



Zadání bakalářské práce

Název:	Sběr dat v herních komunitách
Student:	Luděk Rájecký
Vedoucí:	Ing. Jan Baier
Studijní program:	Informatika
Obor / specializace:	Webové a softwarové inženýrství, zaměření Softwarové inženýrství
Katedra:	Katedra softwarového inženýrství
Platnost zadání:	do konce letního semestru 2022/2023

Pokyny pro vypracování

1. Seznamte se se způsoby sběru dat v rámci herních komunit.
2. Navrhněte vhodnou kategorizaci komunit a nalezněte vhodné zástupce pro dané kategorie. Snažte se minimalizovat počet základních kategorií (maximálně jednotky), hlavním kritériem při rozdělování bude způsob sběru dat dle předchozího bodu.
3. Proveďte stručnou rešerši software využívaného v rámci herních komunit.
4. Navrhněte tzv. "best practices" pro jednotlivé základní kategorie, poukažte na případné problémy při sběru dat a doporučte konkrétní řešení.
5. Veškeré návrhy zpracujte v souladu s běžnými postupy a metodami SW inženýrství.
6. Výsledkem práce bude, kromě pečlivě provedené rešerše, sada návrhů na vhodné zpracování dat v jednotlivých typech herních komunit.

Bakalářská práce

SBĚR DAT V HERNÍCH KOMUNITÁCH

Luděk Rájecký

Fakulta informačních technologií ČVUT v Praze
Katedra softwarového inženýrství
Vedoucí: Ing. Baier Jan
13. května 2021

České vysoké učení technické v Praze
Fakulta informačních technologií

© 2021 Luděk Rájecký. Všechna práva vyhrazena.

Tato práce vznikla jako školní díla na Českém vysokém učení technické v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bez uplatněných zákonných licencí nad rámec oprávnění uvedených v Prohlášení je nezbytný souhlas autora.

Odkaz na tuto práci: Luděk Rájecký. *Sběr dat v herních komunitách*. Bakalářská práce. České vysoké učení technické v Praze, Fakulta informačních technologií, 2021.

Obsah

Poděkování	vi
Prohlášení	vii
Abstrakt	viii
Shrnutí	ix
Seznam zkratk	x
1 Úvod do problematiky	1
1.1 Hráči a hry offline vs. online	2
1.2 Dostupnost informací o hrách	2
1.3 Vyhledávání dat hráči	3
2 Rozdělení sběru dat v herních komunitách	7
2.1 Komunity bez sběru dat	7
2.2 Komunity řešící průtoková data	9
2.3 Komunity s datovým řešením postaveným na kolektivu	11
2.4 Komunity analyzující herní data ve spolupráci s vývojáři	12
2.5 Komunity analyzující herní data bez spolupráce vývojářů	12
3 Případové studie	13
3.1 Průzkum a metodologie	13
3.2 Komunity řešící průtoková data	13
3.3 Komunity s datovým řešením postaveným na kolektivu	15
3.4 Komunity analyzující herní data ve spolupráci s vývojáři	16
3.5 Komunity analyzující herní data bez spolupráce vývojářů	16
4 Specifika sběru dat	19
4.1 Rozdíl mezi západními a východními komunitami ve sběru dat a práci s nimi	19
4.2 Problematika odhadu informací a nahrazení herních vývojářů fanoušky	21
4.3 Výzkum v herních komunitách a možnost zneužití dat	23

5 Doporučení přístupu k problematice pro jednotlivé skupiny	25
5.1 Komunity řešící průtoková data	25
5.2 Komunity s datovým řešením postaveným na kolektivu	25
5.3 Komunity analyzující herní data ve spolupráci s vývojáři	26
5.4 Komunity analyzující herní data bez spolupráce vývojářů	26
6 Závěr	27
6.1 O práci	27
6.2 Úroveň realizace zadání	27
6.3 Výsledky práce	28
6.4 Budoucí směřování návazných prací	28
A Dotazníky případových studií	29
A.1 Prázdný dotazník	29
A.1.1 General part	29
A.1.2 General part regarding data	30
A.1.3 Fast-flowing data (IRC, Discord, Slack, Reddit, etc.)	30
A.1.4 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)	31
A.1.5 Data from developers (game API, cooperation, etc.)	31
A.1.6 Data mining (data mining, reverse engineering, modding, etc.)	31
A.1.7 Etc.	32
A.2 Pokémon komunita	32
A.2.1 With regards to data	32
A.2.2 Fast-Flowing data	33
A.2.3 Collective permanent solution	34
A.2.4 Data from developers	34
A.2.5 Data mining	35
A.3 Doom speedrunning komunita	36
A.4 Elite: Dangerous API komunita	37
A.4.1 General part regarding data	37
A.4.2 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)	38
A.4.3 Data from developers (game API, cooperation, etc.)	38
A.4.4 Etc.	39
A.5 Eiyuu Senki komunita	39
A.5.1 General part regarding data	39
A.5.2 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)	40
A.5.3 Data from developers (game API, cooperation, etc.)	40
A.5.4 Etc.	41
A.6 Genshin Impact komunita, Torn komunita	41
A.6.1 General part	41
A.6.2 General part regarding data	42
A.6.3 Fast-flowing data (IRC, Discord, Slack, Reddit, etc.)	43
A.6.4 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)	43
A.6.5 Data from developers (game API, cooperation, etc.)	44
A.6.6 Data mining (data mining, reverse engineering, modding, etc.)	45
A.6.7 Etc.	45
A.7 Kantai Collection komunita	45

A.7.1	General part	46
A.7.2	Fast-flowing data (Reddit)	46
A.7.3	Fast-flowing data (en Kancolle Wiki Discord)	47
A.7.4	Collective permanent solution (en Kancolle Wiki)	48
A.8	Kantai Collection API	49
A.8.1	General part	49
A.8.2	General part regarding data	49
A.8.3	Data from developers (game API, cooperation, etc.)	50
A.8.4	Data mining (data mining, reverse engineering, modding, etc.)	50

Chtěl bych poděkovat především vedoucímu mé bakalářské práce za spolehlivé vedení a spolupráci.

Dále bych chtěl poděkovat své manželce za bezmeznou trpělivost s mým vypracováváním bakalářské práce z domova v době koronavirové krize.

A chtěl bych poděkovat synovi za motivaci, kterou mi každý den dává.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů, zejména skutečnost, že České vysoké učení technické v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 citovaného zákona.

V Praze dne 13. května 2020

.....

Abstrakt

Tato bakalářská práce se zabývá analýzou sběru dat a jejich následného použití v herních komunitách, kde se zaměřuje na data, která uživatelé jednak sami sbírají, tak následně sami používají, tedy nejde o data o hráčích jako takových. Cílem práce je rozlišit zacházení s daty v různých typech komunit, tyto následně kategorizovat a poté doporučit budoucím zájemcům – jak o vedení herních komunit, tak o programování datových řešení pro tyto herní komunity – buď konkrétní řešení nebo směr, kterým se při vylepšování sběru a užití dat ubírat.

Práce staví hlavně na konkrétních případových studiích, kdy se provedl cílený průzkum u správců primárně větších komunit ohledně užitých technologií a dopadu na jejich komunitu. Výsledkem je sada doporučení pro jednotlivé typy vývojářů napříč různými typy komunit, kde tato doporučení jsou výsledkem analýzy zmíněných případových studií.

Klíčová slova herní komunita, sběr dat, případová studie, herní analytika, interaktivní data, kolektivní výzkum, kolaborativní hraní

Abstract

This bachelor's thesis researches data collecting and use in gaming communities, where the focus is data collected and used by users themselves, meaning the thesis is not interested in data about players as such. The aim of this thesis is to differentiate how data are managed throughout various community types, to categorize them and then recommend interested future users – whether interested in leading gaming communities or programming data solutions for such gaming communities – either an exact solution or a direction in which to progress towards improving collection and use of such data.

The thesis is primarily based off of concrete case studies, where targeted research aimed primarily at administrators of bigger gaming communities was led regarding the used technologies and the impact on their community. The result is a set of recommendations for various types of developers across various community types, where these recommendations are result of mentioned case studies' analysis.

Keywords gaming community, data collecting, case study, gaming analytics, interactive data, collective research, collaborative gaming

Shrnutí

Motivace

Téma práce jsme navrhli jako zadání, jelikož se domníváme, že jde o pole, které si v současnosti zaslouží pozornost, a zároveň je do jisté míry přehlíženo jako pouhá zábava. Jde přitom o finančně stále rostoucí odvětví, v kterém se soustředí značné finanční i materiální prostředky, jejichž objem jen zvýšila pandemie COVID-19 v nedávné době. Z toho důvodu je určitě vhodné toto odvětví prozkoumat a nezapomenout jeho vývoj.

Cíl práce

Cílem práce je klasifikovat jednotlivé typy herních komunit podle toho, jak sbírají herní data a jak s nimi následně zachází. Pro každý klasifikovaný typ by mělo následně být možné vytvořit sadu doporučení, kterým směrem se vydat, či čemu se vyhnout. Z tohoto výsledku by následně měli těžit jak zájemci o vedení herních komunit, či vývojáři nástrojů pro herní komunity.

Postup

První částí postupu je teoretická analýza herních komunit dle typu sběru dat, a to primárně z dvou hledisek – doby trvání jejich uchování danou komunitou a spolupráce s herními vývojáři. Dále pak

bude provedeno několik případových studií formou dotazníku na administrátory hlavně větších herních komunit, které budou ověřovat, zda je toto předpokládané rozdělení platné a dále jaké důsledky z tohoto rozdělení plynou.

Výsledky práce

Výsledkem práce je primárně práce samotná, respektive průzkum učiněný autorem a sada doporučení pro jednotlivé skupiny komunit. Pro budoucí použití je pak přínosem i uvedený dotazník, který se dá případně použít pro opětovné provedení výzkumu v okamžiku, kdy se výsledky práce stanou zastaralými (například z důvodu vzniku nových řešení pro správu informací v herních komunitách).

Závěr

V rámci práce se nám povedlo provést prvotní analýzu a rozdělit si herní komunity do několika prolínajících se typů. Následně jsme provedli cílený průzkum formou případových studií mezi správci datových řešení pro jednotlivé komunity a vyrobili sadu doporučení pro jednotlivé kategorie. Obecně se cíle práce dají považovat za splněné s drobnými výhradami.

Seznam zkratek

AAA hra	Hra produkovaná středně až velmi velkým vydavatelem s velkými náklady na vývoj a reklamu [1]
AN	Author's Note
API	Application Programming Interface
CMS	Content Management System
Eroge	Erotická hra (z japonského エロゲーム, „ero gému“)
FPS	First Person Shooter
GaaS	Game as a Service, hra s průběžně vydávanými updaty bez jasně daného konce vydávání
Gacha	Prvek náhodné odměny, „lootbox“ (z japonského ガチャポン, „gačapon“, popisujícího zvuk původní fyzické podoby stroje dávajícího náhodnou odměnu za malý obnos)
GDPR	General Data Protection Regulation
KanColle	Kantai Collection
Mobage	Mobilní hra (z japonského モバイルゲーム, „mobairu gému“)
RTS	Real Time Strategy
Q&A	Questions and Answers
VOIP	Voice Over IP
VR	Virtuální Realita

Úvod do problematiky

*On my business card, I am a corporate president. In my mind, I am a computer programmer.
But in my heart, I am a gamer. [2]*

Satoru Iwata

Již od nepaměti byla lidstvu dána schopnost hraní a dodnes můžeme vidět schopnost hrát si nejen na lidech, ale i na zástupcích zvířecí říše. S vývojem schopnosti používat nástroje a abstrahovat se ale lidská herní schopnost výrazně vydělila a odlišila od té zvířecí.

Zhruba do poloviny dvacátého století jsme mohli hovořit o tom, že lidé hrají převážně hry fungující na principu vlastního těla (ať již verbální či motorické) nebo s pomocí fyzických pomůcek (dnes bychom řekli například hry deskové). S nástupem počítačů a počítačových her se ale proměňovalo celé odvětví her a za nepříliš dlouho – mluvíme o horizontu maximálně desítek let, což je v porovnání s celkovou historií hraní zanedbatelná doba – se vydělilo počítačové hraní jako samostatná zájmová i obchodní kategorie.

O první historickou počítačovou hru se vedou spory, ale značná část odborné veřejnosti se přiklání k tomu, že touto hrou je Spacewar! z roku 1962, do jisté míry kvůli komerčnímu rozšíření, jakkoliv malému [3]. Již o třicet let později bylo však hraní her prakticky globální záležitostí, mělo za sebou několik kontroverzí, zvyšujících ponětí o počítačovém hraní, a mnohem více jich ještě čekalo a stále čeká.

Devadesátá léta zaznamenala rozvoj hardware jako takového, a s objevením se her jako Doom (1993), Transport Tycoon / Deluxe (1994/95), Quake, Diablo (1996) či StarCraft a Half-Life (1998), které dokázaly tento hardwarový potenciál využít, se podíl počítače jako primárního herního zařízení na trhu rozšířil na úkor do té doby vedoucích arkádových strojů a konzolí. Od roku 2000 jsme pak mohli sledovat až dodnes trvající nárůst mobilního hraní, které v současnosti tvoří zhruba polovinu příjmu herního trhu (85 miliard dolarů v roce 2020), zbytek je rozdělen mezi trhy počítačové (40 miliard dolarů) a konzolové (33 miliard), s tím, že 7 miliard dolarů připadá na nové trhy (hraní na VR zařízeních, cloudové hraní).[4]

Dalším faktorem, který v úvodu musíme zmínit, je rozšíření internetového připojení. Zde se pohybujeme na vzestupu z 11 % v rozvinutých zemích roce 1997 [5] na úroveň odhadovaných 86,6 % v roce 2019 [6]. Tento rozvoj samozřejmě nezanedbatelně proměnil možnosti společného hraní, distribuci her a v poslední řadě i přístup hráčů ke konzumaci her.

1.1 Hráči a hry offline vs. online

Do konce devadesátých let (pokud se bavíme o zemích prvního světa, v ČR pak můžeme sledovat podle regionu pěti- až desetileté zpoždění) se hráč, který chtěl získat novou hru, musel spolehnout na velmi omezené zdroje. Nejtypičtějšími zdroji v této době byly zkušenosti dalších hráčů, typicky spolužáků či kolegů, a časopisy, věnující se herní problematice – v českém prostředí šlo o časopisy Excalibur (1991 – 2000), SCORE (1994 –) a LEVEL (1995 –). Jako další z typů vyzkoušení si hry před jejím zakoupením, které tehdy mohlo znamenat poměrně velký výdaj, můžeme zmínit i šíření demoverzí či dnes takřka zaniklý formát shareware kopií, mnohdy více her na jedné disketě či CD.

Po roce 2000 jsme ale mohli sledovat proměnu herního trhu. Jedním z hlavních hybatelů změny se v roce 2003 stala společnost Valve vydáním jejich online distribuční platformy Steam, která v roce 2013 představovala 75% podíl digitálního trhu s hrami [7] a dnes představuje největšího distributora na tomto trhu, kdy jeho jediným reálným konkurentem je čínská společnost Tencent a její platforma Epic Games, do které Tencent investuje výrazné množství zisku z jejich úspěšné FPS hry Fortnite [8].

Proměna trhu se týkala i množství her – v současnosti je výrazně snazší hru vydat prakticky globálně, kdy například Steam si za vydání hry účtuje jednorázový poplatek 100 dolarů a následně procenta z prodeje. To vedlo v poslední dekádě k výraznému navýšení objemu her a jejich diverzifikaci, kdy dnes Steam nabízí přes 11 000 her.

Spolu s herním trhem se změnil i přístup hráčů ke hrám jako takovým. Tam, kde bylo dříve nutno si pečlivě před koupí vybírat a sledovat recenze, je po nástupu internetu výrazně jednodušší si hru prověřit - ať už v online recenzi dostupné snadným vyhledáním, na YouTube ve videorecenzi či se hráč může podívat na Twitch, zda hru někdo nehraje a dokonce se hráče přímo optat na jeho názor.

1.2 Dostupnost informací o hrách

Již jsme zmínili, že v devadesátých a nultých letech nehrál internet kvůli dostupnosti velkou roli. Kromě ověření kvality her se toto projevovalo i v dalším aspektu hraní her, a to informacích o nich. U většiny her mohl a stále může nastat případ, kdy se hráč „zasekne“ – v nějaké chvíli se zastaví jeho postup hrou a hráč neví, jak pokračovat. Důvodů je více, u žánru adventur z devadesátých let se často stávalo, že hráč přehlédl na některé obrazovce objekt, bez kterého nemohl pokročit ve hře dále, u FPS či RTS žánru mohlo jít o konkrétní taktický či strategický postup ve hře a v neposlední řadě mohla ve hře být chyba, která znemožnila hru dohrát bez ohledu na postup - můžeme zmínit například tituly Fallout 2 (1998) či Septerra Core (1999), které byly v mnoha případech bez oprav nedohrately.

