

Přílohy C. Praktické příklady datových sandboxů

Španělsko

Banco de España se zúčastnila pilotního projektu zaměřeného na syntetická data, který vedla Evropská komise s cílem podpořit vytvoření datového centra v rámci budoucí platformy EU pro digitální finance. Pilotní projekt spočíval v testování postupu generování syntetických dat na základě datových souborů centrální rozvahy (CBI) a úvěrů právnickým osobám (CIR), které jsou k dispozici v datové laboratoři Banco de España (BELab). Počáteční výsledky získané během pilotního projektu byly celkově pozitivní a povzbudivé pro další zkoumání technologie. V datech však byla zjištěna určitá omezení, jako je potlačení odlehklých hodnot, nerovnováha mezi proměnnými a nemožnost sloučení syntetických datových souborů.

Pilotní projekt zdůraznil, že syntéza dat by měla být považována za průřezový projekt, který v rámci orgánu dohledu vyžaduje spolupráci více oddělení a závazek poskytovatele softwaru. Přestože laboratoř BELab zkrátila dobu potřebnou k syntéze dat, byl proces velmi náročný na čas a zdroje a vyžadoval vysoce specializované pracovníky.

Banco de España se domnívá, že syntetická data by mohla být cenným příspěvkem pro sandboxy a přeshraniční testování, avšak omezení zjištěná v datech mohou v některých případech bránit jejich použití. Účastníci trhu by si proto měli být těchto omezení vědomi. Výsledky pilotního projektu by při stanovování podmínek nabídkového řízení nebo budování datového centra neměly být považovány za jediný zdroj informací. K získání spolehlivějších závěrů je zapotřebí další analýza a více času.

Spojené království

Orgán dohledu nad finančními trhy (Financial Conduct Authority, FCA) společně s institucí City of London Corporation úspěšně vytvořil digitální sandbox, přičemž již probíhají dva pilotní projekty, které nabízejí přístup k syntetickým datovým souborům pro testování a vývoj ověření koncepce. První pilotní projekt probíhal od října 2020 do února 2021 a zaměřoval se na zlepšení financování malých a středních podniků, odhalování a prevenci podvodů a podporu finanční odolnosti zranitelných zákazníků. Druhý pilotní projekt probíhal od října 2021 do března 2022 a zaměřoval se na finanční inovace spojené s udržitelností. Kromě syntetických datových souborů nabídl pilotní projekty účastníkům přístup k řadě dalších vývojových nástrojů, jako jsou API a programovací prostředí, stejně jako přístup k odborným mentorům a pozorovatelům.

Účastníci označovali syntetická data za nejhodnotnější prvek, přičemž se zároveň jednalo o prvek s největším potenciálem pro zlepšení. Referenční datové soubory a podrobnější údaje by zejména umožnily účinnější testování a další vývoj produktů. Celkově 84 % respondentů uvedlo, že pilotní projekt urychlil vývoj jejich produktů. I když je obtížné tuto úroveň zrychlení zjistit nebo kvantifikovat, analýza ukazuje, že největším faktorem byl snadný přístup k datům při vývoji počáteční fáze ověření konceptu. Několik účastníků odhadovalo, že svůj vývoj urychlili o 4–6 měsíců, přičemž u jednoho z nich se jednalo o 18–24 měsíců, a to především díky tomu, že nejdříve nemuseli hledat partnera z odvětví a spolupracovat s ním, aby mohli začít s ověřováním konceptu, nebo že získávali či generovali data sami.

Ti, kteří zjistili, že pilotní nástroje jejich vývoj neurychlily nebo nezlepšily, obecně uváděli, že se to týkalo toho, že data byla ke splnění jejich případu použití nějakým způsobem nedostatečná. To bylo důsledkem jedné z těchto možností nebo jejich kombinace:

Nedostatečně podrobná data, zejména nedostatek relevantních typologií nebo chování, které byly potřeba k modelování řešení. Do syntetických dat například nebyly z důvodu složitosti vloženy vzorce transakčních výdajů, které svědčí o kolísavém duševním zdraví zákazníka (v případech použití u zranitelných spotřebitelů).

