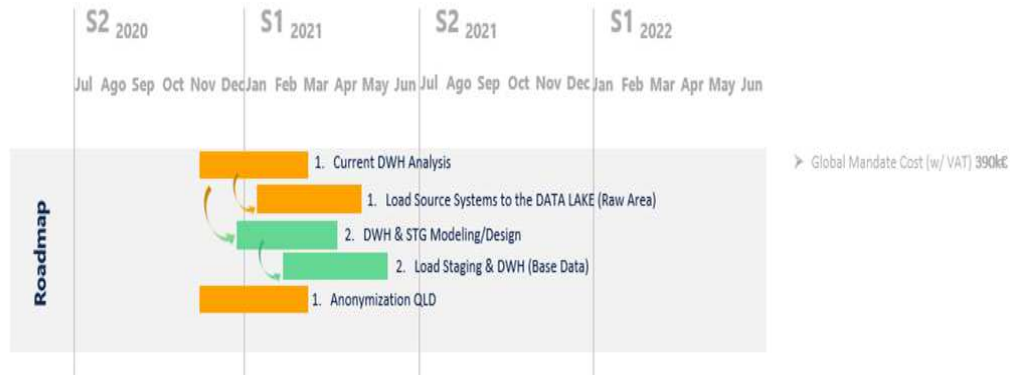
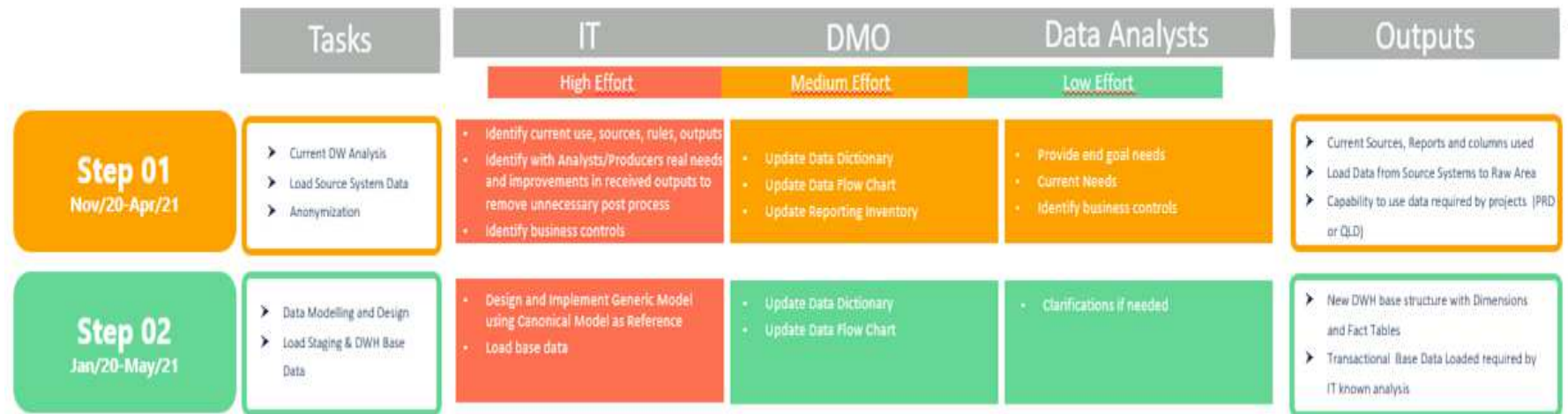
- Operations in each layer increase in complexity and business specific

- Data flows in multiple layers

- Each layer is dependent on previous layer

2.2 – Zoom Phase 1 & 2

Detailed roadmap



- Map Outputs with source systems columns
- Use PRD Anonymized Data in QLD for UAT

Highlights

- Design DWH and Staging Area
- Start feeding DWH and Staging

2.2 - Zoom Phase 1&2 (Cont.)

DWH Constrains for parallel execution – Transition period

FASE 1 & 2 Considerations

- Focus on ODS and Staging Area, with data on Staging Area arriving till end of May
- DWH design will happen till March (Macro Facts and Dimensions)
- This phase should not be done based in specific project requirements but on a global vision architecture to be used by all DTWH clients

CONSTRAINS OF THIS PERIOD

- Any output require specific analysis in DWH/ Datamarts and Visualization to implement during Phase 1 & 2 will create a “box” and won’t have integrated view.
- Any new requests for New DWH will require extra synchronizations and impacts with delays on Decommission project.
- Most issues on current DWH are related with limited vision requests and “box” requests with multiple flows and multiple information from same source, producing similar outputs, implementing projects before DWH vision will start with same error on new DWH.