Pokud hráč v tomto období nechtěl spoléhat v některých případech na náhodu, musel se spolehnout na již zmíněné informační kanály – ať to byli jeho spolužáci, kolegové, či návody z časopisů nebo přiložených databází na CD. Právě tato CD, která kromě demoverzí (pozdeji i plných her) a důležitých oprav obsahovala i databáze cheatů (tj. kódů sloužících k získání výhody hráčem) či návodů, můžeme považovat za první prototyp informačních databází od hráčů k hráčům.

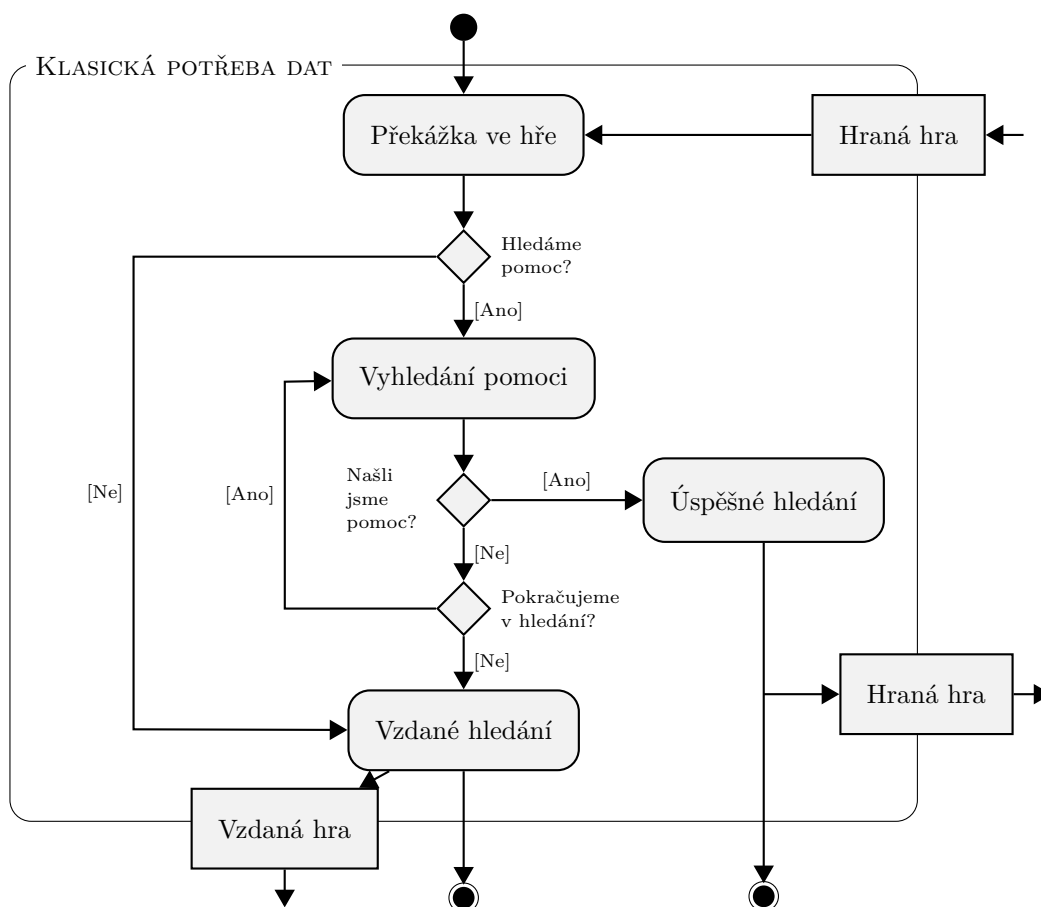
S nástupem internetu se ale tyto informace samozřejmě začaly přesouvat na internet. Nejprve šlo o přesun jedna ku jedné, kdy se mnohé tyto databáze prostě a jednoduše přesunuly z distribuovaných CD na vlastní stránku či stránku časopisu. Takřka zároveň ale s nárůstem titulů a zvětšením rozsahu jednotlivých her začalo docházet k tomu, že se začaly objevovat stránky dedikované jednotlivým hrám - ať už šlo spíše o fanouškovsky zájmové stránky či spíše techničtější

zaměřené stránky s informacemi technického rázu. Spolu s touto proměnou se umenšila i role herních časopisů, které se pomalu přesunuly spíše do role recenzenta, kdy nakonec jim po přesunu většiny herní distribuce na internet jiná role nezbývala.

Další stránkou online prostoru jsou pak sociální sítě – zde jde například o Facebook, vzniklý v roce 2004, který motivoval či nastartoval další sociální sítě jako alternativy. Bohužel díky své nepřehlednosti a nepředvídatelnosti hrál a hraje Facebook roli spíše marketingového média. Z nejdůležitějších sociálních sítí poslední dekadý musíme zmínit spíše streamovací a video sociální platformy jako YouTube a Twitch, prezentační sociální platformy jako Twitter (který hraje výraznou roli spíše v herní žurnalistice) a komunikační sociální platformy, kde zmíníme Reddit či Discord. Všechny tyto platformy – zvláště v době pandemie COVID-19 – nahradily či přinejmenším rozšířily ustálené osobní kontakty se spolužáky, kolegy a přáteli.

1.3 Vyhledávání dat hráči

Jak jsme nastínili, v průběhu let se změnil způsob oběhu dat v hráčských komunitách. Abychom tento oběh a jeho změnu přesněji pochopili a mohli s ní pracovat, musíme nejdříve analyzovat, k jakému posunu ve vyhledávání dat hráči mezi cca. začátkem devadesátých let a dnešním dnem došlo.



Klasický sběr dat byl relativně minimální, jednak kvůli decentralizaci spojené s absencí internetu, jednak kvůli absenci poptávky, která byla pro změnu spojena s nižší dostupností a popularitou her. Ve výsledku se tak potřeba dat jednotlivých hráčů naplňovala případ od případu – hráč začal hledat data až v okamžiku, kdy je potřeboval a konkrétně věděl, která data to jsou. Tuto akci vyhledávání shrnuje následující diagram aktivit na předchozí straně.

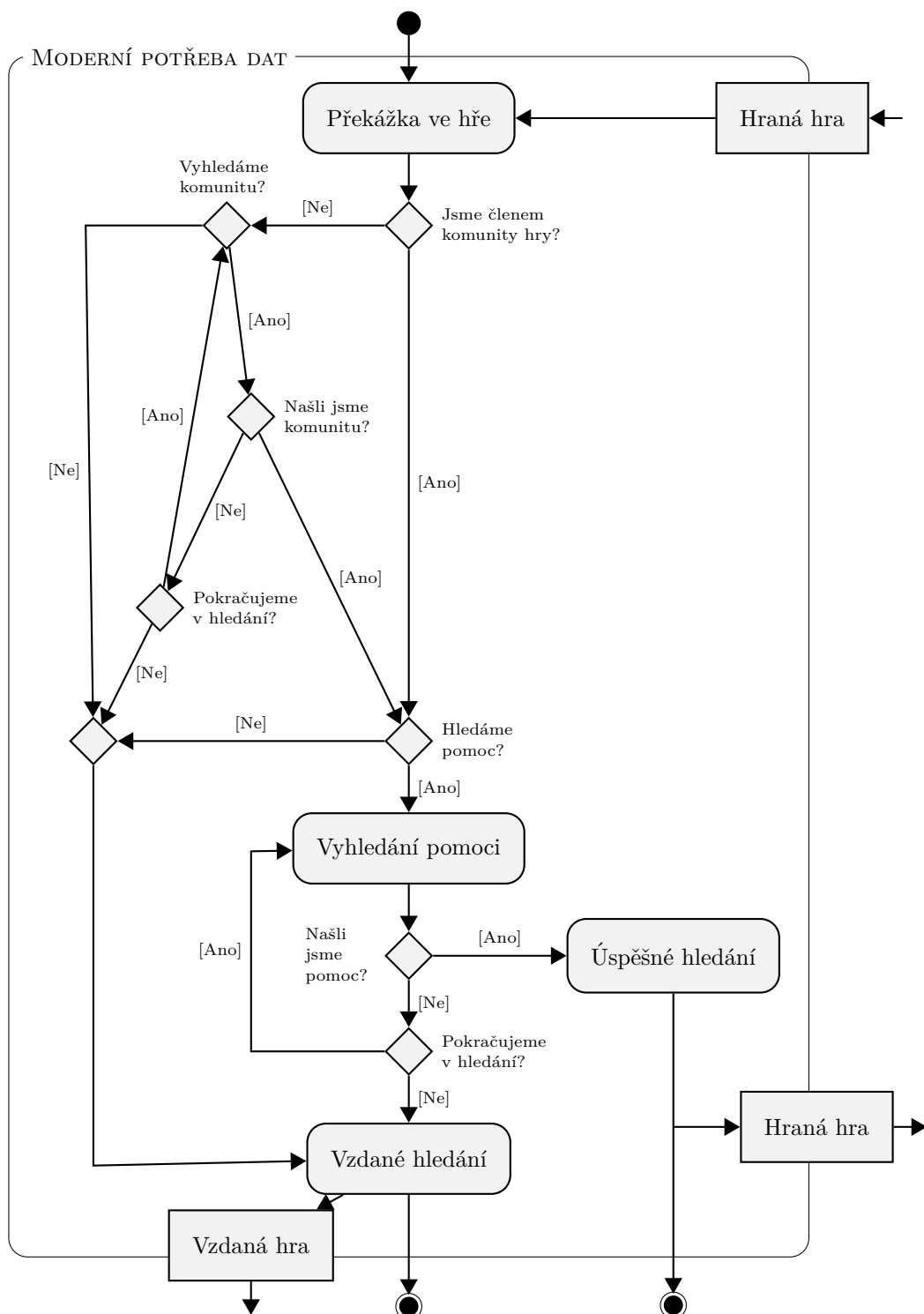
Tento diagram je logicky vcelku univerzálně použitelný pro více typů nízkoúrovňových hledání, proto by ho s mírnou obměnou bylo možné použít například pro druhý typ hledání informací, které dříve hráč potřeboval najít ve své přirozené sociální skupině, a to informaci, zda se vyplatí koupě té které hry. Jelikož tehdejší cena za hru byla vyšší než dnes (a to i pokud započítáme běžnou míru inflace, kdy docházíme k ceně 80 - 100 \$ za hru v druhé půli devadesátých let a k ceně 60 \$ dnes [9]), a míra dostupných informací malá, byl nutný důkladnější výzkum, zvláště před zakoupením AAA hry. V takovém případě můžeme v diagramu nahradit pojmy následujícím způsobem:

- hranou hru nahrazuje hra zvažovaná ke koupi
- hráč namísto pomoci hledá informace, zda se hra vyplatí koupit či ne
- výsledkem pak je buď hra zhodnocená (ať už zakoupená nebo ne) nebo hra, ke které se hráč nedozvěděl dostatek informací a bude muset v případě koupě jít do rizika.

Po nástupu internetu se ale aktivita, kterou jsme na předchozí stránce popsali diagramem, začala proměňovat. Optání se svého sociálního kruhu bylo sice stále dostupné, ale postupně bylo čím dál více v pozadí. Bylo to způsobeno primárně vyšší nabídkou her, kdy se při předpokladu omezeného okruhu přátel nedalo očekávat, že by většina z nich hrála naprosto stejné hry jako my a byla schopna poradit. Tento fyzický sociální svět se tak spíše zúžil na doporučení her, které se líbily nám a případnou následnou diskusi při dalším setkání, zda se doporučená hra líbila tomu, komu jsme ji doporučili.

Do popředí se tak dostaly internetové diskuse nejrůznějšího typu. Zde ale hráči narazili na problém, že hráč nemusí být při naražení na překážku ještě členem herní komunity pro danou hru. Tím se konečně dostáváme k vlastní proměně, ke které došlo v aktivitě hledání herních dat, které sloužily a slouží k pomoci hráči – hráč musel před začátkem vlastního vyhledávání navíc vyhledat i patřičnou herní komunitu (pokud jejím členem ještě nebyl) nebo se obrátit na správnou herní komunitu. Tento druhý bod byl a je neméně důležitý a ve skutečnosti spíše nabývá na důležitosti, protože jeden hráč může kdykoliv být členem více herních komunit – je typické, že dnes je hráč členem značného počtu herních komunit a připojuje se do nových a opouští staré podle toho, které hry právě hraje. V extrémním případě, u her typu GaaS dokonce můžeme hovořit o stavu, kdy hráč je aktivní v té které komunitě podle toho, které hře právě vyšel update či event – obsah hry, který je časově omezený a vede k tomu, že se ho hráč snaží stihnout a dočasně je ve hře a v komunitě aktivnější.

Zde pak máme časem proměnný diagram aktivit, shrnující tyto změny:



Reakcí na tento nový způsob vyhledávání dat po sdružení se hráčů v herních komunitách byl logický krok, kdy tyto komunity začaly sbírat a přechovávat data o hrách – jednak bylo důležité přechovat data tak, aby se o ně nepřišlo, pokud by se obměnili hráči v komunitě, jednak se na ně dalo odkázat nově příchozím a ušetřit čas znovuobjevováním odpovědí na nejběžnější dotazy.

V různých komunitách se ale tato data sbírala a přechovávala různými způsoby. V praktické části se dále tyto způsoby pokusíme v první řadě kategorizovat a utřídit a dále si na případových studiích ověřit, jak s nimi jednotlivé herní komunity pracují a jestli je naše kategorizace funkční.

Rozdělení sběru dat v herních komunitách

The worst thing a kid can say about a homework assignment is that it's too hard; the worst thing a kid can say about a game is that it's too easy. [10]

Henry Jenkins

V teoretickém úvodu jsme zjistili, co komunity vedlo ke sběru dat. Nyní se přesuneme do současnosti a pokusíme se na základě různých herních i neherních vstupů analyzovat různé typy herních komunit a z toho vycházejících způsobů, kterými tyto komunity mohou zacházet se sbíranými daty.

2.1 Komunity bez sběru dat

Jako první si musíme všimnout jistého „slona v místnosti“ – samozřejmě existují jisté skupiny her a okolo nich sdružených komunit, které sběr dat nepotřebují. V dnešní době je to spíše zanedbatelná část dostupných her, ale nemůžeme si dovolit je nezmínit.

Důvodem, který může stát za tím, že okolo hry neexistuje komunita (a tím pádem bez této komunity nemohou ani existovat žádná data ke hře), může být její stáří, zapadlost, případně kombinace obojího. Je to běžný stav pro hry, které byly prodány v malém počtu, na ne běžně rozšířené konzole a v letech, kdy bylo rozšiřování her komplikované.

Logicky je složité uvést příklad takovýchto her, protože jsou zapadlé právě tím, že je nikdo nezná, za zmínku stojí například jihokorejská hra Animal Kingdom z roku 2003 od studia Dataway, ke které se dají dohledat jen zběžné zmínky na internetu, a to převážně o její ruské verzi od vydavatele 1C.

Čas od času může hra z této kategorie být znovuobjevena a mohou k ní vzniknout nové komunity a datové sklady, je to však relativně vzácný případ.

Druhý důvod, který může nastat je sice existující komunita, ale ke hře, která je tak primitivní, že ohledně ní není možné sbírat žádná data. Toto je případ některých raných her, včetně dříve zmiňované první hry Spacewar!, či některých vyřešených her jako jsou piškvorky na síti 3x3 nebo Nim, kde jediná dostupná informace je optimální strategie.

Tato kategorie se samozřejmě neomezuje pouze na staré či deskové hry. Můžeme zde zmínit například Flashovou hru *You Have to Burn the Rope* z roku 2008, kde jediným cílem je přepálit lano na konci hry. Přestože hra je oceňována za její píseň na konci hry, která je delší než samotná herní doba, a existují k ní návody a speedruny – pokusy dohrát hru za co nejnižší čas, často spočívající ve stovkách až desítkách tisíc dohrání za účelem snížení času o sekundy, či jen zlomky sekund –, samotná hra je tak jednoduchá, že návody jsou spíše vtipnou poznámkou k této jednoduchosti. Podobně *Barney's Hide and Seek* je hra vydaná v roce 1993 na konzoli Sega Mega Drive/Genesis a je očividně zaměřena na děti - po několika minutách se začne hrát sama a dohraje se za cca. deset minut sama i bez vstupu hráče, zhruba dvakrát pomaleji, než kdyby hráč skutečně hrál. Opět k ní existují speedruny, ale tyto spočívají buď čistě v doufání v jistou neovlivnitelnou sekvenci, která způsobí dohrání hry o několik vteřin rychleji (pokud jde o kategorii bez ovladače) či o několik jednoduchých sekvencí, které by postupem času odhalil i dedikovaný hráč. Hra je udržována spíše jako meme mezi speedrunovou komunitou a objevila se i na některých charitativních akcích.

