



Data Collaboration Canvas

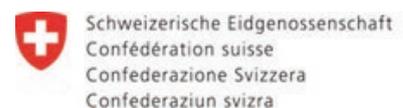
Facilitating Data Innovation Between Organizations SML Working Paper

Nico Ebert, Alice Dal Fuoco, Tim Geppert

In cooperation with:



Funded by:



Kommission für Technologie und Innovation KTI

ZHAW School of Management and Law
St.-Georgen-Platz 2
P.O. Box
8401 Winterthur
Switzerland

Institute of Business Information Technology
www.zhaw.ch/en/sml/institutes-centres/iwi/

Author/Contact
Nico Ebert, Tim Geppert
Nico.ebert@zhaw.ch, Tim.geppert@zhaw.ch

12.2021

DOI: 10.21256/zhaw-2421

Copyright © 2021, ZHAW School of Management and Law

All rights reserved. Nothing from this publication may be reproduced, stored in computerized systems, or published in any form or in any manner, including electronic, mechanical, reprographic, or photographic, without prior written permission from the publisher.

Table of Contents

Table of Contents	3
Purpose	4
Why should an organization use the DCC?	5
An overview of the DCC	6
Identifying data collaboration opportunities	8
1.1. Data Strengths	9
1.2. Data Weaknesses	9
1.3. Data Opportunities	9
Description of a data collaboration opportunity	10
1.4. Mutual Benefit	11
1.5. Required Data	11
1.6. Stakeholders	11
1.7. Constraints	11
1.8. Sharing Risks	11
1.9. Enablers	11
Example	12
1.10. Identification of data collaboration opportunities	12
1.11. Description of a data collaboration opportunity	13
About the DCC	15
Figures	16
Authors	17
Contributors	18

Purpose

The Data Collaboration Canvas (DCC) is a visual framework that allows users to generate ideas for sharing data across (or within) organizations. It can be used by organizations that want to explore the potential of data innovation with other organizations at an early stage of the collaboration to create mutual added value. This simple, visual structuring aid can, among other things, be employed in workshops to identify common potential and hurdles of collaboration.

The DCC helps to identify opportunities for data collaboration between companies or within an organization (e.g., between different divisions or departments).

Why should an organization use the DCC?

The DCC supports organizations by enabling them to explore and structure the potential for data collaboration. Its simplicity allows bad ideas to be discarded quickly and effortlessly. Due to its systematic nature, the potential for data collaboration can be identified in a more targeted way than with a completely individual approach. In addition, essential stakeholders for the further clarification of a cooperation idea can be identified at an early stage.

An overview of the DCC

The purpose of the DCC is to identify potential data collaboration opportunities between two or more organizations (or units with an organization). A data collaboration opportunity means that one or several data users receive data from one or several data owners providing that data for mutual benefit (Figure 1).

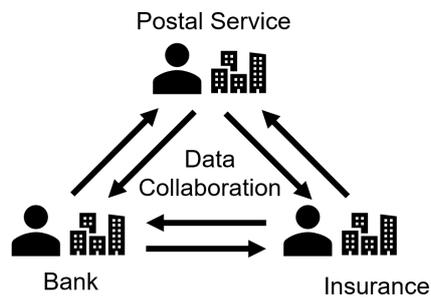


Figure 1: Data collaboration between two or more organizations

The data collaboration canvas follows a two-step process (Figure 2):

1. In a first step (optional), data strengths and weaknesses of the potential partners are documented. As a result, the partners can identify opportunities for data collaboration. These opportunities typically result from dataset combinations (e.g., additional features for the same data points, other data points with the same features). This step becomes obsolete if a data collaboration opportunity to be investigated already exists.
2. In a second step, individual opportunities can be further structured within the DCC. The structuring step is ideally used by heterogeneous teams in which team members have different backgrounds (business, data, technology, legal, or security). The DCC can be used in a workshop setting.