WHAT’S POSSIBLE...

- **Decide in the meanwhile each demand/project where should be implemented to avoid duplicated work and big investments in an application that will be decommissioned...knowing that the structure and organization of the new DTWH cannot be compromised.**
- **Analyse what will be the security and governance rules that need to be implemented to allow opening the data lake information to data scientists work on business side, in the 1st semester of 2021, (evaluation to be done in January by DMO coordination and with cost estimation HW/SFTW).**
- **On the Data Lake in the end of phase 1 & 2 we will have data from the selected sources (applications) related with ex: Client, Credits, Transactions, Contacts..**

Doing parallel this in the new DTWH until May :

- Will create “boxes” on new DWH
- Impact time and effort on DWH Decommission Phase 1&2

Highlights

2.3 – Zoom Phase 3&4 - Decommission

	Tasks	IT	DMO	Data Analysts	Outputs
		High Effort	Medium Effort	Low Effort	
Step 03 Jun/21-Feb/22	<ul style="list-style-type: none"> Decommission Legal Reports Decommission <u>Visit, Fraud</u> Decommission Black Boxes Decommission ... 	<ul style="list-style-type: none"> Technical Analysis Implementation Advice best usage, outputs and processes to maximize DWH platform usage 	<ul style="list-style-type: none"> Update Data Dictionary Update Data Flow Chart Update Reporting Inventory Organization Alignment 	<ul style="list-style-type: none"> Clarifications if needed Change data consumption/usage processes Monitor Business Controls 	<ul style="list-style-type: none"> Load Information into DWH or <u>Datamarts</u> Build Reports/dashboards/ new outputs required by already existing projects
Step 04 May/21-	<ul style="list-style-type: none"> <u>Business MVP requirements</u> Requirements Analysis IT Developments, Tests, 	<ul style="list-style-type: none"> Technical Analysis Implementation Advice best usage, outputs and processes to maximize DWH platform usage 	<ul style="list-style-type: none"> Update Data Dictionary Update Data Flow Chart Update Reporting Inventory Organization Alignment 	<ul style="list-style-type: none"> Business Requirements Define End usage Identify Business Controls 	<ul style="list-style-type: none"> Load Information into DWH or <u>Datamarts</u> Build Reports/dashboards based on Requirements

Decommission Assessment (High Level)

- Reports : 148
- SQL Server
 - o Users : 12
 - o Tables : 17
 - o Columns: 362
- CUBE
 - o Users : 95
 - o Dimension Attributes : 602
 - o Metrics : 652

Draft Decommission DWH effort : 1080 m/D

Highlights

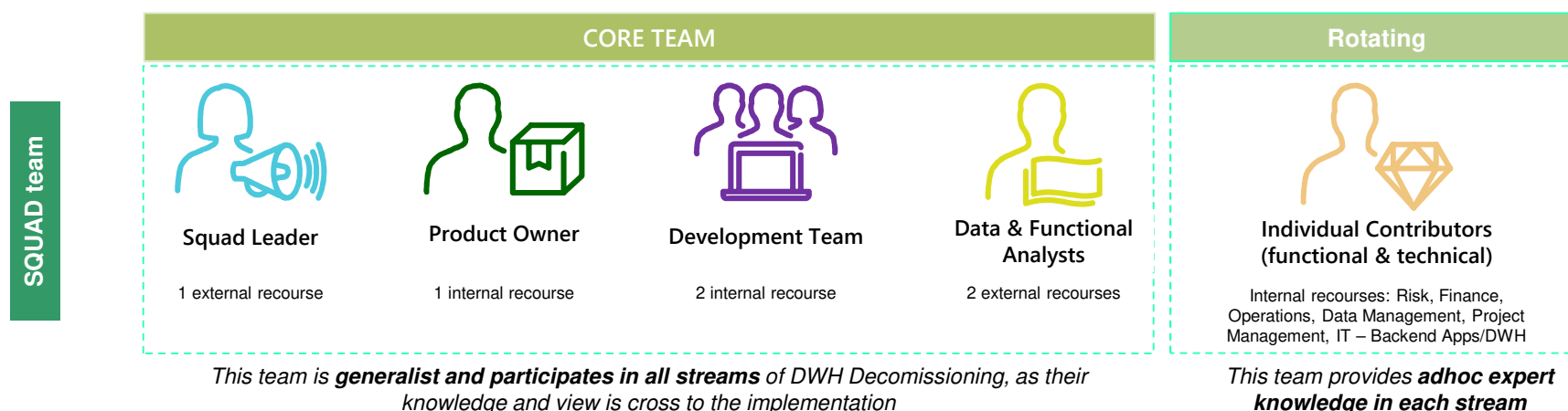
- Roadmap to be defined, leaded by DMO to technical start in May/21
- Business Definition and Clarifications to rules and controls
- Optimize produced output based on Business Needs
- Business availability to help providing global business view and current actions/outputs and needs during Modelling stage