Je velmi důležité v této části poznamenat, že jednoduchost hry musí být opravdu extrémní ve všech ohledech. Existují hry, které jsou velmi jednoduché ovládacím schématem, jako například tzv. idle hry či clicker hry, které nicméně dovolují pomalejší či rychlejší hraní, mohou dovolovat různé strategie a speedruny, mohou obsahovat skryté vlastnosti, atd. a pro mnoho z nich existují vzhledem k jejich jednoduchosti až překvapivě aktivní komunity a komplexní návody či wiki.

Posledním důvodem, který musíme zmínit souvisí mírně s předchozími, ale zaslouží si jisté vypíchnutí – jde o hry, které jsou vyloženě špatné a komunita k nim přistupuje s despektem. Je zde ale obtížné určit, kde je zde přechod mezi první, druhou a třetí kategorií, protože špatné hry bývají často i jednoduché či nedohrately a zároveň kvůli tomu zapadlé. Jako protipříklad zde ale můžeme zmínit některé hry, které jsou natolik špatné, že kolem nich vznikl kult, který je šíří mezi další hráče jako „nejhorší hry všech dob“ – jde o hry:

- *Big Rigs: Over the Road Racing* (2003), zpopularizovanou recenzí *Angry Video Game Nerd* z roku 2014, v které osobitým způsobem kritizuje špatnou kvalitu hry a nemožnost prohrát [11]
- *E.T. the Extra-Terrestrial* (1982) na motivy slavného filmu Stevena Spielberga, kterou proslavila její tak špatná kvalita, že se neprodaly statisíce kopií, které byly následně zakopány v poušti v Novém Mexiku. Toto zakopání bylo dlouhá léta považováno mnohými za městskou legendu, ale v roce 2014 v rámci dokumentu *Atari: Game Over* došlo k vykopání části těchto kopií a ověření faktu, že k zakopání opravdu došlo
- *Hareraiser* (1984), hra na množství tehdejších konzolí, která těžila z navázání na populární knihu *Masquerade* a měla vést ke zlatému šperku jako první ceně po „vyřešení“ hry. V roce 2017 byl její příběh zpopularizován Stuartem Ashenem v rámci *Norwich Gaming Festivalu*, kde se domnívá, že hra je neřešitelná a je prostým pokusem o obránění hráčů o peníze [12]. Každopádně hra jako taková nemá žádný smysl a směřování, které by vedlo k možnosti o ní schraňovat data, a tím pádem spadá do této kategorie.

Závěrem této kategorie je konstatování, že tuto skupinu z logických důvodů nebudeme nadále analyzovat a nebudeme pro ni vyvozovat žádné postupy či doporučení, jelikož by to jednoduše pravděpodobně nebylo možné, a i kdyby se nějak povedlo překonat překážky kladené absencí studovaných vlastností, bylo by to ku prospěchu naprostého minima uživatelů.

2.2 Komunity řešící průtoková data

První typ komunit, který můžeme vydělit a sledovat jejich práci s daty, jsou komunity, kterými data převážně jen protékají. Tato řešení obecně nahrazují komunikaci reálných sociálních skupin, a tak sice obsahují množství informací, ale vyhledávání těchto informací je ve většině případů spíše dodatečným prvkem.

Kořeny těchto komunit můžeme vysledovat až k nejstaršímu typu konverzační skupiny na internetu, tj. elektronickému mailing listu. Toto řešení je dnes už ale v herních komunitách prakticky nepoužívané a nahradila ho jiná řešení:

- **Diskusní fóra** – Okolo roku 2000 relativně obvyklé řešení, nejčastěji ve formátu phpBB, dnes již méně časté, kvůli nutnosti řešení hostingu, nastavení fóra, a následné správy samotného fóra. I tak dnes stále existují herní komunity používající toto řešení, typicky existují buď na fórech spravovaných herními vývojáři samotnými (jako příklad mohou posloužit fóra vývojáře Frontier Developments: <https://forums.frontier.co.uk/>) nebo jsou podskupinou většího fóra, které se zabývá i neherními komunitami (například Games subforum na fórech serveru SomethingAwful.com), v obou případech tak odpadá řešení správy fóra. Velkou výhodou za náročnější správu bývá výrazně lepší systém hledání v datech.
- **Anonymní diskusní fóra** – Herní komunita jen ve velmi volném slova smyslu, jelikož anonymní princip do značné míry znemožňuje konzistentní komunikaci za rámec jednoho diskusního vlákna, na druhou stranu umožňuje anonymní dotazy pro hráče, kteří se bojí, že jejich dotaz bude považován za naivní či hloupý a poznamená to jejich reputaci v komunitě dané hry. Typickým zástupcem je boards.4channel.org/vg a fóra podobného typu, k omezenému vyhledávání v rámci některých vláken slouží jejich externí zálohování.
- **IRC** – Alternativa k fórům, dnes již také poměrně opouštěná. Nastavení IRC místnosti je o poznání jednodušší než nastavení diskusního fóra, na druhou stranu s větším počtem účastníků může vést k chaotické diskusi. Nevýhodou je, že pro mnohé uživatele představuje pokročilé užívání IRC technickou překážku, přes kterou se nejsou schopni přenést. K nastavení vyhledávání v historii konverzací je navíc třeba externě logovat diskusi.
- **Facebook** – Některými komunitami používané řešení hlavně po vzniku Facebooku v roce 2004, dnes se tyto komunity typicky pohybují spíše okolo společností či populárních streamerů dané hry, které mají ke hře založenou stránku. Nevyžaduje od uživatelů kromě Facebook účtu žádnou investici, na druhou stranu je vysoce nepraktický k vyhledávání a často nevrátí ani informace, které v historii diskusí prokazatelně existují.
- **Reddit** – Celkem populární současné řešení k diskusi, fakticky mezistupeň mezi diskusními fóry a Discordem. Je snadné založit subreddit k té které hře (či dokonce jen k části nějaké hry) a nedávne diskuse jsou snadno dohledatelné. Na druhou stranu vyhledávací systém je podobně nekonzistentní jako v případě Facebooku a mnozí uživatelé užívají spíše vyhledávání přes Google.
- **Slack a Discord** – Slack je původní firemní řešení z roku 2013, které nabízí fakticky IRC server se založitelnými místnostmi pro kterékoliv uživatele. Velkou nevýhodou tohoto řešení, motivujícím firmu k platbě, je historie omezená na počet zpráv, tedy větší komunity narazily na tento limit vcelku rychle. I tak ho okolo roku 2013 ho některé herní komunity užívaly, ale v roce 2015 zkopírovala toto řešení firma Discord Inc. (tehdy Hammer & Chisel, Inc.) bez omezení historie. Discord má jiná omezení, ale spíše kosmetického charakteru (počet emoji

na server, užívání animovaných emoji, atd.). To mu dovoluje, spolu s nabídnutím videohovorů, Twitch integrace a VOIP řešení, být v současnosti de facto standardním řešením pro herní komunity. Oproti některým řešením výše navíc musíme zmínit velmi solidní (zvláště s ohledem na počet zpráv v některých komunitách) vyhledávání v historii.

Celkově v této kategorii můžeme v čase sledovat snahu o přechod k jednoduššímu nasazení a vytvoření komunity a snahu o to, aby užívání platformy bylo co nejjednodušší pro uživatele – zatímco například v případě diskusních fór si musel uživatel pro každé fórum pamatovat přihlašovací údaje a sjednocení všech her pod jedno fórum bylo nereálnou představou, na Discordu je uživatel reprezentován jediným účtem pro všechny servery, na kterých diskutuje, přestože jde o rozdílné hry a komunity.

2.3 Komunity s datovým řešením postaveným na kolektivu

Druhým typem komunit, který můžeme analyzovat je typ, který jsme si provizorně nazvali „datové řešení postavené na kolektivu“. Jde o řešení, které se spoléhá na to, že jistá podmnožina uživatelů bude vkládat znalosti do nějakého rozhraní, ke kterému se pak zbytek komunity bude moct dostat a získat z něj data zpět.

Toto řešení má jisté výhody a nevýhody. Mezi nevýhody patří samozřejmě to, že se toto rozhraní musí spravovat – musí se založit, musí se do něj vkládat data (na rozdíl od prvního typu, kde se vkládají data samotnou komunikací) a musí se kontrolovat korektnost vkládaných dat. Na druhou stranu získáme nezanedbatelné výhody, kdy jsou data výrazně lépe vyhledatelná, organizovaná a ověřená; to vše samozřejmě v ideálním případě.

Počet užívaných řešení v této kategorii je výrazně nižší, prakticky jde o tyto tři přístupy:

- **Fandom.com (do roku 2016 Wikia.com)** – Nejběžnější řešení, pod doménou fandom.com může kdokoli založit vlastní wikipedii a subdoménu běžící na systému MediaWiki k libovolnému fandumu a ta může být následně kýmkoliv editována se zachováním historie a případnými kontrolami, v typickém wikipedie stylu. Pro svoji jednoduchost je to řešení velmi oblíbené a i přes jistá omezení a problémy zde můžeme najít velmi rozsáhlé sbírky dat fanouškovských komunit - například nejrozsáhlejší komunita okolo série The Elder Scrolls zde k 9. květnu 2021 reportuje 67 856 stránek. Nevýhodou může pro některé uživatele být nepředvídatelnost služby, kdy vlastník fandom.com rozhoduje o tom, jak bude celá služba vypadat a které vlastnosti a pluginy systému MediaWiki bude používat.
- **Vlastní wikipedie** – Alternativa k fandom.com, kdy si komunita založí wikipedii vlastní. Nevýhodou zde je nutnost řešit hosting a nastavení, na druhou stranu si komunita může zafixovat jistou verzi a vlastnosti, které vyžaduje. Nejběžnější volbou co do stránky softwaru jsou zde MediaWiki a DokuWiki, primárně díky napsání v PHP, což dovoluje jejich snadné rozšíření i mezi méně technicky zdatnou základnu.
- **Vlastní CMS** – Mezi čistě obecně herními komunitami řešení méně časté, na druhou stranu je toto preferovaný přístup například pro herní klany, které přes CMS řeší dva problémy - jak skladování dat pro nové i staré hráče, tak vlastní prezentaci. U některých větších GaaS her se pak můžeme setkat s fanouškovskými blogy, které obvykle též užívají toto řešení, za zmínku stojí například thedailybounce.net sledující novinky ke hře World of Tanks nebo elitedangerous2016.wordpress.com přehledně shrnující historii a plánovanou budoucnost hry Elite: Dangerous.

V této kategorii musíme zmínit i řešení, kterým se věnovat nebudeme, protože jsou vysoce individuální a nejsou dobře sledovatelná bez hluboké analýzy, která je nad časový a obsahový rámec této práce - a to video řešení YouTube a Twitch. Zde máme sice zachovatelná a vyhledatelná data, která jsou ale generována způsobem a v rozsahu, který je jen složitě analyzovatelný. Pro zbytek práce je tedy z datového rozsahu vyloučíme, i když samozřejmě mohou spadnout pod některou z výše zmíněných kategorií (například YouTube video na fandom wikipedii či Twitch klip na Discordu).

2.4 Komunity analyzující herní data ve spolupráci s vývojáři

V třetím typu sběru dat herní komunitou se dostáváme k relativně specializovaným datům - jde o data ve spolupráci s vývojáři. Zde nalezneme jen dvě hlavní kategorie, buď mohou vývojáři poskytnout hráčům nějaké API, či mohou spolupracovat na volné úrovni.

Pokud vývojář poskytuje ke hře API, jde typicky o promyšlenější řešení, kterým se vývojáři snaží hráčům zprostředkovat přístup k některým datům. Zamýšlený dosah se přitom může různit - může jít o jednoduchou implementaci informačního rozhraní jako jsou nejvyšší skóre a veřejná data k účtům, může jít o komplexnější datové informace, a v neposlední řadě může jít o komplexní vhléd do struktury hry pro podporu tvorby modifikací. Tato řešení se tedy obvykle řeší případ od případu, i když sdílejí jisté společné prvky, vycházející z jejich podstaty (možný update cyklus, jak se k nim přistupuje, atd.)

Spolupráce na volné úrovni bývá nejčastěji realizována jednou z předchozích kategorií, dejme tomu blogem vývojáře nebo Q&A sekcí na fórech vývojáře. Může ale jít i o některé nezmíněné možnosti, jako například přímá komunikace s hráči (kteří následně mohou data zveřejnit na běžném datovém řešení) nebo prezentace dat pro hráče přímo ve hře, způsobem, který je jasně míněn hráčům jako takovým spíše než herním postavám (tedy vývojář používá tzv. metagaming).

2.5 Komunity analyzující herní data bez spolupráce vývojářů

Posledním typem sběru dat je analýza herních dat bez spolupráce vývojáře – konkrétně v této části máme na mysli data, ke kterým se běžný hráč nedostane, tedy analýza herního kódu, reverse engineering, přístup k neveřejnému API vývojáře, analýza souborů pro datamining a získání leaků, atd. Tuto kategorii vydělujeme primárně proto, že často nebývá pokryta na předchozích řešeních, případně pouze v oddělených částech (soukromá místnost na Discordu či IRC, separátní subreddit, atd.), aby se některé citlivější části nedostaly mezi širší komunitu, kde by mohly způsobit více škody než užitku. V mnoha případech jde o data, která ve skutečnosti tak citlivá nejsou, ale v případě, že by byla, mohla by způsobit problémy, tedy se jedná o jistou prevenci těchto skupin.

Kapitola 3

Případové studie

By the time there is a case study in your specific industry, it's going to be way too late for you to catch up. [13]

Seth Godin

3.1 Průzkum a metodologie

Počáteční průzkum byl proveden přes vyhledávání službou Google, variací klíčových slov této bakalářské práce (tj. různé kombinace frází gaming community, data collecting, case study, gaming analytics, interactive data, collective research, collaborative gaming). Přes nalezení několika prací na podobnou tematiku jsme nenalezli žádnou, která by jí vyhovovala v dostatečné míře, aby nám posloužila například k citaci. Asi nejpodobnější z nalezených byla práce s názvem Player Dossiers: Analyzing Gameplay Data as a Reward [14], ale i ta se zabývala spíše motivací hráčů pomocí dat. Drtivá většina nalezených prací pak pracovala spíše s hráči jako daty a s tím spojenou problematikou (úniky dat, hodnota takových dat pro společnosti, atd.)

Vzhledem k absenci relevantních dat jsme se rozhodli pro cílené případové studie formou dotazníků zaslanych konkrétním správcům jednotlivých komunit. Z časových důvodů (beroucí v ohled kontaktování tázaných, čas na odpovědi, čas na zpracování a vypracování) jsme kontaktovali zhruba deset správců různých komunit napříč herní scénou, a to jak větších a ustálených, tak menších a novějších. V tuto chvíli nastaly jediné vážnější problémy, kdy nebylo akceptováno sto procent dotazníků a bylo potřeba doplnit je komunitami dalšími. Obecně ale platí, že pokud mohli, byli tázaní vcelku ochotní a vstřícní odpovídat.

V této části práce budeme italikou vyznačovat citace z přílohy A, tedy dotazníků případových studií. Každému dotazníku odpovídá jedna herní komunita, jejímž názvem budou citace uvedeny, s výjimkou komunity Kantai Collection, kde byly dotazníky vyplněny dva, kde ale tento fakt nijak nesnižuje snadnou dohledatelnost citací.

3.2 Komunity řešící průtoková data

První a nejrozsáhlejší kategorií, kterou dotazník pokrýval, jsou komunity, řešící průtoková data. Zde můžeme prohlásit, že pod tuto skupinu více či méně spadá každá komunita, kdy jsme se nesetkali z žádnou z nich, která by tuto problematiku neřešila.

Za zajímavé v této části můžeme prohlásit, že služba Discord se v současnosti prakticky stala de facto standardem komunikace v oblasti herních médií. Setkali jsme se jen s velmi málo velmi drobnými komunitami, které využívaly alternativních komunikačních kanálů (IRC, vlastní fóra), a tyto bohužel navíc reagovaly spíše pomaleji a nejsou ve studii zahrnuty. Kromě toho, že každá komunita ve studii se nachází na Discordu, je ještě potřeba zmínit, že na něm došlo i ke kontaktování všech tázaných. Některé odpovědi byly řešeny přes e-mail, ale všichni tázaní se nacházeli na Discordu a u všech zde došlo k prvotnímu kontaktu, byť u některých komunit to bylo přes automatického bota pro kontaktování moderátorské skupiny jako celku.

Co se vlastních získaných dat týká, názory správců se liší ohledně konkrétní užitečnosti Discordu (Doom Speedrun: *This is used for communication, but doesn't offer much in terms of searching* oproti Eiyuu Senki: *(Our community uses data from the Discord, AN) A lot.*), ale ve výsledku se obecně shodují, že Discord je dostatečně robustní a užitečný napříč různými velikostmi komunit a v současnosti jeho řešení nemá přes jisté výhrady konkurenci (Genshin Impact: *I personally think other company-based competition sucks. Specific fan tools that serve specific needs, however, are great and I would love to incorporate a soft version of those tools so that this project is the one stop shop for all the needs.*).