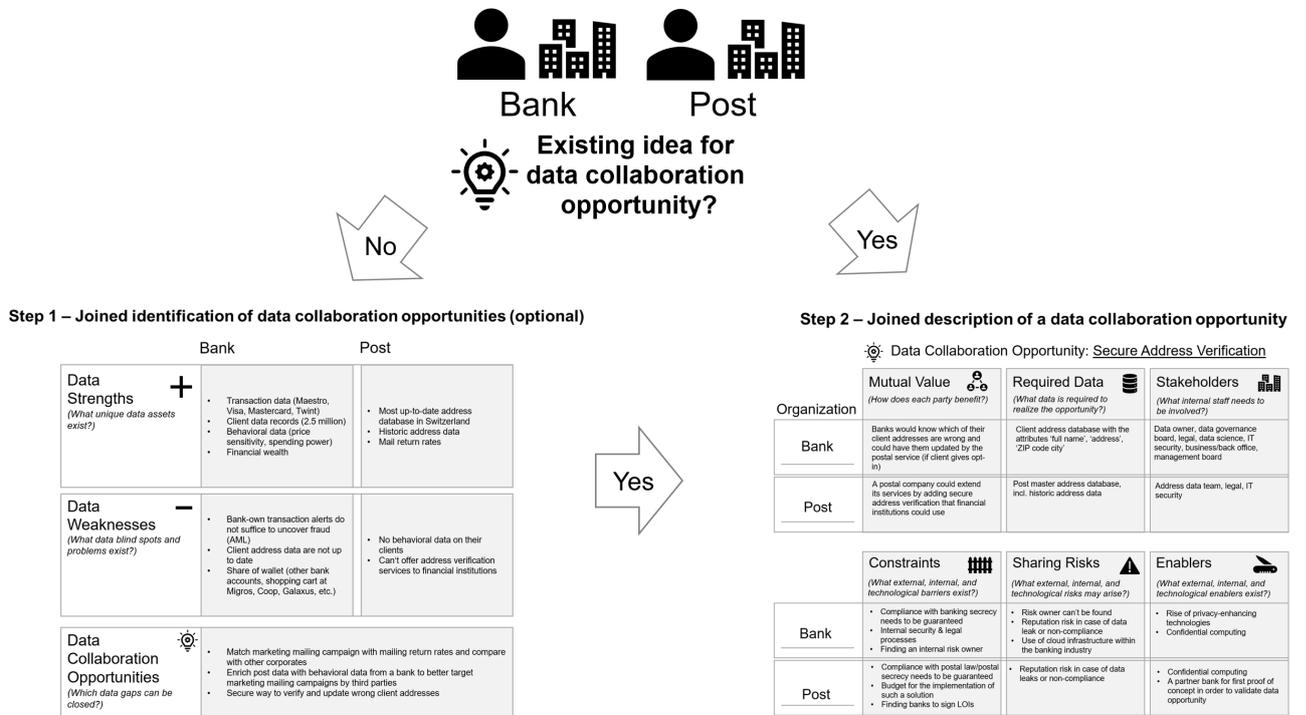


Figure 2: Two-step process to identify potential data collaboration opportunities

Identifying data collaboration opportunities

Organization: _____ Organization: _____

<p>Data Strengths </p> <p><i>(What unique data assets exist?)</i></p>		
<p>Data Weaknesses </p> <p><i>(What data blind spots and problems exist?)</i></p>		
<p>Data Collaboration Opportunities </p> <p><i>(Which data gaps can be closed?)</i></p>		

Figure 3: Part one of the DCC

This part of the DCC (Figure 3) helps users identify opportunities for collaboration between organizations (e.g., Party A and Party B). Imagine a workshop with different companies from different industries. To start, each company representative considers their organization’s data strengths and weaknesses and writes them down. Next, the representatives discuss these findings. Based on this exchange, ideas for data collaboration may evolve and can be written down as opportunities. This discussion enables an outside view of the data provided by each participant and facilitates innovation. Finally, the data collaboration opportunities are prioritized or rejected.

1.1. DATA STRENGTHS

Each party (e.g., a bank) describes the unique data assets that exist within its organization. These can be specific insights into customers that competitors do not have. The goal for users is not to describe all the data assets they possess. They should focus on those they consider unique for their organization and those which might be especially important for the other organization(s).

1.2. DATA WEAKNESSES

Each party describes their organization's data blind spots and other problems they may have. They consider what additional data would give them a competitive advantage or constitute a gain for them, and what new information about their customers they would like to have.

1.3. DATA OPPORTUNITIES

The parties now identify opportunities based on comparing the information provided by the parties involved (e.g., overlaps). They discuss the problems that could be solved using data provided by the other party. They write down who gives what data to whom to enable the collaboration to proceed. A data opportunity does not necessarily have to be a mutual exchange of data. One party can also pay another for its data.

Description of a data collaboration opportunity

This part of the DCC (Figure 4) facilitates the description of specific collaboration opportunities. The DCC enables the evaluation from the perspective of two or more participating organizations. In a first step, different data opportunities have been identified. Now, one data opportunity is analyzed in more detail. Together, the representatives of the companies involved in the data opportunity fill out a form (see figure below). This information will enable them to discover whether the opportunity that has been identified is feasible. If so, it can be discussed further with experts (e.g., legal, risk, IT); if not, it can be rejected (e.g., because there is no mutual benefit).