2.3 – Zoom Phase 3&4 - Collaboration Approach

Decision to be taken

Organization: team and profiles

We propose the creation of a dedicated squad: a cross-functional, autonomous and high performing execution teams with have end-to-end responsibility for the delivery process, with an unique mission that guides the work to do, with a static and rotating component. To be implement in May/2021.



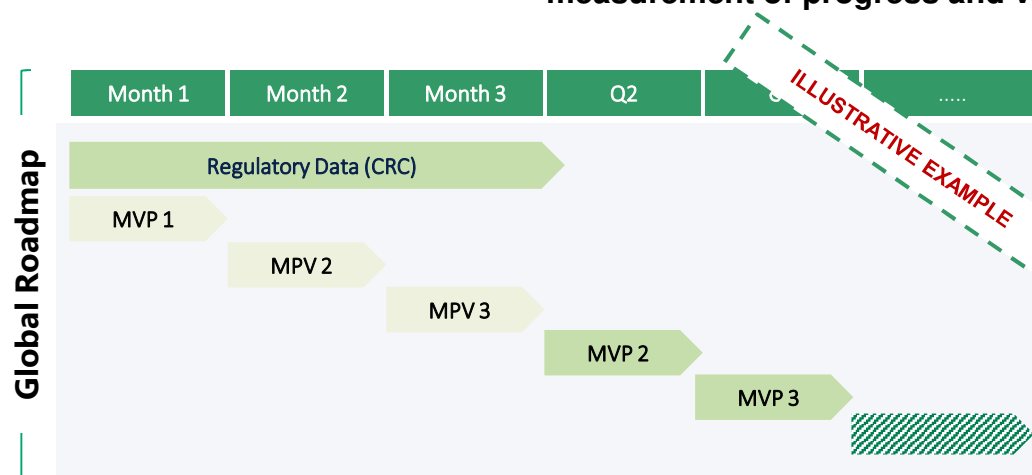
- ✓ This model allows **core team to exponentially increase their knowledge** during the different projects and be capable of **making faster deliveries**, reducing the learning cost and being able of efficiently transmit to other contributors the goals, methodology applied and DMS processes.
- ✓ On the other hand, **rotating teams come aboard according to the project as individual contributors**. For example, a Finrep project probably would have a representative of Accounting as PO and Risk Team as Individual contributors, as well as all experts in the systems from which the information comes from (current DWH, SICLID, Litigation team...).
- ✓ The team will use **agile elements** as a way of working dynamically and delivering sprints results and MVPs.

Please see More detail in appendix 4

2.3 - Zoom Phase 3&4 - Collaboration Approach (cont.)

Planning – Global Roadmap composition

The delivery will happen in sprints (1 month duration) and MVPs – minimum viable projects, allowing control and measurement of progress and value creation



- Using the Data Prioritization Log, MVPs will be defined and incorporated to the backlog according to the added value.
- Each MVP will be decomposed into **Sprints**. Each Sprint has a goal of what is to be built, a design and flexible plan that will guide building it, the work, and the resultant product increment → i.e. 30 variables created within a time box (8,5 days).