Za jedinou reálnou alternativu v současnosti můžeme považovat reddit (KanColle: *reddit is the frontpage of the internet (unironically), so having a subpage on there for the community is often the first stepping stone for someone interested in the topic to find more resources* a Pokémon: *Reddit was chosen as it was clear that it would surpass Bulliten Board style sites in general, and was a powerful free platform.*), kdy za jeho primární výhody můžeme považovat relativní stálost dat oproti Discordu a koncentrovanější diskusi.

Obě uvedené služby navíc mají velmi jasnou výhodu snadné škálovatelnosti (Pokémon ohledně možného přesunu: *As far as I am aware no, it's difficult to with communities in the 6 or 7 digit member counts relocate, and Reddit and Discord are already sufficient for our needs.*), kterou žádná ze starších služeb nenabízí v takové míře, což představuje pro správce jak důvod pro přesun na tyto platformy, tak důvod nechuti pro případný přesun jinam. Jakákoliv nová konkurence by tedy musela nabízet výrazně větší jednoduchost užití při zachování obdobně silné škálovatelnosti.

Mezi hlavní nevýhody Discordu a Redditu pak patří dle předpokladů omezená možnost vyhledávání (Doom: *(Discord, AN) doesn't offer much in terms of searching* a Genshin Impact: *Discord has a pretty decent indexing feature, but as a programmer, I wish I could do more with the data.*), kterou některé komunity suplují vyhledáváním externím (Eiyuu Senki: *For searching data, we refer players to the JP wiki and use google translate for navigating*). Dále trochu překvapivě omezená správa rolí, která by ale údajně měla být v blízké budoucnosti vylepšena (KanColle: *Teamspeak has better ways to distribute and manage rights along the Server members (compared to Discord, AN* a Pokémon: *Discord is lacking in the moderation front, with the majority of holes patched by automated processes such as bots developed and maintained by the community*) a v neposlední řadě je – v případě dotázání se – mírně zmiňována absence záloh (KanColle: *n/A, it isn't possible to backup Discord Servers. No great effect as it is supplementary to the wiki* a Genshin Impact: *I personally don't care that the data is lost. What I do care about is that the company hosting the solution is no longer reputable at this point.*).

Co se největší výhody Discordu a Redditu týká, je to pravděpodobně fakt, že jsou to uživateli známé služby, s kterými se většina z nich setkala a ví, jak k nim přistupovat. Zde musíme zmínit velmi výstižný komentář z komunity Genshin Impact: *I didn't pick the solution: I was drawn to it. With custom implementation methods, they become more tech savvy, very niched, and might get lost in the WWW if it's not well designed/gets enough traction. Using popular tools, like Discord and Fandom, shortcuts the recognition (at the cost of losing direct control over the data).*

Tento komentář přesně vystihuje současnou situaci, kdy se kruhově posilují nejsilnější služby už jen faktem, že je uživatelé používají, a vyžadují, aby nové komunity byly založeny v rámci stejné služby. Postupem času je tak stále náročnější prorazit s novým, menším řešením.

3.3 Komunity s datovým řešením postaveným na kolektivitu

Druhou nejčastější kategorií vyplněnou v dotazníku bylo datové řešení, které jsme pro jednoduchost nazvali kolektivním. Jde převážně o řešení ve formátu fandom wikipedie, s některými konkrétními výjimkami.

Nejprve se budeme věnovat fandom wikipedii, tedy službě dostupné na webové stránce fandom.com. Zde mezi její výhody patří jednoduchost a přímocíarost jejího užívání, spojená se zvykem užívání její markdownové syntaxe (Genshin Impact: *My big deciding factor is the startup/maintenance cost of any project I start. As for similar solutions, I've known Fandom's system ever since 2014, and their system is one of the best systems to easily manipulate every aspect of the data*). Co se dat týká, v současnosti mezi preference patří spíše obecná data, která nejsou aktuálně tzv. „v kurzu“, protože u těch se vyplácí držet je spíše na živých službách prvního typu (Pokemon: *(Data on Discord and Reddit are, AN) Usually information about what pokemon are obtainable, and new characters is obtained at this stage, rather than specific niche data which is stored on wikis and database sites.*

Toto řešení má samozřejmě i své nevýhody, mezi kterými musíme zmínit hlavně zpoplatnění služby reklamami a nemožnost ovlivnit směr vývoje služby (KanColle: *Because of ads and the global developers ruining the back and frontend of the Fandom wiki, they migrated to the en Kancolle wiki. I dont see a reason to switch from the current platform a Torn: I have considered fleeing to Fandom, but as this is the official wiki, I am now striving to make it a better place.*) a omezenou možnost detailního ovlivnění vzhledu (Genshin Impact: *More control over the design would be ideal: making sure that the look of the site services several different platforms (desktop, mobile, print)*).

Tím se dostáváme k alternativě k fandomové wikipedii, a tou je vlastní instalace wikipedia řešení. V dotazovaných komunitách šlo primárně o MediaWiki řešení (KanColle: *For displaying Content, we are using Mediawiki*). Z pozorování a dotazníků jde zhruba odvodit, že vlastní instalaci preferují technicky zdatnější vývojáři, případně větší komunity. Zajímavostí je instalace Bulbapédie v Pokémon komunitě, která je velmi stará (říjen 2004), vznikla prakticky souběžně se službou fandom a svým rozsahem (přes 44 tisíc článků) se vyrovná největším fandomovým komunitám.

Poslední možností je kompletně vlastní řešení skladování kolektivních dat, buď vlastnoručně napsané nebo řešené přes CMS. Je to řešení méně časté, hlavně pro časovou náročnost, proto je využíváno spíše pro specifické potřeby, které se liší případ od případu (Doom: *The demo archive tracks all the runs and has some basic information in it, but it's primary purpose is preserving history in the runs themselves*). Pokud řešení spadá čistě pod CMS, bývá spíše přehledem informací v čase, na způsob novinového serveru. Zde se nám nepodařilo kontaktovat žádného autora CMS, ale již z náhledu například na <https://thedailybounce.net/> – běžící na CMS Wordpress – je toto snadno analyzovatelné i bez potvrzení.

3.4 Komunity analyzující herní data ve spolupráci s vývojáři

Překvapivě i v dnešní době je tento přístup relativně málo rozšířený. Z dotazníků se nám podařilo najít jedinou aplikaci tohoto přístupu, a to cíleně – konkrétně jde o Elite: Dangerous API. Na trhu samozřejmě v tomto stylu funguje více her, ale spíše jde o sdílení rozhraní k tvorbě modů (modifikací), publikování vyloženě herního API je věc méně častá, i když se samozřejmě může objevit, typickým příkladem v tomto ohledu je hra Minecraft [15].

Je do jisté míry pravděpodobné, že za důvodem zveřejnění herního API vývojáři Frontier Developments stojí fakt, že hra jako taková je procedurálně generována – vyskytuje se v ní na 4 miliardy hvězdných systémů, z toho přes třicet tisíc jich je obydlených. V této situaci se zájem vývojářů logicky soustředí na robustnost hraní a procedurálního generování a zbývá méně času a sil pro správu okrajových mechanik, jako je například vyhledávání komodit ve vesmíru pro obchodování nebo koordinace hráčů.

V takové chvíli bylo logickým krokem vývojářů vydat API přístupující k částem hry (Elite API: *Obvykle pohodlnější, rychlejší a ucelenější přístup k informacím než přímo ve hře, plus přidaná hodnota v podobě dalších výstupů (vyhledávání, přehledy, statistiky, atp.) a uživatelského obsahu (galerie, logbooks, diskuse, atp.)*), které mohou zpracovat hráči externě a zpřístupnit je jiným hráčům (Elite API: *Podle velmi konzervativního odhadu (s ohledem na sporá oficiální data o počtu aktivních hráčů) je Inara pravidelně používána minimálně 20% hráčské základny. Realisticky spíše více.*) ve formátu, který bude mít navrch nad vývojáři, jelikož autoři webového rozhraní s tímto mohou mít více zkušeností, na vývoj mají více času než vývojáři a nemusí se omezovat integrací do hry samotné, což přináší značnou úsporu času (Elite API: *Inara nabízí možnost jak dané informace na základě vstupních požadavků rychle najít, pohodlněji než hra samotná. Řada informací není také ve hře snadno či uceleně přístupná.*).

Je vcelku zajímavé, že touto cestou se v současnosti nevydává více vývojářů, zvláště v okamžiku, kdy si tuto spolupráci pochvalují všechny zúčastněné strany (Elite API: *Zajímavostí ohledně 3rd party vývojářské komunity kolem Elite: Dangerous je poměrně provázaný ekosystém, kdy například lokální uživatelské nástroje jako EDMC, EDDiscovery, EDDI a další kromě jiného i posílají část dat přes ZeroMQ dalším nástrojům a ty pak naopak mohou poskytnout nějaké informace zpět. Stejně tak mezi jednotlivými nástroji a websites do určité míry funguje (volitelná) provázanost přes vlastní API, exporty dat podobně. Domnívám se, že v porovnání s drtivou většinou nástrojů a komunit jiných her, které jsou obvykle striktně izolované, je to poměrně unikátní.*). Je ale možné, že se tento způsob šíření informací v budoucnu rozšíří u her typu GaaS, kde již můžeme sledovat jisté náznaky sdílení informací vývojářem nejen v podobě aktualizací, ale i živých informací (například mapa ke hře Genshin Impact vývojáře Mihoyo, přeložená do třinácti jazyků: <https://webstatic-sea.mihoyo.com/app/ys-map-sea>).

3.5 Komunity analyzující herní data bez spolupráce vývojářů

Posledním typem, který jsme v dotaznících analyzovali byly komunity, které se spoléhaly na analýzu her bez pomoci vývojářů. Zde jsme se zaměřili na tři v teorii zmíněné varianty - jednak využívání soukromého vývojářského API (KanColle: *My team manages a Chrome extension (KC3Kai) that parses data from the game API as the user plays and presents extra information that the game does not present. Some of this data is submitted to a database (TsunDB) for research on game*

mechanics by user consent.), jednak průzkumu herního kódu jako takového (Doom: *All the game data is readily available, from basic delineation of what is there all the way to modern editing tools fully exposing all the content.*) a jako poslední variantu jsme zběžně analyzovali reverse engineering původní hry Diablo – toto ne formou dotazníku, ale z veřejně dostupného videa o tomto výkonu.

Využívání soukromého vývojářského API je zajímavou variantou v případě, že je dostupné. Jde navíc o jakousi „šedou zónu“ v případě, že vývojář chce spolupracovat například se zahraničními hráči a není mu to umožněno vydavatelem, což je názor, který se neformálně objevuje v KanColle komunitě.

Na druhou stranu je v takovém případě logicky spolupráce výrazně jednostranná a jakýkoliv problém, i nezávinný uživatel API, může zkomplikovat spolupráci. Tyto problémy se pohybují od drobných a překonatelných po stav, kdy se API už ve skutečnosti nedá vůbec využít (KanColle: *The developers have recently implemented an IP ban for users outside Japan, and so obtaining metadata from the developers have become more difficult.*).

Důležité je také zmínit, že i když je tento způsob v této kategorii relativně jednoduchý, stále jde o dosti náročnou práci, která může být takřka na hranic využití jedincem a praktičtější bývá využít týmovou spolupráci (KanColle: *Initially, I used these to personally collect data and form my own analysis. It was extremely time consuming to obtain enough data for meaningful analysis. Later on, I collaborated with Tsubakura to implement a community submission feature in KC3Kai to submit data to TsunDB. This has made it a lot easier to collect the data I need.*). Odměnou za tuto snahu ale bývá jasný přínos pro komunitu v podobě přesných interních dat ze hry, jako jsou hodnoty zdraví, zranění či přesné navigace mapami, kterážto data mohou být do jisté míry záměrně vývojáři skrývána (KanColle: *KC3Kai helps users calculate how changes they make might affects gameplay. For example, it helps calculate potential damage or hidden variables such as accuracy and evasion per strategy setup. I think that the reason why this data is not provided by the game is because of the initial appeal of the game: a community driven approach to find how the mechanics work.*).

Co se průzkumu kódu jako takového týká, kontaktovali jsme Doom speedrun komunitu, která se ve hře silně vyzná. Tato znalost je takové míry, že je pro hráče, který se ve hře nevyzná na úrovni kódu, prakticky nemožné například uchopit jisté herní mechaniky nebo rozpoznat falešný speedrun od reálného, na což ve svém videu upozorňuje známý YouTube speedrunner Karl Jobst [16].

Bohužel z praktického hlediska se nenaplnilo naše očekávání, že tato data jsou v komunitě nějak specificky držena – jsou sice dostupná na wikipedii, ale od hráčů se spíše očekává přímý přístup do kódu a vyznání se v něm, wikipedie je tedy spíše prvotním návodem, který nasměrovává hráče k základnímu porozumění (Doom: *There isn't a speedrunning-focused wiki or anything like that, so access to information via this kind of resource is limited in a lot of ways.* a Doom: *This data is available on the wikis but any serious runner is familiar with the editing tools and will open up the game data directly if they want to check some specifics.*).

Jako poslední zmíníme případ kompletního reverse engineeringu hry, konkrétně prvního Diabla [17]. V tomto případě se jednalo o náročnější analýzu pro obecný zájem komunity (vypuštění staré hry ve formátu open source a její případné portování například na přenosné konzole).

Bohužel tento příklad můžeme probrat spíše pro zajímavost, jelikož jde o nesmírně náročnou činnost, která v tomto konkrétním případě byla vývojáři výrazně usnadněna, a to neúmyslným zahrnutím ladícího .sym souboru (obsahujícího mimo jiné jména funkcí a proměnných) ve verzi pro PlayStation 1. Bez tohoto souboru by reverzní engineering hry sice nebyl nemožný, ale zabral by výrazně více času a je zde otázkou, zda by se do této činnosti autoři překladu pouštěli.

Za velmi zajímavé můžeme považovat doznání pravděpodobného autora zachování tohoto souboru Garyho „Gaze“ Liddona v komentářích pod videem, že uniklo jeho tehdejší pozornosti jako vedoucího programátora. I tak ale faktem zůstává, že zdrojové soubory dnešních větších her unikají spíše jako celá data a v některých případech za ně může být požadováno výkupné, jako v případě získání interních kódů a dalších materiálů vývojáře CD PROJEKT RED [18], jehož data byla následně dražena na černém trhu.

Specifika sběru dat

My advice would be to write what is most personal and specific to your experience or your life. And your voice will emerge and because of its specificity, it will be universal.

David Schwimmer

V této sekci rozebereme různé specifické detaily, které mohou mít vliv na samotný výzkum. Spadají sem různé detaily, které se buď na studii neprojevily, ale v některých případech by ji mohly ovlivnit a uvádíme tyto detaily jako varování pro budoucí výzkum, případně naši studii ovlivnily, ale nespádají pod žádnou (či všechny) výzkumné skupiny, a tedy je patřičné je zmínit odděleně.

4.1 Rozdíl mezi západními a východními komunitami ve sběru dat a práci s nimi

Prvním faktem, který musíme zmínit je ten, že veškerý náš výzkum a do značné míry i předpoklady se týkají tzv. západní komunity hráčů. Ve skutečnosti ale existují minimálně dvě velké skupiny hráčů, a je možné, že existuje i další jemnější dělení (například americké versus evropské hraní či hraní japonské versus čínské). Pro zbytek této sekce ale budeme pracovat s označením těchto bloků jako západního, který bude pokrývat euroamerickou sféru a tzv. východního, který bude pokrývat sféru asijskou.

Toto rozdělení je důležité proto, že každý z těchto bloků pracuje s informacemi do jisté míry jinak. Zjednodušeně bychom do začátku mohli říct, že západní komunity pracují s informacemi kolem hry jako s kolektivním vědomím, které je potřeba doplňovat, zatímco východní komunity častěji pracují s informacemi kolem hry jako s mírou prestiže, která určuje postavení hráče v komunitě na základě toho, kolik informací má.

Je samozřejmě důležité před zbytkem tohoto bloku poznamenat, že se v něm – jak je již možná zřejmé – do menší či větší míry dopouštíme generalizace. Jistě existují západní komunity, které nebudou sdílet veškerá data a budou existovat východní komunity, které budou mít své kolektivní wiki a alternativy. Obecně jde ale o poznatek, který vychází jak pozorování, tak z jisté míry rozdílu v mentalitách západních a východních komunit.

Tento rozdíl bychom mohli shrnout pod opět generalizující poznatek, že komunita si vždy žádá to, čeho se jí nedostává – obdobně jako například s preferencemi některých jedinců. Tedy tam, kde západní národy a přeneseně komunity jsou postavené na právech a postavení jedince,

přes které tzv. „nejede vlak“ a někteří jedinci mohou považovat kolektivní vědomí za zajímavou alternativu, tam východní národy a komunity postavené na kolektivní snaze mohou vidět vydělení jedince za zajímavější.

