PayPal Under the Microscope - Fact Sheet

PayPal's Business Model

During 2021, there were 19.3 billion payment transactions made through PayPal which sum up to 1.25 trillion dollars. The relationships in this payment network are strong predictors of behaviors and great risk indicators that can reveal fraud. However, a much richer set of data items becomes available for online purchases, including an itemized bill or information about the buyer. It is expensive to collect and store seemingly unrelated data of 426 million PayPal users. This means that it must be used in some way that brings PayPal profit.

Minimum data requirements for handling payments:

- order total
- the receiving merchant
- an authenticated payment

Data privacy concerns are also related to consumer behavior. Payment providers can collect and connect purchase details at a large scale to create a behavioral profile of the consumer. **This allows PayPal to build up comprehensive consumption profiles across sites.**

PayPal applies several **commercial strategies** to address private customers. Their commercials try to trigger internal needs like simplicity, velocity, diversity, and safety through "fast checkout", extensive buyer protection, instalment payment option etc.

Example:
„Willste? Kriegst. Das ist der PayPal Unterschied."

(engl.: „You want it? You get it. That's the PayPal difference.")

PayPal introduced the merchant **Application Programming Interface (API)** with a basic credit card payment feature. Through this, any small corporation can start their business with near-zero investment and accept payments over the internet via PayPal.
APIs are increasingly used in everyday items when using apps on our smartphones, but outsourcing data storage to a cloud operator leads to data privacy issues. The more APIs are used, the more vulnerable our privacy gets when a cloud is attacked.
The recruitment of smaller marketing and payment services/providers opens a gate for PayPal for acquiring much richer datasets and a significantly more well-rounded estimation of their users' profiles.

The **PayPal User Agreement** is a document that consists of far more words than those of other companies, and therefore takes sufficiently more time to read. Experts claimed that PayPal violates the principle of transparency and users are disadvantaged. Based on a scientific text analysis, PayPal's Terms & Conditions are formally incomprehensible and the time and effort that users must invest is unreasonable.
With more recent versions, PayPal made their Terms & Conditions shorter, but they still have the longest document compared to other companies. Additionally, the sentences in the terms and conditions and privacy statements are often passive constructions and complicated.

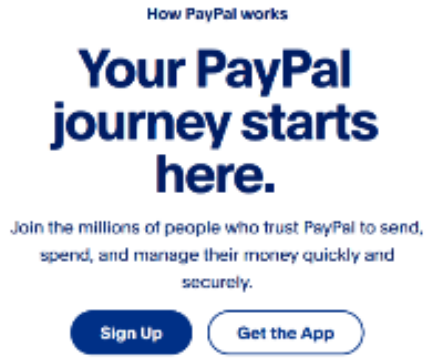
- Words and phrases as **'possibly'** and **'among other things'** make it even more vague and harder to understand.
- The average user will not spend **almost 1.5 hours** on reading a document that they have no influence on, and which they cannot refuse, if they want to keep the possibility of using the service.

Online service providers try to deceive users to give access to their personal data voluntarily, which is done by exploiting cognitive biases through **nudging and implementing dark patterns**. Examples are user logins at starting pages that show an obvious login with personal accounts like Facebook, Google and a less bright "create a separate account" button and a hidden skip button (often in the upper right corner). The intention seems to be accessing friends lists, photos and likes, and therefore getting information that is basically not necessary for the service by using intentionally bad design, user-experience tricks, and knowledge from Human-Computer Interaction research.

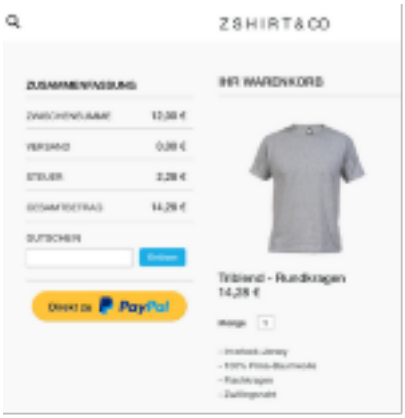
Dark patterns become effective by their frequency of occurrence which can be derived from the psychological investigations on these strategies. Their focus seems to lie on cognitive information processing and the psychological constitution. The fundamental need to belong highly influences the effectiveness of dark patterns. This highly increases peoples' readiness to provide personal data.