CRC proposal:

When creating MVPs from known data for example, CRC or other sets with broad information available, deliveries can be accelerated up to ~ 8,5 days for a 30 variables “pack”

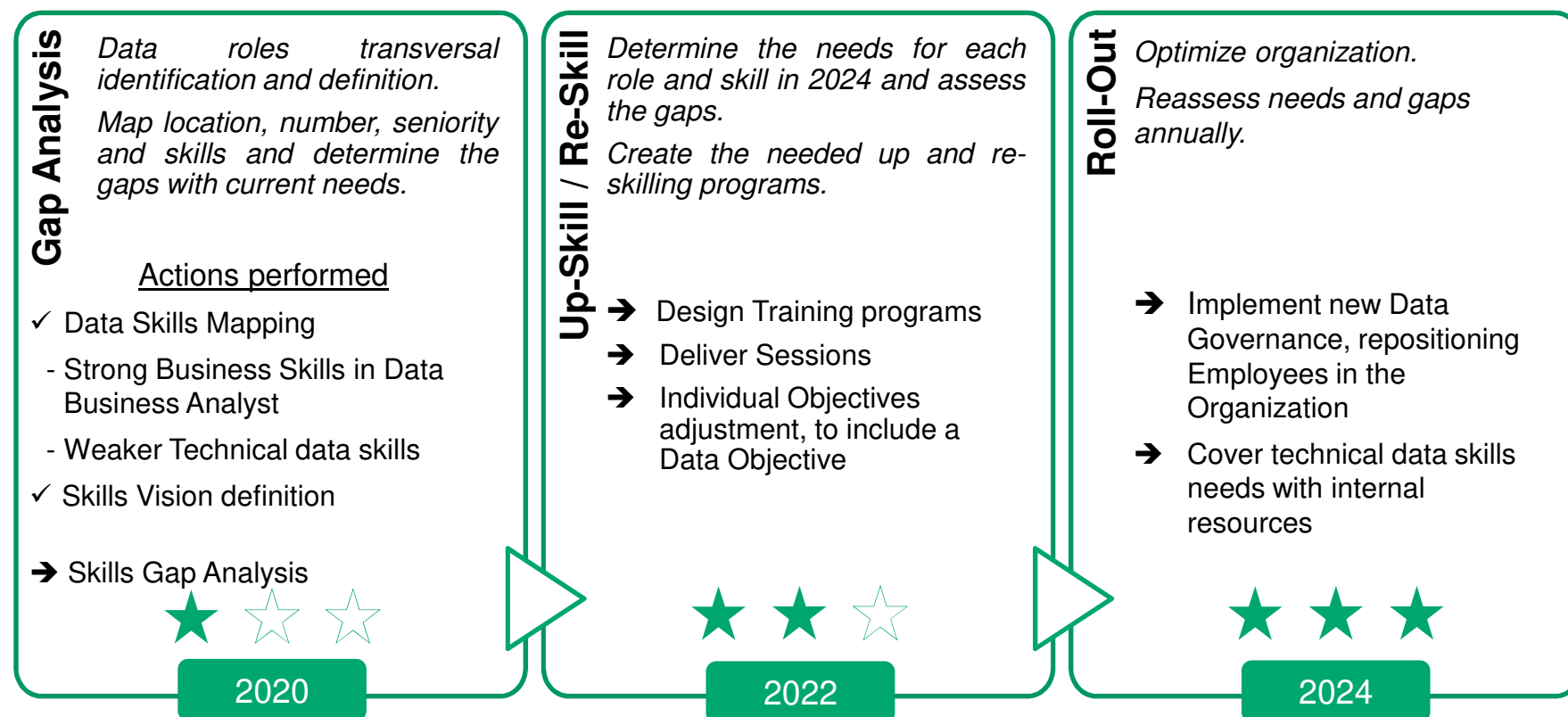
~3 months for decommissioning all 250 CRC variables

Phase	Target	Most likely	Worst-case sc.
Requirements (F)	1,0	2,0	4,0
Design (F)	2,0	4,0	6,0
Development (T)	3,0	6,0	8,0
Testing (F)	1,0	2,0	7,0
Deployment (T)	1,0	2,0	3,5
Review (F)	0,5	1,0	1,0
	8,5 days	17,0 days	29,5 days

(F) Functional Roles – PO, Squad Leader, functional and data analyst
(T) Technical Roles – Development team

3 - Data Skills/Profile

Data Strategic Workforce Planning – DMO / IT / RH



KPI PF Forward | # Stars Data Strategic Workforce Planning

Proposal Individual Data Objectives for 2021

- ▶ Employees w/ Data Roles: RH/DMO work with Managers to define a data objective aligned to overall strategy
- ▶ Other Employees: Participation in Data Quality Training