Tím se konečně dostáváme ke konkrétní problematice. Pokud budeme nějakou dobu sledovat západní komunity okolo některých her, nejspíš si dříve či později všimneme, že znalost dat je pro hráče jednou z nejvyšších priorit. Jde to až do takových extrémů, že můžeme mít komunity, které mají řešit nějaké individuální problémy, ale nejsou toho schopny, protože veškeré problémy berou jako problém kolektivní. Tím ovšem vzniká jistá nevyváženost mezi tím, jak jsou problémy kladeny a jak jsou problémy řešeny, kdy na jedné straně máme zdroje omezené (vývojáře) a na druhé straně máme zdroje v extrémním případě neomezené (kde počet hráčů může škálovat výrazně rychleji než počet vývojářů.

V západním světě existují – pokud zanedbáme rezignaci na tento typ problémů – hlavní dva proudy, jak tuto problematiku řešit:

- Prvním způsobem řešení je akceptování toho, že hráč nevystupuje jako jedinec a drastické navýšení složitosti předkládaných problémů. Zde můžeme zmínit buď hry jako *Elite: Dangerous*, které používají problémy jako šifrované audio [19], velmi vágní nápovědy, a jiné problémy, kterým se musí věnovat vyšší počet hráčů, ideálně koordinovaně, aby byli schopni nalézt řešení. Často nejde ani o nalezení konkrétního řešení, ale o zúžení okruhu, v kterém se má v rámci herní procedurálně generované galaxie hledat řešení, do takové míry, že se následně dá „pročesat“ větším počtem hráčů. Nejnotoričtější hledanou záhadou v univerzu *Elite: Dangerous* je pak planeta označovaná jako *Raxxla*, která má údajně obsahovat portál do jiné galaxie a k níž fakticky neexistují logická vodítka.

Druhým stylem navýšení složitosti je zvýšení počtu kombinací, které může hráč prozkoumat, ideálně exponenciální řadou. Touto cestou se vydávají některé hry jako *Final Fantasy X* nebo *Path of Exile*, z novějších například *Assassin's Creed: Valhalla*, které implementují tzv. strom schopností (skill tree), kde se hráč může vydat různými směry v síti provázaných schopností a vytvořit si unikátní postavu. Do jisté míry jde o falešnou dichotomii, protože ve hře, která je postavená na postupu hrou musíme zvážit, že výrazně nevýhodné schopnosti nám zabrání v hraní a dohrání hry. I tak ale vzniká astronomicky vysoké množství kombinací, kterými se mohou hráči vydat. Ve výsledku je tedy pro hráčskou komunitu nemožné analyzovat všechny cesty a různé průchody jsou akceptovány jako optimální podle toho, že schopnosti a bonusy na sebe navazují, jak je vývojář navazoval za sebe při tvorbě stromu.

- Druhým způsobem řešení je naprostý protiklad řešení prvního, a to rezignování na problémy, které vyžadují hlubší analýzu a jejich přenastavení na problémy, které hráč má řešit jako jedinec. Opět se zde můžeme omezit na dvě ukázkové podskupiny.

První podskupina prohlašuje, že hráč by se měl soustředit na hraní samotné a ne na extrémní metagaming. Mezi hry tohoto typu můžeme zařadit například *Diablo 3* – *Diablo 2* obsahovalo relativně jednoduchý strom schopností, který ale hráče omezoval ve smyslu, že hráč měl jednu postavu, která realizovala průchod tímto stromem. Výsledkem byly mimo hru hráči vedené diskuse ohledně relativně komplikovaných „optimálních stromů schopností“ pro jednotlivé postavy ve hře. Vývojáři ale ve vývoji *Diabla 3* sice zanechali stromy schopností, ale umožnili jejich přepínání, kdy konkrétní postava může vyzkoušet jakoukoliv kombinaci schopností. Výsledkem je fakt, že jediný hráč si na šesti, případně sedmi s rozšířením, postavách může vyzkoušet jakoukoliv kombinaci, zda mu vyhovuje, což drasticky snižuje čas potřebný k testování a omezuje komunitní kolektivní řešení. Samozřejmě nadále existují sepsané ideální kombinace

pro různé typy hráčů a postav, ale výsledkem je zvýšená interaktivita hráčů, kdy hráč nemusí strávit delší čas zvyšováním úrovně postavy, aby si kombinaci vyzkoušel.

Druhá podskupina se zaměřuje spíše cestou náhodnosti a reakcí hráčů. Z většiny jde o takzvané „roguelites“ – hry vycházející z původní hry Rogue, které se jí více či méně blíží a staví na sadě pravidel žánru roguelike. Nemusí jít ale jen o hry tohoto žánru, spadají sem i některé tahové strategie jako UFO: Enemy Unknown (1994) a jeho následovníci XCOM: Enemy Unknown (2012) a XCOM 2 (2016). Principem těchto her je, že hráč má v danou chvíli jen omezenou znalost budoucího vývoje a omezenou nabídku zdrojů (zbraní, staveb, jednotek, atd.) a je nucen si vybrat z toho, co je dostupné. Ve výsledku pak, pokud hra není striktně naplánována, toto vede k tomu, že ke hře neexistuje komplexní analýza, pouze sada obecnějších doporučení a srovnání (kdy může například existovat výrazně výhodnější zbraň či stavba za jednotkovou cenu jako přehlédnutí vývojáře).

Protikladem k těmto západním problémům a jejich řešením je pak jistý styl her východních, kdy východní hráči sice používají kolektivní řešení, ale jak vývojáři, tak hráči samotní směřují k jisté znalostní prestiži, kdy existuje množství hráčů, kteří například dohráli všechny konce hry, ale existují specifické kousky informací – typicky distribuované omezeně geolokací –, které nemusí být v komunitě sdíleny a zvyšují prestiž hráčů, kteří se k nim dostali. A i v případě, že jsou tyto informace sdíleny v komunitě, stále může hrát značnou roli fakt, zda se k informaci hráč dostal na internetu nebo v reálném životě.

Jako příklady těchto východních her můžeme uvést například klasickou hru Pokémon v její karetní podobě, kdy existují karty distribuované na limitovaných akcích (z nichž některé se týkaly soutěžení pomocí digitální verze hry), které nešly získat jinak. Z čistě digitálních her pak můžeme uvést sérii her NieR, u které existují části příběhu, vysvětlení a neoficiální konce, které byly sdíleny jen omezenému množství hráčů na zpoplatněných živých setkáních v rámci Japonska [20]. Posledním příkladem, který zde můžeme uvést jsou živé události (eventy, opět zpoplatněné) v japonské komunitě Kantai Collection, kde se hráči dovídají informace o nových updatech a eventech ve hře s předstihem oproti hráčům na internetu či čtenářům herních časopisů. Jako konkrétní příklad můžeme uvést Zuiun Festival v japonském zábavním parku Fuji-Q, kdy se kromě standardního zábavního a doprovodného programu hráči měli šanci dozvědět, že jedna z herních postav obdrží ve hře vylepšení [21].

Bohužel tento rozdíl v přístupu západních a východních komunit k informacím není možné aplikovat důsledněji v rámci práci, kde jsme se omezili pouze na tento popis. Bylo by jistě zajímavé rozebrat tyto rozdíly, jejich vznik a dopad na přístup k informacím (kam by například mohl spadnout i dopad zavedení GDPR v rámci EU), ale hlubší výzkum předpokládá dobrou znalost asijských jazyků, ideálně alespoň japonštiny a čínštiny, na komunikační a lepší úrovni. Autoři této práce bohužel touto znalostí nedisponují a omezují se tedy v této části primárně na zprostředkované a obecné informace.

4.2 Problematika odhadu informací a nahrazení herních vývojářů fanoušky

Jak bylo částečně naznačeno v části věnující se herním API, v posledních letech narůstá problematika objemu hry jak z hlediska vývojářů, tak z hlediska hráčů. Značné množství hráčů není schopno abstrahovat a pochopit náročnost vývoje, hlavně komplikovaných her s provázanými systémy. Následně dochází ke konfliktům mezi jednotlivými hráči a v extrémním případě i vývojáři, kdy hráči nejsou schopni a ochotni akceptovat vysvětlení.

Tento problém je nejvíce zastoupen ve hrách středního a velkého objemu, protože u malých indie (nezávislých) her jsou hráči ještě relativně schopni pochopit objem tvorby vztažený na čas. Jak ale roste objem dat a samotné hry, je toto pochopení náročnější.

Co dále ztěžuje toto pochopení je jistý konflikt obsahu a mechanik. Existují hry obsahové, jako například *Grand Theft Auto V*, kde si hráč za konkrétním prvkem hry může představit konkrétní dobu, kterou na ní vývojář strávil. U těchto her bývá jediným problémem odhad množství prvků ve hře, protože hráč jich vidí jen jistou část. Druhým extrémem jsou hry mechanické, jako například *Minecraft* – zde vývojář prakticky netvoří žádné konkrétní prvky (vyjma procedurálně umístěných bloků, které tvoří zanedbatelný čas vývoje) a je sice náročnější odhadnout čas k vytvoření mechanik jako je procedurální generování, ale ve výsledku nějaké odhady existují a dají se aplikovat.

Nejtěžší pozici pak mají hry kombinující oba postupy, jako například *No Man's Sky* či *Elite: Dangerous*. Pro hráče je jednak složitější odhadnout čas vývoje mechanik hry, a zároveň nemusí kvůli rozloze hry nikdy spatřit propracované prvky (assety), na kterých tým vývojářů spálil mnoho času. To v extrémních případech vede k nespokojenosti hráčů mnohdy vyjádřené jednoduchým výkřikem: „Čemu se ti vývojáři poslední měsíc / rok / dekádu věnovali, vždyť v tom updatu nic není!“

Aby vývojáři neměli jednoduchý život, je tento problém navýšen o problém rychlejšího konzumování obsahu. Hry s fixním obsahem jsou sice propracované, ale zvláště v poslední dekádě (kvůli dostupnosti zvýšené grafické kvality) je pro vývojáře velmi náročné generovat obsah na takové úrovni a v takovém objemu, aby uspokojil veškerou herní komunitu. Zde si můžeme mírně vypomoci a vzít jako ukázkou některou hru typu GaaS, například *Genshin Impact*. Tento má fixní dobu mezi updaty, šest týdnů, což je pro naši analýzu ideální. Některé updaty jsou menší a některé větší, navíc v rámci velkého updatu jsou distribuována data pro menší eventy, které jsou rozprostřeny mezi data updatů, aby byl hráč stále lákán zpět. I tak lze ale odhadnout, že těchto šest týdnů vývoje poskytne hráčům maximálně jednotky dnů (nižší desítky hodin) obsahu, než je tento plně zkonsumován / odehrán.

Není pak překvapením, že je množství her, které rezignují na komplexní obsah a zaměřují se spíše na jednoduchou interakci s živými protihráči, jelikož tato zvyšuje znovuhratelnost a nepředvídatelnost hry a ve výsledku i její zajímavost. Z posledních let můžeme z této kategorie zmínit hitové hry *Fall Guys: Ultimate Knockout* (2020) a *Among Us* (vydána 2018, zpopularizována roku 2020) a předtím multiplayerové hry typu FPS jako *Team Fortress 2* (2007) nebo různé verze hry *Counter Strike*.

Poslední problematikou v této kategorii ztěžující hráčům pochopení informací o hře a vývojářích jsou přehnané výroky vývojářů. Zde se omezíme na konkrétní citace, z kterých je patřičně pochopitelné, že zvláště hráčům, kteří se nevyznají v reáliích vývoje mohou tyto výroky fakticky znemožnit odhad toho, co je ve vývoji her reálné a co ne:

- *The team programmed some of the physics for aesthetic reasons. For instance, Duncan insisted on permitting moons to orbit closer to their planets than Newtonian physics would allow. When he desired the possibility of green skies, the team had to redesign the periodic table to create atmospheric particles that would diffract light at just the right wavelength. [22]*
- *„I could take this coffee cup, I could get into my spaceship, I could fly to another planet, wonder out into a forest, I could drop my coffee cup in the forest, I could leave and come back a month later. And if someone hasn't taken the coffee cup, the coffee cup would still be there.“ [23]*

Všechny tyto problémy postupem času vedou k vyšší popularitě řešení, kdy vývojáři převádí část vývoje na hráče, čímž odpadá značná část těchto problémů. Nejmarkantnějším příkladem

může být v části API zmiňovaný Elite: Dangerous, který má větší množství externích řešení založených na vypuštěném API. Někteří hráči a vývojáři by ale mohli argumentovat například i tituly od vývojáře Bethesda, kde některé hry mají postupem času více obsahu v hráči vydaných modifikacích než v originální hře, typicky jde o některé hry série Fallout a The Elder Scrolls. Bude tedy jistě zajímavé a pro budoucí vývoj důležité sledovat vývoj na tomto poli v budoucnosti, protože může výrazně změnit dynamiku uchovávaných dat v herních komunitách a vnímání hráčů ohledně toho, která data jsou důležitá, a která ne.

4.3 Výzkum v herních komunitách a možnost zneužití dat

Posledním problémem, na který jsme při realizaci práce narazili, byla jistá neochota některých komunit podstoupit výzkum, vedoucí k prodlevám a realizaci mimo vytipované komunity.

U některých komunit si nemůžeme být stoprocentně jisti, jelikož nedošlo ani k opětování komunikace administrací (například Super Mario komunita), ale u některých komunit jsme odpověď dostali a byla jednoznačně ve smyslu neakceptování dotazování se ohledně komunity. Citujeme odpověď ze soukromé konverzace Doom komunity, 6. dubna 2021:

Hello there. I am the Administrator of the Official DOOM Discord Community. Due to Privacy Reasons we cannot share any data related to usage statistics or similar metrics. Thank you for understanding.

I přes trvajícím ujišťování, že nám nejde o data o komunitě, ale o data využívaná komunitou byla naše komunikace ukončena.

Tato neochota je pravděpodobně vázána na fakt, že Discord Doom komunita je vázána na oficiální kanály jako vývojáře hry, vydavatele a další oficiální společnosti pohybující se kolem hry. Ve světle bezpečnostních skandálů napříč IT světem v minulosti (uniklá hesla či účty) a zpřísňující se administrativy (směrnice GDPR Evropské Unie) je pak bohužel pochopitelné, že některé komunity jsou velmi opatrné ohledně jakékoliv komunikace, která se ptá na data uživatelů nebo jejich skladování.

Podobný problém jsme zaznamenali při kontaktování Hollow Knight komunity, což je fakt, který nás velmi mrzel, protože tato komunita je velmi aktivní a sdružuje Discord, fandom wikipedii a aktivní speedrun komunitu. Odpověď administrace ale byla ve velmi omezeném smyslu (cituji ze soukromé konverzace s uživatelem PowderedWater, 16. dubna 2021: *i do believe i understand your idea, but from our end we really just search up what we need to know. the most extensive data collection would be how our bots log things like warns/mutes/bans/joins/leaves and those are private on purpose.*) a přes další upřesnění komunikace zanikla a nepovedlo se nám ji navázat nebo najít vhodnější komunikační kanál či osoby.

Ve výsledku tyto problémy ovlivnily bakalářskou práci jen do menší míry, ale je to jistě problém, na který je potřeba dávat pozor do budoucích výzkumů, protože bezpečnostní prvky a jistá zdrženlivost administrací herních komunit se budou spíše zvyšovat než naopak.

Doporučení přístupu k problematice pro jednotlivé skupiny

Getting information off the Internet is like taking a drink from a fire hydrant. [24]

Mitch Kapor

5.1 Komunity řešící průtoková data

V této části se bohužel moc navrhnout nedá - konkurovat v současnosti Redditu a Discordu by byl oříšek i pro větší týmy, natožpak pro jedince.

Navrhovaný přístup je proto celkem jednoduchý - ať již jste vývojář či správce komunity, je pravděpodobně ve vašem zájmu, abyste si vytvořili tyto kanály, protože jinak si je komunita založí dříve či později sama a pak může být obtížnější se propracovat ke správě této služby – zjednodušeně se dá říct, že nasdílet nebo předat řízení služby se dá vždy, přebrat řízení služby je nepoměrně těžší. Na Discordu existuje i možnost nechat si ověřit, že jste oficiální vývojář, které můžete využít.

Nepodceňovali bychom toto založení jen kvůli velikosti hry, pokud nejde například o vyložení za víkend vyvinutou hru typu těch, které vznikají na festivalu Ludum Dare, apod. Vždy se může stát, že vaše hra zaujme hráče a raketově vylétne nahoru a pak již bude pozdě tyto záležitosti řešit. Pokud čekáte, že globálně vaše hra zaujme stovky až tisíce hráčů, mít tyto kanály ošetřené je ve vašem zájmu.

5.2 Komunity s datovým řešením postaveným na kolektivu

Tyto komunity jsou asi nejnáročnější na rozhodnutí. Z hlediska technické práce musíme spíše doporučit vlastní instalaci některého z wikipedia řešení. Pro rozšířenost PHP na serverových instalacích a robustnost těchto řešení bychom asi doporučili MediaWiki či DokuWiki (mj. také

protože jejich syntax je de facto standardem a může spíše vyhovovat uživatelům), ale na trhu se nachází značné množství dalších alternativ, které lze zvážit.