 Data Collaboration Opportunity: _____

Organization	Mutual Value  <i>(How does each party benefit?)</i>	Required Data  <i>(What data is required to realize the opportunity?)</i>	Stakeholders  <i>(What internal staff members need to be involved?)</i>

	Constraints  <i>(What external, internal, and technological barriers exist?)</i>	Sharing Risks  <i>(What external, internal, and technological risks may arise?)</i>	Enablers  <i>(What external, internal, and technological enablers exist?)</i>

Figure 4: Second part of the DCC

1.4. MUTUAL BENEFIT

What is the motivation, and what are the benefits of a data collaboration for each party? For a data collaboration to be feasible, all parties must benefit. Possible benefits include monetary compensation, receiving other data in return, the ability to compare one's data with that of other companies, and an improved process compared to existing solutions. Benefits could also be related to offering customers new or better services and a better image/reputation.

1.5. REQUIRED DATA

The organizations need to describe the specific data to be used in the context of an opportunity. They may have an initial idea of what data they need (e.g., customers' location data), which they need to specify here. The data may be specific (e.g., the records of a customer) or general (e.g., statistical data). The type of data will have an influence on subsequent issues to be addressed concerning risks and constraints. In the case of personal data (e.g., customer records), additional consent by the customer may be required.

1.6. STAKEHOLDERS

The stakeholder(s) involved in this opportunity need to be described in detail. These can be an internal department, such as marketing or operations, and support functions, such as IT, legal, and risk management. Stakeholders can also be external vendors.

1.7. CONSTRAINTS

What currently prevents the data owners from sharing data? Constraints are typically technological (e.g., there is no software interface), internal (e.g., there are no processes), or external (e.g., competition, legal regulations).

1.8. SHARING RISKS

What are the risks for the parties in realizing the data opportunity? Examples of risks are breach of confidentiality, negative effects on reputation, illegal cooperation (e.g., antitrust law), poor data quality, and data breaches.

1.9. ENABLERS

What are potential enablers related to the data collaboration opportunity? Enablers are typically issues that are technical (e.g., secure data transfer, data anonymization), organizational (e.g., non-disclosure agreements, certificates, insurances), or based on the environment (e.g., third parties, legal regulations).

Example

Below is an example of two parties using the data collaboration canvas: a retail bank and a postal services provider.

1.10. IDENTIFICATION OF DATA COLLABORATION OPPORTUNITIES

Figure 5 shows the outcome of the first step to identify data collaboration opportunities. The retail bank has identified different data strengths and data weaknesses. It has a large customer database and individual transaction data for each customer (e.g., credit card usage). It can generate behavioral data (e.g., price sensitivity) for its customers, and it can distinguish between different levels of financial wealth within its customer base. However, the bank also has some weaknesses. The transaction data the bank possesses is not sufficient to uncover the fraudulent actions by customers. Also, the bank often does not know when customers change their address after a move. Finally, the retail bank would be interested in its customers' "share of wallet", such as what other bank accounts they have and what expenses they incur that are not shown in the bank's transactions?

	Bank	Post
Data Strengths  <i>(What unique data assets exist?)</i>	<ul style="list-style-type: none"> Transaction data (Maestro, Visa, Mastercard, Twint) Client data records (2.5 million) Behavioral data (price sensitivity, spending power) Financial wealth 	<ul style="list-style-type: none"> Most up-to-date address database in Switzerland Historic address data Mail return rates
Data Weaknesses  <i>(What data blind spots and problems exist?)</i>	<ul style="list-style-type: none"> Bank-own transaction alerts do not suffice to uncover fraud (AML) Client address data are not up to date Share of wallet (other bank accounts, shopping cart at Migros, Coop, Galaxus, etc.) 	<ul style="list-style-type: none"> No behavioral data on their clients Can't offer address verification services to financial institutions
Data Collaboration Opportunities  <i>(Which data gaps can be closed?)</i>	<ul style="list-style-type: none"> Match marketing mailing campaign with mailing return rates and compare with other corporates Enrich post data with behavioral data from a bank to better target marketing mailing campaigns by third parties Secure way to verify and update wrong client addresses 	

Figure 5: Illustrative result of part one of the DCC

The other party, the postal service provider, has an up-to-date record of address data of its customers and a history of their address changes. They also know which addresses are outdated. However, the postal service provider does not have any behavioral data regarding its clients. It can offer address verification services to companies, but due to specific legal reasons, it cannot offer these services to banks because that would require the banks to reveal their customer data. This reduces the value of the postal service provider's data assets.