BNP PARIBAS
PERSONAL FINANCE



✓ Done
→ To do

3.1 - Data Skills/Profile

Zoom IT Data Resources

Current situation

- In a total of 66 internal resources we have ~**10 specialist** in data »» **15%**.
- An extra temporary team of **+6 external resources** it's being structured to work in the project of decommissioning of the old DTWH (only one missing).
- New DTWH manager process launched.
- Diversification of suppliers ongoing to reduce dependency in external support (BI4ALL, CAPGEMINI, SINGULARITY DE, CLOSER, HABBERTEC,...)
- Existing functional data knowledge in technical teams.

Market constraints

- Active market in technical data competences, with lack of resources available and increase of costs.
- RFP done to the market of outsourcers without good results.

Target

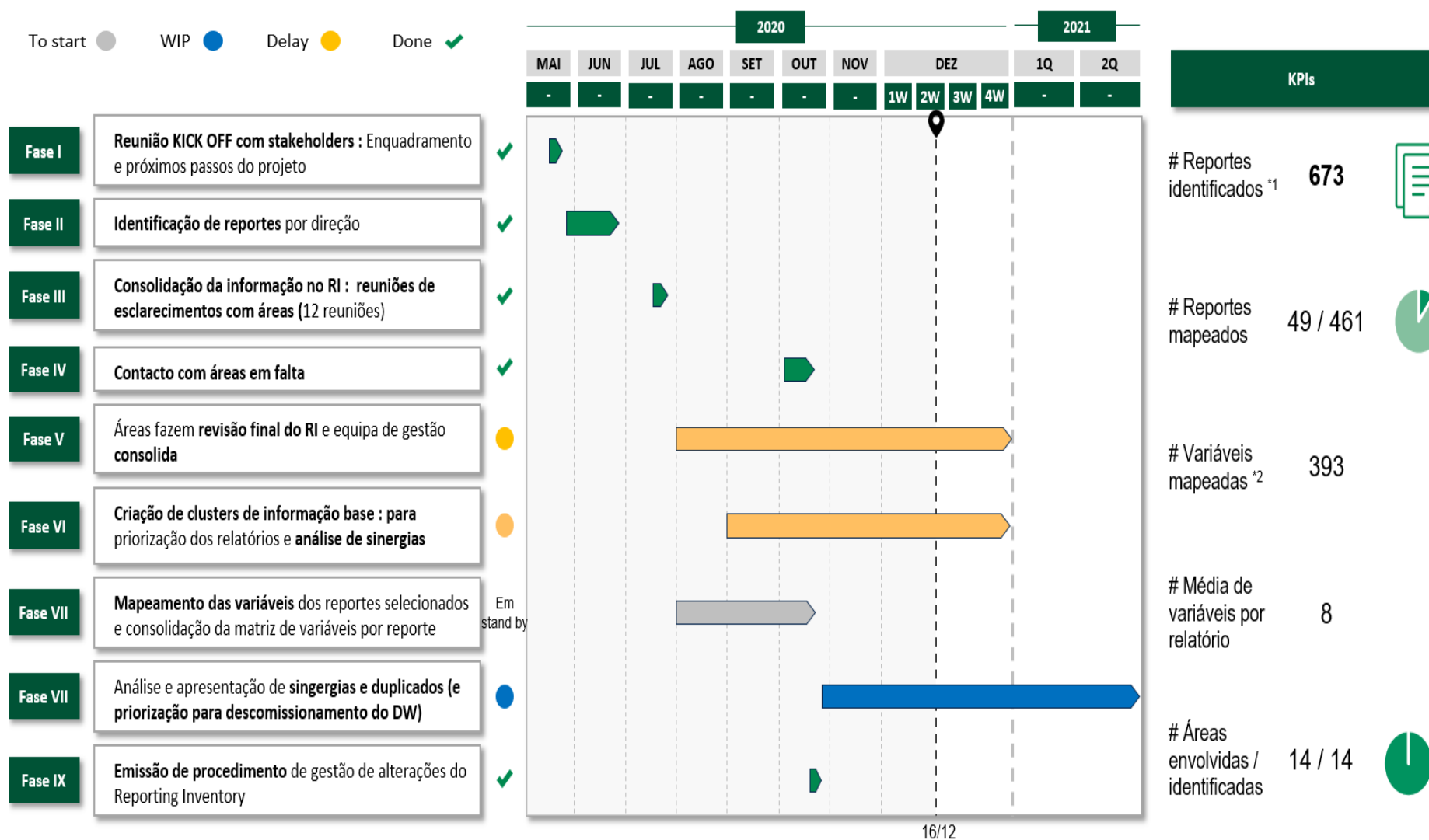
- Define future data organization (product vision or task force)
- Reinforce internal competences (trainee)
- Increase data resources (if validated)