Pouze v případě, že si vývojář či správce komunity opravdu není jist technickou zručností či se nedostává času a dalších kapacit, bychom doporučili založení stránky na službě fandom. Je dost pravděpodobné, že tato stejně vznikne, ale oficiální wikipedie vývojáře mívají často výhodu prvních a oficiálních dat, která je činí v očích veřejnosti zajímavějšími.

Co se instalace CMS či jiné novinkové platformy týká, lze ji doporučit pouze v případě, že fanoušek či vývojář chce informovat o hře dlouhodoběji a pro krátkodobé informace využít průtokové služby prvního typu. V tom případě můžeme doporučit vlastní instalaci služby Wordpress, ideálně však s patřičnými bezpečnostními pluginy (tj. moduly), jelikož slabiny této instalace jsou díky jejímu rozšíření široce známy. Alternativou může být spuštění blogu na službě wordpress.com, kde ale autor může velmi rychle narazit na omezení nastavení vzhledu a jiných funkcí.

5.3 Komunity analyzující herní data ve spolupráci s vývojáři

Ve výsledku se tato část liší případ od případu a nejde perfektně doporučit konkrétní technologie a postupy, spíše navrhneme obecné a komunikační rady.

Pro vývojáře zde musíme doporučit publikaci herního API, pokud to hra jen trochu logicky umožňuje. Samozřejmě je nutná patřičná sanitace výstupů, ale odměnou je vstup hráčů do hry jako takové, který není srovnatelný s žádným jiným způsobem zde uvedeným. Nejvyšším přínosem zde může být fakt, že vývojář jen s malým zpožděním vidí, po kterých informacích má komunita hlad a může je – pokud na to má kapacity – patřičně vylepšit ve hře.

Důležité je zde zmínit, že málokdy bývá zpětné vylepšení ku neprospěchu věci, jelikož hru hraje sto procent hráčů, zatímco externí aplikaci využívající API sto procent hráčů nikdy nevyužije. Zároveň autoři API často berou vylepšení jako službu sobě a komunitě a i v okamžiku, kdy je hra vylepšena a jejich aplikace API přestane být k užítku, bývají schopni najít novou oblast, ve které mohou vytvořit novou informační hodnotu pro komunitu.

Pro vývojáře v této části doporučíme primárně komunikaci s vývojářem, v druhé řadě pak důkladné nastudování poskytnutého API. Je slušností vývojáře upozornit na bezpečnostní chyby, na druhou stranu slušný vývojář by měl uživatele API předem upozornit na jeho aktualizace a poskytnout nějaký čas na seznámení se s novou verzí před jejím vypuštěním do ostrého provozu.

5.4 Komunity analyzující herní data bez spolupráce vývojářů

Reverzní inženýrství je vysoce specializované téma, a tím pádem bohužel v této části také nemůžeme nabídnout více než obecné poznatky. Kde je to možné, je žádoucí komunikovat s vývojářem, i za cenu sníženého komfortu. Tam, kde to možné není je vhodné se připravit na celkem náročnou část vývoje.

V této části je pak vcelku jasné, že jakákoliv snaha bude výrazně usnadněna navýšením vývojářů v týmu, kdy se snaha o vývoj může pohybovat mezi velmi těžkou až plně nemožnou. Doporučujeme tedy v případě, že se vývojář do této práce chce pustit, aby se pokusil v komunitě najít další vývojáře, které tento směr může zajímat a koordinovat svůj postup s nimi.

Určitě je také v těchto komunitách dbát více než jindy na zálohování a s ním spojené procesy, jelikož ztráta dat zde může být výrazně větším problémem, co se časové náročnosti týká.

Kapitola 6

Závěr

Po konci hry se král a pěšec potkají ve stejné krabici.

Italské přísloví

6.1 O práci

V této teoretické práci jsme se věnovali rozboru herních komunit a jejich zacházení s daty, sbíranými členy komunity a dále využívanými členy těchto společenství. Provedená práce se dá rozdělit do několika primárních částí:

- Rozbor teoretického zázemí, kde jsme si rozebrali historii, která vedla k dnešnímu chápání herních komunit a tomu, jak tyto komunity pracují s herními informacemi.
- Případové studie, kdy jsme cíleně kontaktovali správce jednotlivých komunit s dotazníkem ohledně dat, která spravují pro komunitu a detailů těchto řešení.
- Analýza získaných dat, kde jsme probrali získané případové studie a snažili si ověřit, zda odpovídají naším předpokladům z teoretické části práce.
- Doporučení jednotlivým skupinám, kdy jsme vybrali některá doporučení pro vývojáře jak řešení pro komunity, tak samotných her, a snažili se je klasifikovat a přehledně vypsát, aby byla přínosem pro čtenáře této práce.

6.2 Úroveň realizace zadání

- **Seznamte se se způsoby sběru dat v rámci herních komunit.**
K seznámení se s jednotlivými typy sběr došlo v patřičné míře a máme za to, že jsme pokryli všechny typické případy sběru a využití dat.
- **Navrhněte vhodnou kategorizaci komunit a nalezněte vhodné zástupce pro dané kategorie. Snažte se minimalizovat počet základních kategorií (maximálně jednotky), hlavním kritériem při rozdělování bude způsob sběru dat dle předchozího bodu.**
Jednotlivé kategorie jsme navrhli dle zadání, konkrétně v počtu čtyř hlavních typů. Zde musíme

zmínit náročnost takového dělení, kdy se do značné míry jednotlivé typy sběru dat v komunitách překrývají. Na druhou stranu jsou tyto kategorie celkem jasně odlišitelné používaným softwarem či užitými postupy.

- **Proveďte stručnou rešerši software využívaného v rámci herních komunit.**
Rešerše byla provedena a její výstup byl popsán primárně v kapitole Rozdělení sběru dat. V případových studiích jsme si pak potvrdili užití převážně moderních typů těchto řešení.
- **Navrhněte tzv. „best practices“ pro jednotlivé základní kategorie, poukažte na případné problémy při sběru dat a doporučte konkrétní řešení.**
Navrhli jsme patřičná řešení pro jednotlivé kategorie dat a různé typy vývojářů v kapitole Doporučení přístupu k problematice pro jednotlivé skupiny. Některé kategorie obsahovaly doporučení konkrétnější, některé obsahovaly spíše obecnější instrukce.
- **Veškeré návrhy zpracujte v souladu s běžnými postupy a metodami SW inženýrství.**
Hlavně v teoretické části – ale i jinde – jsme na tento bod dbali a snažili se používat patřičné metody softwarového inženýrství a analýzy.
- **Výsledkem práce bude, kromě pečlivě provedené rešerše, sada návrhů na vhodné zpracování dat v jednotlivých typech herních komunit.**
Požadovaná sada návrhů je samostatnou, již zmíněnou kapitolou, která by měla vyhovovat jako výstup práce.

6.3 Výsledky práce

Hlavním výsledkem práce je sada doporučení jednotlivým typům vývojářů dle požadovaného rozsahu věnování se komunitě. Do jisté míry je zklamáním, že tento výsledek je relativně obecný, na druhou stranu se nám ale v rámci práce podařilo specifikovat obecnější místa a místa, která budou potřebovat další konkrétnější výzkum, což považujeme za jednoznačný přínos práce do budoucna.

V rámci práce jsme se také snažili shrnout teoretické poznatky z průběhu průzkumu, které považujeme za další možné (byť volnější) nasměrování do budoucích prací, případně které mohou pomoci v dalším výzkumu.

6.4 Budoucí směřování návazných prací

Máme za to, že v rámci práce se podařilo obsáhnout hlavní směry sběru dat a jejich zpracování v herních komunitách. Co se dalšího výzkumu týká, po analýze případových studií máme za to, že většina budoucích směrů je spíše vyčerpána, jelikož jsou to buď směry obecné a není příliš reálná realizace konkurenceschopného řešení (pro průtoková data), případně jde o směry naopak vysoce konkrétní (API a reverzní inženýrství), kde konkrétní řešení k té které hře či komunitě může vydat na samostatnou práci.

Ve výsledku je dle našeho názoru možné pokračovat ve výzkumu primárně na poli konkrétních řešení stálých komunitních dat, kde by bylo možné více do hloubky prozkoumat různé verze různých instalací wikipedií a případně navrhnout, či realizovat řešení, které by vylepšilo práci herním komunitám, například integrací do služeb průtokového typu jako jsou Reddit či Discord.

Dotazníky případových studií

A.1 Prázdný dotazník

- The questions are mostly fairly open and it's up to you how you want to answer them. I appreciate more information and try to provide some examples, but ultimately it's up to you how you answer and how much you answer.
- If you think you cannot answer any question (don't know how, feel that it's a security issue, etc.) just write it down, there is no problem with that. More information on why you cannot answer is welcome, but I understand that in some cases even this information can be security issue, so answers like „I cannot answer this“ are fine, too.
- You can also answer the question for my understanding only and specify you don't want it published. In such case I recommend sending the answers to me via more secure channel (like password protected archive and password via separate channel, etc.) and I won't put it in the thesis.
- The questions are separated into few blocks, first a general one for everyone, then I split them into blocks based on my preliminary view on how the data is collected from players. This is mostly guidance, feel free to answer as much as you know, but it might help you skip questions you don't know answers to or that are not related to your case.
- That said, if you are in a community that handles multiple types of input and you know someone who might be willing to provide more information that you cannot (for example you manage community wikipedia and you know the administrator of the same community's subreddit and discord), I'd appreciate if you could pass these questions to them.
- If you have any questions, feel free to contact me on Discord or mail, I am usually online most of the day because of my job and can reply fairly swiftly.
- Send the answers to me via any reliable channel. E-mail or pastebin are great, even Discord is okay via its text file option.

That said, here are the questions, thank you for your time:

A.1.1 General part

- Briefly introduce yourself (name or nickname) and what gaming community you (help) manage

- Write down which data types (from the four groups further below) you manage
(If you can think of any type of data that players use to get information on the game and it is missing, feel free to mention it.
I intentionally didn't cover things like YouTube/Twitch here, because they aren't managed by one person and are individual and independent efforts.)

A.1.2 General part regarding data

- Which solution do you use to keep this type of data?
(I'll keep referring to this as a „your solution“ only from now on)
- What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)
(More information is better, but I understand this question in particular can be security risk in some cases.)
- If you picked this solution yourself, why did you pick it over the competition (Discord over IRC, own wiki over Fandom wiki, etc.)? If someone else picked the solution and not you, why do you think they picked it?
- Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?
- What would make you switch to another solution (aside of the current one becoming unusable for any reason)?
- What could be most improved on your solution (in case it's more things pick two or three most desired)?
- Is there some feature you'd like to see that the competition of your solution has and your solution not?
- If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?
- If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?
- Why do you think players use your solution to search for data to begin with?

A.1.3 Fast-flowing data (IRC, Discord, Slack, Reddit, etc.)

- How much do you think this type of service is used by players in your community to get some data they need?
- How good is your solution of this type with searching data, is it easy for your players to find the data they need?
- How much do you or your moderators need to oversee the data that flows around the community?
- Have you ever met intentional spreading of fake data (misinforming new players, fake leaks, etc.) on this type of solution? If yes, how well could it be solved with your solution's tools?

A.1.4 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)

- How much do you think this type of service is used by players in your community to get some data they need?
- What do you feel is the most searched thing regarding the game and why?
- This type of solution is usually aimed at searching to begin with, but do you think the way players get to the data they want can be improved?
- How much do you or your moderators need to oversee the data that flows in?
- How much do you or your moderators need to oversee the editing of the data that is already in?
- Have you ever met intentional vandalism on this type of solution? If yes, how well could it be solved with your solution's tools?

A.1.5 Data from developers (game API, cooperation, etc.)

- How much do you think this type of service is used by players in your community to get some data they need?
- What type of data does your solution provide and why do you think it's not in the game or - in case it is in the game - it is duplicated outside of the game and possibly more searched there?
- How easy it is to get the data you want and/or need from the developers?
- Is there something you'd like to see improved in getting the data to the players?
- Is there some type of feedback loop, both from you and players using this solution, towards the developers? How much, in your opinion, do you think it helps the game?
- Is there some common misconception you see players make, even though reliable data from developers are available and provided?
If yes, why do you think these misconceptions exist in your community? Is it the fault of the developers, of the solution, of the community, someone/something else?

A.1.6 Data mining (data mining, reverse engineering, modding, etc.)

- How much do you think this type of service is used by players (or possibly a subset like speedrunners) in your community to get some data they need?
- How easy it is to get the data you want from the game?
- Since it's usually harder to get this type of data than the others, do you feel that it tends to be kept more secret or do you feel that people share the data as soon and openly as possible?
- Have you ever run into issues with the developer not wanting to have data from their game available publicly? If yes, how was the situation solved?
- Specifically for very hard-to-perform communities (speedrunning, etc.), how big do you estimate is the retention of the players if the techniques are kept public? That is, if one hundred people read „speedrunning wiki and tips“, how many of them will stick around and actually try speedrunning or at least actively join the community?

A.1.7 Etc.

- Anything that you feel might help the reader understand how/why/when/etc. data flows from the game/developers/players to you and back to players that you feel might have been missing in the previous questions

A.2 Pokémon komunita

I am u/Haruka-sama on Reddit Josh#6734 on Discord, I moderate the r/Pokemon subreddit and discord server.

The subreddit / discord usually manages Fast-Flowing data, (community submissions).

I however have attempted to provide some insights into other types of data collected by the Pokemon community at large where I can, I can't call myself an expert but I hope it can be found to be informative.

Some notes: Throughout this I'll be using some acronyms you should be aware of, these are commonly used in the Pokemon community:

Acronym	Meaning
GF	Game Freak
TPCI	The Pokemon Company International
/vp/	The Pokemon board on 4chan

Useful links: Here are some links I'll reference throughout this.

Site	URL
r/Pokemon	https://reddit.com/r/pokemon
r/Pokemon Discord	https://discord.gg/pokemon
Bulbapedia	https://bulbapedia.bulbagarden.net/wiki/
Pokemon Database	https://pokemondb.net/
Serebii	https://serebii.net
Smogon	https://www.smogon.com/
/vp/	https://boards.4chan.org/vp/
Official Pokedex	https://www.pokemon.com/pokedex/

A.2.1 With regards to data

Which solution do you use to keep this type of data?

We use Reddit and Discord to keep our data, as well as some moderation specific information stored on a Virtual Private Server (Cloud Server).

What technical background is behind your solution?

Reddit and Discord are cloud run services, as far as I know Reddit is hosted on AWS (Amazon Web Services) and Discord GCP (Google Cloud Platform)

Moderation related data is hosted on a Hetzner VPS with a self-managed PostgreSQL database.

If you picked this solution yourself, why did you pick it over the competition? If someone else picked the solution and not you, why do you think they picked it?

I personally did not choose to create the Reddit/Discord communities myself.

Reddit was chosen as it was clear that it would surpass Bulliten Board style sites in general, and was a powerful free platform.

Discord is a similar thing, originally the discord was created in tandem with an IRC chat r/Pokemon managed, however as the Discord server continued to grow in size the IRC chat was eventually phased out.

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

As far as I am aware no, it's difficult to with communities in the 6 or 7 digit member counts relocate, and Reddit and Discord are already sufficient for our needs.

What would make you switch to another solution?

I personally cannot provide an answer to this.

What could be most improved on your solution

On the reddit side, there are few improvements to be made, the system is fairly robust as is.

Discord is lacking in the moderation front, with the majority of holes patched by automated processes such as bots developed and maintained by the community.

Is there some feature you'd like to see that the competition of your solution has and your solution not?

Slack has threads (essentially channels derived from individual messages or replies on Reddit)

Discord is currently lacking this feature, however this is expected to be released in the next few months.

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts.

Reddit has mirror sites which replicate data, data posted on r/Pokemon is backed-up in the moderation database, however if data is lost due to how discord and reddit work it would be impossible to put the data back in its original place, instead sites would need to be developed to host the information, and our platforms would need to link to those sites where needed.

How long do you think it would take to deploy and how much would it affect the community?

A usable site could be developed and deployed in less than a day, with improvements made over weeks and months subsequent.

Why do you think players use your solution to search for data to begin with?

Reddit and Discord inherently have a sense of community, I think that is a powerful draw in of itself to our platform.

A.2.2 Fast-Flowing data

r/Pokemon and its discord server very commonly hold discussion channels for newly released or upcoming games as a way to quickly propagate new information:

You can find an example of that here (Search for messages from late 2019, type in: #pokemon-sush before: 2019-12-01)

How much do you think this type of service is used by players in your community to get some data they need?

Prior to and at release time it's likely the majority of players use services such as Reddit and Discord to retrieve information, as wikis and other database sites usually take a few weeks to fully provide the information.