Nudging strategies use cognitive shortcuts observed in (irrational) decision making to make irrational behavior more explainable and even predictable. Social framing was found to be an exceptionally effective social behavioral nudge. Since it employs descriptive social norms as prescriptive rules it can be used effectively, and is already adopted by instances like governments, employers, and businesses. Also, different patterns are used depending on the language/country selected.

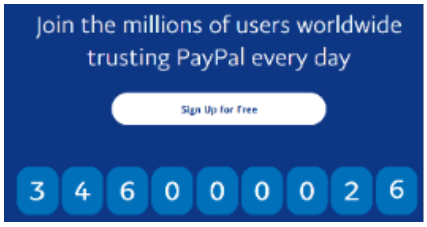
Examples:
"7 people are looking at this right now!", "This deal expires in 15 minutes!".
Those examples make users act with a kind of urgency by turning an abstract principle into a concrete nudge.



A dark pattern using different button colors to make the user sign up to PayPal.



Express checkout button using bright colors that stick out. It implies "PayPal security" and therefore increases the confidence in the shops.



PayPal addresses the users' fundamental need to belong. The counter doesn't really count the actual number of users. It starts at 346.000.000 users every time one opens the website.

PayPal also provide detailed information in which they help businesses to achieve the PayPal "advantages" by telling the customer that PayPal is available at their service. The guide explains in detail how to place the express checkout button precisely to keep the users' attention and to make the user buy items from their shop.

The goal is to „blend digital and physical commerce to create seamless customer experiences".

Big Data, Privacy and PayPal

The sum of an individual's personal data forms a **digital identity** which can be traded as a commodity or capital. It can be later utilized to feed predictive models. This leads to a "data dictatorship" in which "we are no longer judged on the basis of our actions, but on the basis of what all the data about us indicates our probable actions may be". Because our digital identity expresses one's identity, such data must be protected.

PayPal collects the following data and more for advertising and marketing, analysis, personalization of the product, app functionality and to enhance their users' experience, to which users agree to in the terms and conditions:

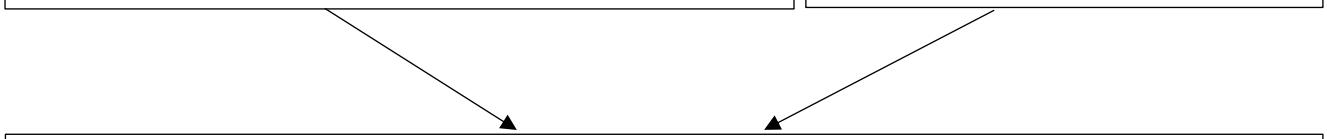
- Shopping history, contact data, browsing history, financial information, localization

This conflict with the General Data Protection Regulation (GDPR) of the European Union (EU), which states that the principle of **'data minimization'** should be applied. This means collecting as little information as possible for a purpose. The problem is that the EU law doesn't define what 'adequate, relevant and limited' data is. Therefore, PayPal may collect any data if they can justify it with their business interests, which are not obvious to the users.

Why is this a problem?
The more data gets collected, the greater the success for artificial intelligence (AI) and machine learning algorithms (ML) to find hidden patterns in that data to predict our behavior and desired outcomes.

Proxy discrimination
Making a prediction about a member of a protected class (e.g., race), a non-protected variable that highly correlates with the protected class, a so-called proxy,
can be used to make such a prediction. When algorithms are trained on large datasets to make predictions, they rely on proxies to make the most accurate prediction. This leads to using group characteristics that again amounts to stereotyping and discrimination, because individuals are treated the same as other people that have behaved similarly in the past.

Predictive analytics
Based on large amounts of data from many different people, behavior patterns and interests are predicted for individuals who decided to not share their data. Because many other people decided to share private information, a sensible private space is violated for people who do not share private information.



Example: Credit Scoring
AI is used to convert data points about an individual into a credit score while making use of **proxy discrimination** and **predictive analytics**. This might lead to unfair and discriminating decisions, because individuals might receive a worse credit score simply because they behave similarly like someone else who is indeed not trustworthy. Therefore, the vicious cycle of dirscrimination, unequal changes and the gap between poor and rich is exacerbated. In addition, the way an algorithm derives a prediction is not readable to humans, making it impossible to explain why an individual received a certain credit score.