Appendix 1: Group Assessment (via DPO)

Sub category	Declaration	Declaration description	Priority	Owner	Deadline	Closed	Status
Data Protection Policy	2.13	The Group Privacy and Personal Data Protection Policy or equivalent local Policy adapted to the Group	High	Cecilia	31/08/2020	18/05/2020	Closed
Local DP Network	4.15	There is sufficient execution power in the 1LoD	Medium	Cecilia	28/02/2021	30/11/2020	Closed
Training & Awareness	5.12	Completion of results of the mandatory e-learning are captured and reported to the DPO	Medium	Cecilia	28/02/2021	07/04/2020	Closed
Role specific DP training	6.11	DP Training is delivered to all new staff as part of the induction process (beyond e-learning)	High	Cecilia	31/08/2020	29/09/2020	Closed
Regular awareness	7.11	DPC / SPOC raises awareness on DP matters in the entity	High	Cecilia	31/08/2020	13/07/2020	Closed
DPO Certification	8.11	The DPC / SPOC has CIPP certification or equivalent	Medium	RH/CR	28/02/2021		Ongoing
DPO Certification	8.12	The DPC / SPOC has a record of individual GDPR training captured on My Development (or equivalent)	Medium	RH/CR	28/02/2021	24/07/2020	Closed
DPO Certification	8.13	The DPC / SPOC regularly attends industry events or DPA forums for raised awareness	Low	Cecilia			Ongoing
DPO Certification	8.14	The DPC has attended at least one Group DP Seminar in the last 12 months	Low	Cecilia		27/05/2020	Closed
RoPA Governance	9.11	Entity level local RoPA documented procedure for maintaining accurately the RoPA, with the DPO	Critical	Eduard	31/05/2020	12/11/2020	Closed
RoPA Governance	9.12	Documented procedures in place to capture new processing activities in the RoPA	High	Eduard	31/08/2020	12/11/2020	Closed
RoPA Accuracy and	11.11	There is a mechanism in place (and evidenced) to measure the completeness and accuracy of the RoPA	High	Eduardo	31/08/2020	25/09/2020	Closed
RoPA Accuracy and	11.12	There is a systematic tool in place for maintaining the RoPA	Medium	Eduard	28/02/2021	14/09/2020	Abandoned
RoPA Monitoring	12.11	Relevant DPC / SPOC have verified completeness of the elements of GDPR Article30	Critical	Cecilia	31/05/2020	25/03/2020	Closed
RoPA Monitoring	12.12	Relevant DPCs have performed sample testing of ROPA content	Critical	Cecilia	31/05/2020	25/03/2020	Closed
DPIA Training and monito	15.12	Training to 1LoD on how to use DP Hub has taken place	High	Eduardo	31/08/2020	13/11/2020	Abandoned
DPIA for existing processin	16.11	It is clear which existing processes qualify as high risk personal data processing	Critical	Eduardo	31/05/2020	28/09/2020	Closed
DPIA for existing processin	16.12	DPIAs have been performed on all existing processes that qualify as high risk personal data	Critical	Eduardo	31/05/2020		Ongoing
DPIA for existing processin	16.13	There is a plan that addresses the performance of DPIAs on existing processes that qualify	Critical	Eduardo	31/05/2020	25/09/2020	Closed