In case the above is not zero, what do you feel is the most searched thing regarding the game and why?

Usually information about what pokemon are obtainable, and new characters is obtained at this stage, rather than specific niche data which is stored on wikis and database sites.

How good is your solution of this type with searching data, is it easy for your players to find the data they need?

Reddit provides a search feature, however unless a post has been pinned by a moderator it's not likely to be seen by a player quickly searching for information.

To that end, it's likely that when a game is announced or released a „Megathread“ will be created to allow the community to share data efficiently.

On the discord end, usually a channel will be created for discussion and useful posts will be pinned by moderators to allow other community members to easily find them.

How much do you or your moderators need to oversee the data that flows around the community?

Moderation of fast flowing data is key, especially to remove fake or misinformed data. At release time it's not unlikely that (especially on reddit) the majority of community submissions will be removed or marked as misinformation by moderation staff.

Have you ever met intentional spreading of fake data (misinforming new players, fake leaks, etc.) on this type of solution? If yes, how well could it be solved with your solution's tools?

Reddit and Discord both support deletion of user submitted data, On Reddit's side it's also possible to mark a post as invalid or pin a comment to such effect.

A.2.3 Collective permanent solution

The r/Pokemon subreddit directly doesn't hold a permanent solution like this, instead we would link to Bulbapedia or Serebii. I can provide some insight into how those platforms are used however.

How much do you think this type of service is used by players in your community to get some data they need?

I would gather than for games after release the majority of information would be gathered from these services.

What do you feel is the most searched thing regarding the game and why?

I would, from personal experience assume information regarding Pokemon evolution requirement, stats and movesets are the most searched throughout the community.

This type of solution is usually aimed at searching to begin with, but do you think the way players get to the data they want can be improved?

Yes, I think this can be shown well enough by example, searching something like „Pikachu moveset“ into Google will provide results pretty much exclusively made up of the list of sites I linked above.

How much do you or your moderators need to oversee the data that flows in?

This rarely happens in our situation, as usually only during release time will users add their own information.

How much do you or your moderators need to oversee the editing of the data that is already in?

Usually, if information is proven to be false moderators will remove it or just mark the information as inaccurate, rather than personally editing it.

Have you ever met intentional vandalism on this type of solution? If yes, how well could it be solved with your solution's tools?

I personally have not.

A.2.4 Data from developers

Data from developers directly isn't often used in the Pokemon community at large, but here is some insight.

How much do you think this type of service is used by players in your community to get some data they need?

This usually depends on what information the player is looking for, for information on upcoming releases far in the future the official site or CoroCoro magazine pages are about as much information as one will find.

After a game is released however I would gather most players avoid using official information from GF/TPCI just due to how comparatively sparse it becomes. Take this for example:

Bulbasaur's Official pokedex page

Bulbasaur's Bulbaapedia page

You can clearly see the fan run site bulbapedia holds maybe 100 times more information on it than the official site hosted by TPCI.

It's this lack of information from GF/TPCI which has sparked the creation of communities in of themselves.

What type of data does your solution provide and why do you think it's not in the game or - in case it is in the game - it is duplicated outside of the game and possibly more searched there?

r/Pokemon's data solutions are in most cases fairly niche use case, the data is available in-game via trial and error or data-mining but storing and providing the information publicly is obviously preferable for most players.

How easy it is to get the data you want and/or need from the developers?

It's very uncommon for data used in our solutions to have come from the developers.

Is there something you'd like to see improved in getting the data to the players?

Through necessity the community has grown to not need information from the developers anymore, whether that's a good thing or not I couldn't say definitively. However I would not say this means I wouldn't like to see GF try and provide more comprehensive resources for communities.

Is there some type of feedback loop, both from you and players using this solution, towards the developers? How much, in your opinion, do you think it helps the game?

I think solutions created by not just r/Pokemon help the community immensely. Personally I would not have developed such an interest in Pokemon or it's Competitive scene without tools such as Bulbapedia or Serebii existing.

Is there some common misconception you see players make, even though reliable data from developers are available and provided?

In Pokemon's case, no not really, GF/TPCI provide limited information in the first place unfortunately.

A.2.5 Data mining

Data-mining isn't usually done by r/Pokemon directly however data collected from this will usually become fast-flowing data which passes through r/Pokemon and its discord server.

How much do you think this type of service is used by players (or possibly a subset like speedrunners) in your community to get some data they need?

The majority of information on pokemon database sites i.e. Bulbapedia, Serebii, Marriland, Veekun etc. will be derived from data-mining pokemon ROMS.

How easy it is to get the data you want from the game?

I'm not an expert in this myself, however given how quickly information is leaked out usually I would hazard a guess that there pose little challenges.

Since it's usually harder to get this type of data than the others, do you feel that it tends to be kept more secret or do you feel that people share the data as soon and openly as possible?

In Pokemon's case, not really usually as soon as someone gets a copy of the ROM datamined information will appear on /vp/ and begin being propagated to various other sites including the subreddit community.

Have you ever run into issues with the developer not wanting to have data from their game available publicly? If yes, how was the situation solved?

Nintendo, TPCI and GF are pretty well known for wanting to keep information out of the public, at least until product release date, good example of this was leaks of Pokemon: Sword & Shield which occurred on various discord servers. [See Article 1]

Usually as we obtain the information through these individuals by proxy, we're unaffected by this.

Specifically for very hard-to-perform communities (speedrunning, etc.), how big do you estimate is the retention of the players if the techniques are kept public? That is, if one hundred people read „speedrunning wiki and tips“, how many of

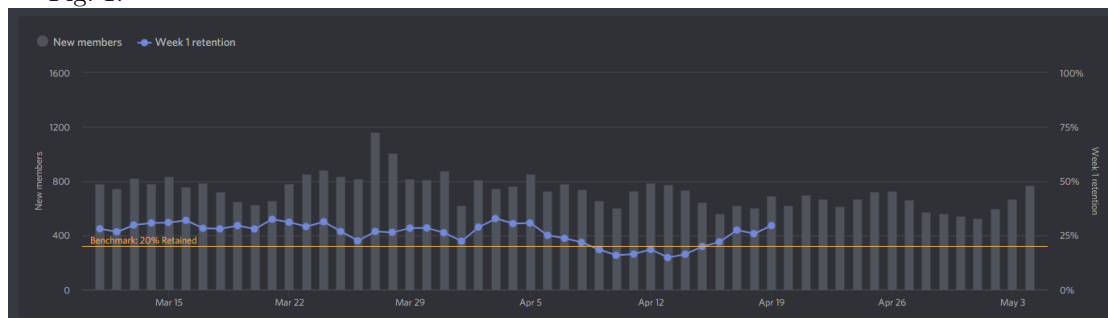
them will stick around and actually try speedrunning or at least actively join the community?

The general Pokemon community tries to be as inclusive as possible so there are very few bars to entry, So this question does not particularly apply.

n.b. According to discord's own data, the r/Pokemon subreddit discord retains 30% of members who join after 7 days. (500 members per week). [See Fig 1]

Article 1:
IGN: Pokemon Company Given Legal Permission to Pursue Sword and Shield Leakers
<https://www.ign.com/articles/2019/12/14/pokemon-company-lawsuit-subpoena-sword-shield-leak-leakers-discord-4chan>

Fig. 1:



A.3 Doom speedrunning komunita

General:

I'm kraflab, I run the „hell on earth“ discord devoted to classic doom speedrunning (the original from the 90s, on PC). I also built and host the current doom speed demo archive, which tracks the history of doom speedrunning, which is recorded on demo files.

Discord:

This is used for communication, but doesn't offer much in terms of searching. There are some basic posts that are stickied that provide lists of links to more information, but it isn't really a good place for permanent information storage, so it isn't used for this purpose. We don't do any „management“ in the sense of storing and tracking things - the discord doesn't even have basic things like bots. It's more or less just a platform for discussion.

Permanent Solutions:

There are wikis maintained by different subcommunities that have general information on the game, which also includes things relevant for speedrunning, and various forums as well. I'm not involved in the moderation or operation of any of those, so I can't comment on that side of things. There isn't a speedrunning-focused wiki or anything like that, so access to information via this kind of resource is limited in a lot of ways. The demo archive tracks all the runs and has some basic information in it, but it's primary purpose is preserving history in the runs themselves. It doesn't sound like that's the type of data you're asking about, so I'll skip it.

Game Data:

This game has been open source for decades and has one of the most active modding communities for any game, historically. As such, most of the questions here aren't really applicable. All the game data is readily available, from basic delineation of what is there all the way to modern editing tools fully exposing all the content. This data is available on the wikis but any serious runner is familiar with the editing tools and will open up the game data directly if they want to check some specifics.

A.4 Elite: Dangerous API komunita

Briefly introduce yourself (name or nickname) and what gaming community you (help) manage

Dlouhá léta používám nickname Artie a jsem jediným autorem jedné z největších website zaměřené na hru Elite: Dangerous - Inara (<https://inara.cz>). Další netřeba rozvádět, výsledná práce doufám hovoří za vše. :)

Write down which data types (from the four groups further below) you manage (If you can think of any type of data that players use to get information on the game and it is missing, feel free to mention it.

I intentionally didn't cover things like YouTube/Twitch here, because they aren't managed by one person and are individual and independent efforts.)

Website, API, herní databáze.

A.4.1 General part regarding data

Which solution do you use to keep this type of data?

Website, webová služba.

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

Klasické řešení v podobě Linux s Apache, PHP a MariaDB (bez uvedení verzí) na dedikovaném serveru. Nepoužít žádný framework, CMS, atp. vyjma ZeroMQ pro sdílení dat.

If you picked this solution yourself, why did you pick it over the competition (Discord over IRC, own wiki over Fandom wiki, etc.)? If someone else picked the solution and not you, why do you think they picked it?

Vzhledem k nutnosti hodně specifických řešení, byla to nejlepší volba.

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

Ne, není žádný konkrétní důvod měnit, řešení je plně vyhovující.

What would make you switch to another solution (aside of the current one becoming unusable for any reason)?

Nic, vzhledem k rozsahu projektu jiné řešení ani nepřipadá příliš v úvahu.

What could be most improved on your solution (in case it's more things pick two or three most desired)?

Bez odpovědi.

Is there some feature you'd like to see that the competition of your solution has and your solution not?

Bez odpovědi.

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?

Ano, vše je zálohováno, takže řešení je obnovení ze záloh.

If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?

Zaleží na rozsahu problému, od pár minut až několik hodin, s dopadem na komunitu mizivým nebo omezeným (menší rollback), opět v závislosti na rozsahu problému.

Why do you think players use your solution to search for data to begin with?

Obvykle pohodlnější, rychlejší a ucelenější přístup k informacím než přímo ve hře, plus přidaná hodnota v podobě dalších výstupů (vyhledávání, přehledy, statistiky, atp.) a uživatelského obsahu (galerie, logbooks, diskuse, atp.).

A.4.2 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

Podle velmi konzervativního odhadu (s ohledem na sporá oficiální data o počtu aktivních hráčů) je Inara pravidelně používána minimálně 20% hráčské základny. Realisticky spíše více.

What do you feel is the most searched thing regarding the game and why?

Data o herních obchodovatelných komoditách, vybavení lodí a jejich modifikace (engineering), informace o stanicích a podobný typ dat jsou rozhodně nejvyhledávanější.

Inara nabízí možnost jak dané informace na základě vstupních požadavků rychle najít, pohodlněji než hra samotná. Řada informací není také ve hře snadno či uceleně přístupná.

This type of solution is usually aimed at searching to begin with, but do you think the way players get to the data they want can be improved?

Jako vždy u všeho platí - vše se dá vylepšit. ;)

How much do you or your moderators need to oversee the data that flows in?

Minimálně, obvykle je to jen o řešení nějakých specifických vyjímek. Systém je nastaven tak, aby pokud možno vyžadoval co nejméně ručních zásahů.

How much do you or your moderators need to oversee the editing of the data that is already in?

Opět minimálně, jedná se zpravidla jen o korekce nečekaných menších problémů a nebo implementaci nových dat či definic nutných pro novou verzi hry.

Have you ever met intentional vandalism on this type of solution? If yes, how well could it be solved with your solution's tools?

V omezené míře, ano. V části případů se jednalo o nepochopení jak věci fungují, v části o cílený zlý úmysl. Drtivá většina podobných škod je odstraněna automaticky díky častým automatizovaným updatům dat, menší část se musí opravit ručně, pokud je to nezbytné.

A.4.3 Data from developers (game API, cooperation, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

Nelze přesně specifikovat, integrální část website. Předpoklad je, že drtivá většina uživatelů, tedy odhadem stejné nebo velmi podobné procento hráčů specifikované výše.

What type of data does your solution provide and why do you think it's not in the game or - in case it is in the game - it is duplicated outside of the game and possibly more searched there?

Všechna klíčová data hry, které je realisticky možné obsáhnout a dávají nějaký smysl a užitek větší skupině uživatelů.

How easy it is to get the data you want and/or need from the developers?

Záleží na typu dat, některá je možno dostat relativně jednoduše díky přímé podpoře od vývojářů (API), některá se musí vkládat ručně.

Is there something you'd like to see improved in getting the data to the players?

Větší/snadnější dostupnost některých typů dat, i s ohledem na automatizaci procesů.

Is there some type of feedback loop, both from you and players using this solution, towards the developers? How much, in your opinion, do you think it helps the game?

Do určité míry. Nelze však odhadnout jaký dopad to má na vylepšování hry (věc vývojářů hry), rozhodně má však nějaký dopad na vylepšování cest jakou jsou některá data dostupná 3rd party vývojářům.

Is there some common misconception you see players make, even though reliable data from developers are available and provided?

If yes, why do you think these misconceptions exist in your community? Is it the fault of the developers, of the solution, of the community, someone/something else?

Část hráčů nechápe, jakou cestou se data za hry dostávají do nástrojů, websites, atp. třetích stran jako Inara. Drtivá část nepochopení obvykle pramení z toho, že uživatelé prostě nečtou informace a vysvětlení, která jsou jim předkládána.

A.4.4 Etc.

Anything that you feel might help the reader understand how/why/when/etc. data flows from the game/developers/players to you and back to players that you feel might have been missing in the previous questions

Zajímavostí ohledně 3rd party vývojářské komunity kolem Elite: Dangerous je poměrně provázaný ekosystém, kdy například lokální uživatelské nástroje jako EDMC, EDDiscovery, EDDI a další kromě jiného i posílají část dat přes ZeroMQ dalším nástrojům a ty pak naopak mohou poskytnout nějaké informace zpět. Stejně tak mezi jednotlivými nástroji a websites do určité míry funguje (volitelná) provázanost přes vlastní API, exporty dat podobně. Domnívám se, že v porovnání s drtivou většinou nástrojů a komunit jiných her, které jsou obvykle striktně izolované, je to poměrně unikátní.

A.5 Eiyuu Senki komunita

Briefly introduce yourself (name or nickname) and what gaming community you (help) manage

Dlouhá léta používám nickname Artie a jsem jediným autorem jedné z největších website zaměřené na hru Elite: Dangerous - Inara (<https://inara.cz>). Další netřeba rozvádět, výsledná práce doufám hovoří za vše. :)

Write down which data types (from the four groups further below) you manage (If you can think of any type of data that players use to get information on the game and it is missing, feel free to mention it.

I intentionally didn't cover things like YouTube/Twitch here, because they aren't managed by one person and are individual and independent efforts.)

Website, API, herní databáze.

A.5.1 General part regarding data

Which solution do you use to keep this type of data?

Website, webová služba.

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

Klasické řešení v podobě Linux s Apache, PHP a MariaDB (bez uvedení verzí) na dedikovaném serveru. Nepoužít žádný framework, CMS, atp. vyjma ZeroMQ pro sdílení dat.

If you picked this solution yourself, why did you pick it over the competition (Discord over IRC, own wiki over Fandom wiki, etc.)? If someone else picked the solution and not you, why do you think they picked it?

Vzhledem k nutnosti hodně specifických řešení, byla to nejlepší volba.

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

Ne, není žádný konkrétní důvod měnit, řešení je plně vyhovující.

What would make you switch to another solution (aside of the current one becoming unusable for any reason)?

Nic, vzhledem k rozsahu projektu jiné řešení ani nepřipadá příliš v úvahu.

What could be most improved on your solution (in case it's more things pick two or three most desired)?

Bez odpovědi.

Is there some feature you'd like to see that the competition of your solution has and your solution not?

Bez odpovědi.

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?

Ano, vše je zálohováno, takže řešení je obnovení ze záloh.

If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?

Zaleží na rozsahu problému, od pár minut až několik hodin, s dopadem na komunitu mizivým nebo omezeným (menší rollback), opět v závislosti na rozsahu problému.

Why do you think players use your solution to search for data to begin with?

Obvykle pohodlnější, rychlejší a ucelenější přístup k informacím než přímo ve hře, plus přidaná hodnota v podobě dalších výstupů (vyhledávání, přehledy, statistiky, atp.) a uživatelského obsahu (galerie, logbooks, diskuse, atp.).