The retail bank and the postal service provider discuss different data collaboration opportunities. The bank could use the mail return rates of the postal service to create more efficient mailing campaigns. The post could enrich its customer data with behavioral data to target third parties' mailing campaigns. Both could also work on a lawful and secure way to verify the bank's customer addresses using the postal service's know-how.

1.11. DESCRIPTION OF A DATA COLLABORATION OPPORTUNITY

The two parties decide to develop the idea of "secure address verification," (Figure 6) which would allow the bank to have up-to-date customer addresses without revealing customer data to the postal service provider. They identify a mutual benefit: The bank would save money because mailing campaigns would reach their target audience, and the postal service could offer secure address verification as a new service also to other banks. Both companies identify their customer address data as "required data" They specify the internal stakeholders needed (e.g., the data owners and the legal department).

In a second step, the two companies discuss their current constraints. The retail bank has high standards with regard to bank secrecy and complicated internal security and legal processes. With regards to risk management, an internal stakeholder (e.g., a business line manager) willing to take the risk of using this service does currently not exist. The postal service is regulated under postal law and would need more than one bank to build a business case. Both companies fear reputational risk (e.g., in case of data leakage). However, they have also identified technological enablers that could provide a secure way of exchanging data in compliance with the law.

 Data Collaboration Opportunity: Secure Address Verification

Organization	Mutual Value  <i>(How does each party benefit?)</i>	Required Data  <i>(What data is required to realize the opportunity?)</i>	Stakeholders  <i>(What internal staff needs to be involved?)</i>
Bank	Banks would know which of their client addresses are wrong and could have them updated by the postal service (if client gives opt-in)	Client address database with the attributes 'full name', 'address', 'ZIP code city'	Data owner, data governance board, legal, data science, IT security, business/back office, management board
Post	A postal company could extend its services by adding secure address verification that financial institutions could use	Post master address database, incl. historic address data	Address data team, legal, IT security

	Constraints  <i>(What external, internal, and technological barriers exist?)</i>	Sharing Risks  <i>(What external, internal, and technological risks may arise?)</i>	Enablers  <i>(What external, internal, and technological enablers exist?)</i>
Bank	<ul style="list-style-type: none"> Compliance with banking secrecy needs to be guaranteed Internal security & legal processes Finding an internal risk owner 	<ul style="list-style-type: none"> Risk owner can't be found Reputation risk in case of data leak or non-compliance Use of cloud infrastructure within the banking industry 	<ul style="list-style-type: none"> Rise of privacy-enhancing technologies Confidential computing
Post	<ul style="list-style-type: none"> Compliance with postal law/postal secrecy needs to be guaranteed Budget for the implementation of such a solution Finding banks to sign LOIs 	<ul style="list-style-type: none"> Reputation risk in case of data leaks or non-compliance 	<ul style="list-style-type: none"> Confidential computing A partner bank for first proof of concept in order to validate data opportunity

Figure 6: Illustrative result of part two of the DCC

About the DCC

The DCC was developed as part of the research project “Confidential Data Analytics Based on Trusted Execution Environments”, which was funded by Innosuisse, and the expert group “Privacy Technologies for Data Collaboration“ of the Swiss Data Innovation Alliance. The authors would like to thank the many people who contributed via discussions, interviews, document reviews, and participation in a workshop. They would especially like to acknowledge the contribution of the parties listed below.

Figures

Figure 1: Data collaboration between two or more organizations	6
Figure 2: Two-step process to identify potential data collaboration opportunities	7
Figure 3: Part one of the DCC	8
Figure 4: Second part of the DCC	10
Figure 5: Illustrative result of part one of the DCC	12
Figure 6: Illustrative result of part two of the DCC	14

Authors

Nico Ebert (ZHAW), nico.ebert@zhaw.ch

Alice Dal Fuoco (PostFinance)

Tim Geppert (ZHAW), tim.geppert@zhaw.ch

Contributors

Markus Barmettler (Comparis)

Ilija Braun (PostFinance)

Stefan Deml (Decentriq)

Sebastian Eckhardt (Swiss Re)

Georg Russ (Mobiliar)

Aetienne Sardon (Swisscom)

Marcel Schöngens (SIX Group)

David Sturzenegger (Decentriq)

Zurich University
of Applied Sciences

School of Management and Law

St.-Georgen-Platz 2
P.O. Box
8401 Winterthur
Switzerland

www.zhaw.ch/sml



swissuniversities



European Business Schools
Ranking 2021