Appendix 2: Zoom Reporting Optimization

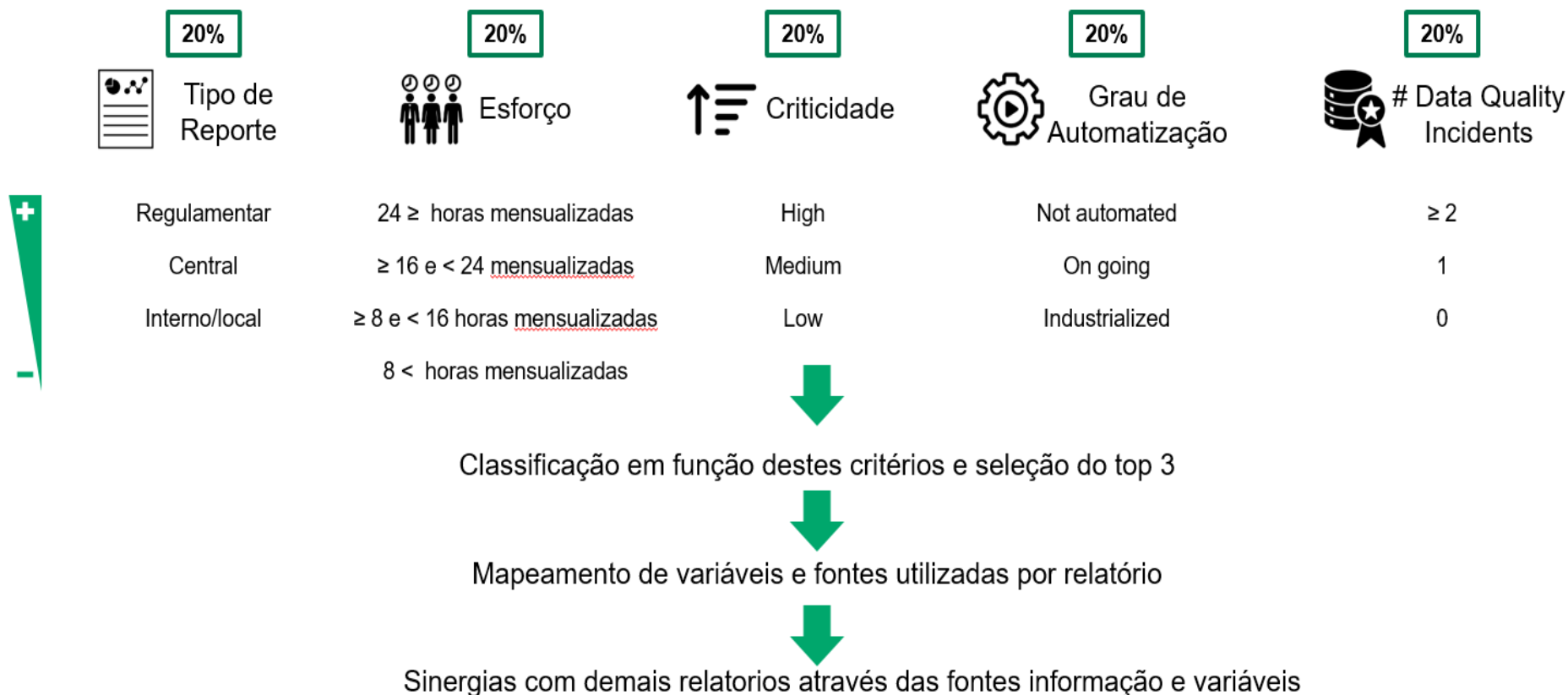
Roadmap Entregáveis



Appendix 2: Zoom Reporting Optimization (Cont.)

Classificação de Reportes, Sinergias e Descomissionamento DW

A estratégia de descomissionamento do DW deverá ser alavancar com as informações recolhidas no âmbito da atualização do Reporting Inventory



Appendix 3: Zoom Data Skills/Profile

Assessment realized in Oct/Nov 2020 by HR*

Data Analytics	COO	D REC. & CONT	D OPER & EFIC OPER	D AFS	D B2C&CP	D MKT	D RISK	DFIN	IT DADM	DMO	RISK ORC
DA01 Data Business Analyst	5	2		1	1	1					
DA02 Data Scientist	2					3	3				
DA03 Data Analyst			1			5	4	5			
DA04 Data Engineer									4		
DA05 Analytics Architect									2		
Data Management											
DAM01 Chief Data Officer										1	
DAM02 Referential Officer										1	
DAM03 Data Quality Analyst							2	2			
DAM04 Data Protection Officer											1
DAM05 Data Custodian (IT) Officer									1		

* High Level assessment based only on DF without challenged by Head of departments
N.B: Some employees may have several responsibilities in parallel.