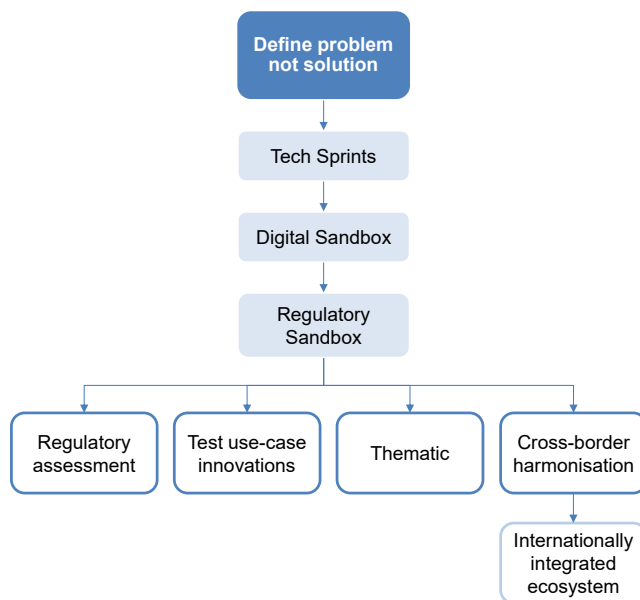
Požadované datové soubory nebyly k dispozici. Například velké objemy nestrukturovaných dat, jako jsou texty stížností spotřebitelů, k trénování a ověřování technik zpracování přirozeného jazyka.

Datové soubory nebyly referenčně propojeny. Například různé datové soubory byly generovány nezávisle, takže vzorce chování nebo charakteristiky syntetického jedince „Johna Smithe“ se neshodovaly s „Johnem Smithem“ v samostatném datovém souboru.

I přes tato omezení však účastníci zaznamenali užitečnost datových souborů pro „bootstrapping“ návrhu produktu. I v případech, kdy data nemohla být použita k upřesnění algoritmického modelu, měla význam při poskytnutí datových modelů, datových struktur a formátů, které byly reprezentativní pro to, s čím by pracovali v reálném produkčním prostředí (FCA, 2021^[1]).

Je třeba také poznamenat, že ve Spojeném království byl proces vývoje FinTechových řešení optimalizován díky dlouholetým zkušenostem a inovacím v tomto odvětví. Byl zaveden systematický přístup, který začíná definicí problému a nechává FinTechové společnosti prozkoumat možná řešení. Dalším krokem je Tech Sprint, který zahrnuje definování problému a poskytnutí datových souborů k vývoji ověření konceptu. Fáze digitálního sandboxu navazuje na ověření konceptu a dále testuje potenciální řešení a jejich životaschopnost. A konečně, regulační sandbox posuzuje regulační požadavky na minimální životaschopný produkt (MVP) vyvinutý během předchozích fází. Tento postupný proces poskytuje jasnou cestu pro vývoj a implementaci inovativních FinTechových řešení, což umožňuje hladký přechod od konceptu k produktu připravenému k uvedení na trh (Obrázek A C.1).

Obrázek A C.1. Digitální a regulační sandboxy založené na systematickém přístupu orgánu FCA Spojeného království



Existuje několik zemí, které implementovaly datové sandboxy, včetně Austrálie, Singapuru, Spojeného království a Kanady. Jediným sandboxem zaměřeným na finanční inovace, který účastníkům dosud nabízel přístup k finančním datům, je však orgán FCA, přičemž poslední dvě kohorty nabízejí datový atribut.

Austrálie

Australská komise pro hospodářskou soutěž a spotřebitele (ACCC) zřídila sandbox pro práva spotřebitelů v oblasti dat (CDR)¹ na podporu vývoje a testování řešení souvisejících s těmito právy ze strany podniků. Komise ACCC vytvořila „falešný“ registr, „falešného“ držitele dat a „falešné“ úložiště týkající se CDR pro příjemce dat, které podnikům pomáhají vyvíjet a testovat řešení v oblasti CDR v jejich vlastním IT prostředí. Toto falešné prostředí bylo z platformy GitHub staženo více než 20 000krát a od účastníků CDR a poskytovatelů platforem získalo pozitivní zpětnou vazbu. Komise ACCC nyní spouští hostované prostředí sandboxu, které navazuje na tyto zkušební nástroje, což podnikům umožňuje testovat řešení v oblasti CDR s ostatními účastníky a ověřovat technická řešení v rané fázi cyklu vývoje softwaru. Tento sandbox má potenciál snížit bariéru pro podniky, které se k CDR chtějí připojit, a zvýšit kvalitu řešení. Komise ACCC očekává, že sandbox bude účastníky CDR široce přijat a umožní testování společných platforem s jinými systémy před jejich rozsáhlou implementací.

Referenční dokumenty

FCA (2021), *Supporting innovation in financial services: the digital sandbox pilot*, [1]
<https://www.fca.org.uk/publication/corporate/digital-sandbox-joint-report.pdf>
 (accessed on 20 March 2023).

Poznámka

¹Austrálie uvádí typ dat dostupných pro sandbox prostřednictvím tohoto odkazu: <https://github.com/ConsumerDataRight/sandbox>.



From:

Supporting FinTech Innovation in the Czech Republic

Regulatory Sandbox Design Considerations

Access the complete publication at:

<https://doi.org/10.1787/081a005c-en>

Please cite this chapter as:

OECD (2023), “Praktické příklady datových sandboxů”, in *Supporting FinTech Innovation in the Czech Republic: Regulatory Sandbox Design Considerations*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/d1320b5f-cs>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.