A.5.2 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

Podle velmi konzervativního odhadu (s ohledem na sporá oficiální data o počtu aktivních hráčů) je Inara pravidelně používána minimálně 20% hráčské základny. Realisticky spíše více.

What do you feel is the most searched thing regarding the game and why?

Data o herních obchodovatelných komoditách, vybavení lodí a jejich modifikace (engineering), informace o stanicích a podobný typ dat jsou rozhodně nejvyhledávanější.

Inara nabízí možnost jak dané informace na základě vstupních požadavků rychle najít, pohodlněji než hra samotná. Řada informací není také ve hře snadno či uceleně přístupná.

This type of solution is usually aimed at searching to begin with, but do you think the way players get to the data they want can be improved?

Jako vždy u všeho platí - vše se dá vylepšit. ;)

How much do you or your moderators need to oversee the data that flows in?

Minimálně, obvykle je to jen o řešení nějakých specifických výjimek. Systém je nastaven tak, aby pokud možno vyžadoval co nejméně ručních zásahů.

How much do you or your moderators need to oversee the editing of the data that is already in?

Opět minimálně, jedná se zpravidla jen o korekce nečekaných menších problémů a nebo implementaci nových dat či definic nutných pro novou verzi hry.

Have you ever met intentional vandalism on this type of solution? If yes, how well could it be solved with your solution's tools?

V omezené míře, ano. V části případů se jednalo o nepochopení jak věci fungují, v části o cílený zlý úmysl. Drtivá většina podobných škod je odstraněna automaticky díky častým automatizovaným updatům dat, menší část se musí opravit ručně, pokud je to nezbytné.

A.5.3 Data from developers (game API, cooperation, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

Nelze přesně specifikovat, integrální část website. Předpoklad je, že drtivá většina uživatelů, tedy odhadem stejné nebo velmi podobné procento hráčů specifikované výše.

What type of data does your solution provide and why do you think it's not in the game or - in case it is in the game - it is duplicated outside of the game and possibly more searched there?

Všechna klíčová data hry, které je realisticky možné obsáhnout a dávají nějaký smysl a užitek větší skupině uživatelů.

How easy it is to get the data you want and/or need from the developers?

Záleží na typu dat, některá je možno dostat relativně jednoduše díky přímé podpoře od vývojářů (API), některá se musí vkládat ručně.

Is there something you'd like to see improved in getting the data to the players?

Větší/snadnější dostupnost některých typů dat, i s ohledem na automatizaci procesů.

Is there some type of feedback loop, both from you and players using this solution, towards the developers? How much, in your opinion, do you think it helps the game?

Do určité míry. Nelze však odhadnout jaký dopad to má na vylepšování hry (věc vývojářů hry), rozhodně má však nějaký dopad na vylepšování cest jakou jsou některá data dostupná 3rd party vývojářům.

Is there some common misconception you see players make, even though reliable data from developers are available and provided?

If yes, why do you think these misconceptions exist in your community? Is it the fault of the developers, of the solution, of the community, someone/something else?

Část hráčů nechápe, jakou cestou se data za hry dostávají do nástrojů, websites, atp. třetích stran jako Inara. Drtivá část nepochopení obvykle pramení z toho, že uživatelé prostě nechtou informace a vysvětlení, která jsou jim předkládána.

A.5.4 Etc.

Anything that you feel might help the reader understand how/why/when/etc. data flows from the game/developers/players to you and back to players that you feel might have been missing in the previous questions

Zajímavostí ohledně 3rd party vývojářské komunity kolem Elite: Dangerous je poměrně provázaný ekosystém, kdy například lokální uživatelské nástroje jako EDMC, EDDiscovery, EDDI a další kromě jiného i posílají část dat přes ZeroMQ dalším nástrojům a ty pak naopak mohou poskytnout nějaké informace zpět. Stejně tak mezi jednotlivými nástroji a websites do určité míry funguje (volitelná) provázanost přes vlastní API, exporty dat podobně. Domnívám se, že v porovnání s drtivou většinou nástrojů a komunit jiných her, které jsou obvykle striktně izolované, je to poměrně unikátní.

A.6 Genshin Impact komunita, Torn komunita

A.6.1 General part

Briefly introduce yourself (name or nickname) and what gaming community you (help) manage

I am Helcostr (aka Echo), and I am one of the community admins for Fandom's Genshin Impact Wiki as well as one of the community admins for Torn's Official Wiki (text RPG).

Write down which data types (from the four groups further below) you manage (If you can think of any type of data that players use to get information on the game and it is missing, feel free to mention it.

I intentionally didn't cover things like YouTube/Twitch here, because they aren't managed by one person and are individual and independent efforts.)

I manage all the data listed except for the method of general storage of the data.

A.6.2 General part regarding data

Which solution do you use to keep this type of data?

(I'll keep referring to this as a „your solution“ only from now on)

Genshin: We store data in Fandom's MediaWiki system (<https://www.mediawiki.org/>): we as users are not aware of the specifics.

Torn: We store data in a server provided by Torn: we as users are not aware of the specifics.

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

(More information is better, but I understand this question in particular can be security risk in some cases.)

Genshin: This is a cloud based solution where we don't worry about the solution and we let Fandom take care of the storage for us.

Torn: This is a cloud based solution that is managed in the dev house, and we let the devs take care of it for us.

If you picked this solution yourself, why did you pick it over the competition (Discord over IRC, own wiki over Fandom wiki, etc.)? If someone else picked the solution and not you, why do you think they picked it?

Genshin: I didn't pick the solution: I was drawn to it. With custom implementation methods, they become more tech savvy, very niched, and might get lost in the WWW if it's not well designed/gets enough traction. Using popular tools, like Discord and Fandom, shortcuts the recognition (at the cost of losing direct control over the data).

Torn: this is a very big WIP project, and as it's the official wiki, I am always looking to improve it, even though it's not the most prettiest solution just yet.

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

Genshin: My big deciding factor is the startup/maintenance cost of any project I start. As for similar solutions, I've known Fandom's system ever since 2014, and their system is one of the best systems to easily manipulate every aspect of the data (MediaWiki is a pretty decent markup language).

Torn: I have considered fleeing to Fandom, but as this is the official wiki, I am now striving to make it a better place.

What would make you switch to another solution (aside of the current one becoming unusable for any reason)?

Genshin: If we were contacted directly by the source material (in this case, the Genshin Impact team), I would be willing to migrate the information away from our fan wiki toward the official wiki. Hopping from fan wiki to fan wiki isn't in my consideration unless there is a huge community wide synchronization to merge into a specific solution.

Torn: If the dev team is struggling and wants to offload the weight to a 3rd party company, I will follow their lead.

What could be most improved on your solution (in case it's more things pick two or three most desired)?

Genshin: More control over the design would be ideal: making sure that the look of the site services several different platforms (desktop, mobile, print)

Torn: Faster communication between the wiki dev and the wiki designer.

Is there some feature you'd like to see that the competition of your solution has and your solution not?

Genshin: I personally think other company-based competition sucks. Specific fan tools that serve specific needs, however, are great and I would love to incorporate a soft version of those tools so that this project is the one stop shop for all the needs.

Torn: We are in an isolated bubble, and as all projects are volunteer projects, both fan tools as well as the official wiki are pretty much at the same quality level.

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?

I personally don't care that the data is lost. What I do care about is that the company hosting the solution is no longer reputable at this point.

However, there are built in tools to obtain backup snapshots of the wiki.

If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?

Time-wise, it shouldn't be an issue. As long as we are in the process of recovery, the community can't stay mad forever. If we show no signs of recovery, it makes sense for the community to turn to other solutions.

Why do you think players use your solution to search for data to begin with?

Genshin: I believe it's the familiarity, as well as the wealth of information all in one location.

Torn: it's official.

A.6.3 Fast-flowing data (IRC, Discord, Slack, Reddit, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

This is very much the raw source of data, since it's centered around communication. This kind of information is very temporary, and can get lost without someone immortalizing the information.

In case the above is not zero, what do you feel is the most searched thing regarding the game and why?

People are always looking for the latest information/to lazy to go digging through collective permanent solution.

How good is your solution of this type with searching data, is it easy for your players to find the data they need?

Discord has a pretty decent indexing feature, but as a programmer, I wish I could do more with the data.

How much do you or your moderators need to oversee the data that flows around the community?

Genshin: When we have time, we monitor the fast flowing information to make sure there are no leaks.

Torn: we don't have an in game policy against information, although we do have a team separate from the wiki team who is privy to future plans to the game itself. This team is under an NDA like, and are punished for leaking information.

Have you ever met intentional spreading of fake data (misinforming new players, fake leaks, etc.) on this type of solution? If yes, how well could it be solved with your solution's tools?

Genshin: Escalation of warn, mute, ban.

Torn: we don't moderate.

A.6.4 Collective permanent solution (Fandom wiki, own wiki, CMS, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

Very much the homepage of information.

What do you feel is the most searched thing regarding the game and why?

People are looking for the latest reliable information.

This type of solution is usually aimed at searching to begin with, but do you think the way players get to the data they want can be improved?

In an ideal world, we would be able to pull up articles from context. But for now, we must take context, translate it into words, search, and then find the source of data that satisfy our needs.

How much do you or your moderators need to oversee the data that flows in?

We attempt to prevent edit wars, and pure speculation: the community checks itself and works mostly autonomously.

How much do you or your moderators need to oversee the editing of the data that is already in?

Not much. We adopt the policy of „all edits are in good faith“. We do have some tools to help us patrol the wiki at a macro level, but we don't fact check every micro detail unless something makes us intrigued.

Have you ever met intentional vandalism on this type of solution? If yes, how well could it be solved with your solution's tools?

Genshin: I don't treat vandalism as intentional: some people are expressive and want to vandalize public property. But for vandalism in general, escalation of warn, short ban, longer ban, etc.

Torn: contributors are whitelisted, and so vandalism is more punishable, since each edit is not anonymous; it's tied to a user (multi account isn't allowed).

A.6.5 Data from developers (game API, cooperation, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

It is the source of the information, and gives our collective permanent solution validity.

What type of data does your solution provide and why do you think it's not in the game or - in case it is in the game - it is duplicated outside of the game and possibly more searched there?

Duplicated, and then added on.

How easy it is to get the data you want and/or need from the developers?

Genshin: Very hard.

Torn: Not too hard.

Is there something you'd like to see improved in getting the data to the players?

Genshin: A good API for data is always handy, I would love to see it.

Torn: explicit explanation on how things work is great (for now, we only have community information on how things work). (there is a game api that can be accessed if basic information is needed)

Is there some type of feedback loop, both from you and players using this solution, towards the developers? How much, in your opinion, do you think it helps the game?

Genshin: there is no communication

Torn: this is a very newbie project that is growing as the community is helping shape it's decision for the api. Every decision made is focused at ease of development, and not good against protection.

Is there some common misconception you see players make, even though reliable data from developers are available and provided?

If yes, why do you think these misconceptions exist in your community? Is it the fault of the developers, of the solution, of the community, someone/something else?

The data is tricky to get to, and misconceptions breed. It's partially because the information is not live/up to date/data is misinterpreted.

A.6.6 Data mining (data mining, reverse engineering, modding, etc.)

How much do you think this type of service is used by players (or possibly a subset like speedrunners) in your community to get some data they need?

This data is used and crunched into bite sized information for people to digest.

How easy it is to get the data you want from the game?

Very very hard: security is the game's goal, and information is our goal: those two don't mix well.

Since it's usually harder to get this type of data than the others, do you feel that it tends to be kept more secret or do you feel that people share the data as soon and openly as possible?

Secret is what people turn to first. But information that isn't used is pretty much dead information, and the digging would have been moot.

Have you ever run into issues with the developer not wanting to have data from their game available publicly? If yes, how was the situation solved?

Genshin: Obtaining information about future content is a big nono.

Obtaining information about current content is fine: as we could have experienced the content through normal means.

Their PR states that leaks lead to hype, and because content is still in development, hype can turn into anger even before the content is released.

Torn: Reverse engineering math formulas is pretty much the only thing we can do.

Specifically for very hard-to-perform communities (speedrunning, etc.), how big do you estimate is the retention of the players if the techniques are kept public? That is, if one hundred people read „speedrunning wiki and tips“, how many of them will stick around and actually try speedrunning or at least actively join the community?

I'd say a small fraction would be interested in helping speed run. My goal isn't to cater to one audience, but rather to a general public. Speed running information, like bugs, are best suited in the after section, where the bugs and quirks, and patchnotes lie. Speed running tactics are best suited in a blog format, as they are per-person and not indicative of the main content. A wiki dedicated to collecting speed running information would be the best way to aggregate this data.

A.6.7 Etc.

Anything that you feel might help the reader understand how/why/when/etc. data flows from the game/developers/players to you and back to players that you feel might have been missing in the previous questions

I treat any wiki as a shadow copy of the original source (whether it be a game, tv show, or real life). A reader should be able to take the information from the encyclopedic entries and replicate/relive the scenario for themselves. But I don't promote the duplication of Intellectual Property: I only promote the archiving of such „information“.

A.7 Kantai Collection komunita

A.7.1 General part

Hello, I am Raphael Kratz (DoktorKaputt/Grunilg) and manage the Kantai Collection subreddit (<https://www.reddit.com/r/kancolle/>) as well as being an editor for the english kancolle wiki (http://en.kancollewiki.net/Kancolle_Wiki), moderator of the accompanying Discord and a Contributor to the English KC Patch (<https://github.com/InochiPM/KanColle-English-Patch-KCCP>) The Data types I manage are therefore Fast-flowing, Collective/permanent, and a bit of Developer Data.

A.7.2 Fast-flowing data (Reddit)

Which solution do you use to keep this type of data?

reddit

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

The Cloud (reddit)

If you picked this solution yourself, why did you pick it over the competition (Discord over IRC, own wiki over Fandom wiki, etc.)? If someone else picked the solution and not you, why do you think they picked it?

reddit is the frontpage of the internet (unironically), so having a subpage on there for the community is often the first stepping stone for someone interested in the topic to find more resources

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

n/A, you would need to ask Rui for redditlore. Possibly the himeuta forums?

What would make you switch to another solution (aside of the current one becoming unusable for any reason)?

n/A

What could be most improved on your solution (in case it's more things pick two or three most desired)?

More slots to pin relevant threads

Easier editing of the subreddit wiki

Unified design, not having to account for new, legacy and mobile users

Is there some feature you'd like to see that the competition of your solution has and your solution not?

see above

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?

There is no backup made by me personally, but there are plenty of crawlers and I hope that reddit itself has a backup of their content

If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?

No longer than a couple of days, as we have the assets and CSS to recreate the community theme saved off-site. Since reddit is more in-the-moment and focused on contemporary content rather than archiving, it wont affect the community too much.

Why do you think players use your solution to search for data to begin with?

It's more permanent than Discord and allows for more structured discussion (Both in markdown options and comment tree layout). Every Topic is self-contained and once it lost its relevancy, it disappears.

How much do you think this type of service is used by players in your community to get some data they need?

Reddit is mostly used as an information source during patch day as one community member (StoneFlame) is fast at translating the patchnotes and quest descriptions in one permanent thread (instead of scattered messages across a board or Discord channel)

In case the above is not zero, what do you feel is the most searched thing regarding the game and why?

Patchnotes during Patchday and Event Guides during Event time.

How good is your solution of this type with searching data, is it easy for your players to find the data they need?

If they are looking for current topics of the community, they will find a thread relatively quickly, otherwise they need to query the hivemind and ask a question themselves. Depending on quality, time and willingness of the other community members they might get a helpful answer or none.

How much do you or your moderators need to oversee the data that flows around the community?

It's quite a slow subreddit, so I usually read all posts that are made in a day within 15 minutes or so.

Have you ever met intentional spreading of fake data (misinforming new players, fake leaks, etc.) on this type of solution? If yes, how well could it be solved with your solution's tools?

Yes. The offending Members will get warned and get banned upon repeat offense

A.7.3 Fast-flowing data (en Kancolle Wiki Discord)

Which solution do you use to keep this type of data?

Discord, a chat Service for gamers

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

The Cloud (Discord)

If someone else picked the solution and not you, why do you think they picked it?

It's a supplementary and more immediate way of communicating with other editors instead of using the discussion pages on the wiki

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

n/A

What would make you switch to another solution (aside of the current one becoming unusable for any reason)?

n/A

What could be most improved on your solution (in case it's more things pick two or three most desired)?

Better bot integration to access wiki pages easier

Is there some feature you'd like to see that the competition of your solution has and your solution not?

Teamspeak has better ways to distribute and manage rights along the Server members

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?

The wiki still stands, as will probably the other Kancolle Discord Community Servers

If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?

n/A, it isnt possible to backup Discord Servers. No great effect as it is supplementary to the wiki

Why do you think players use your solution to search for data to begin with?

There is something to be said of the power of the hivemind os to speak. The game as it is is complex enough so some kind of live support and more community input might be desired by some players to make the right decision regarding their gameplay.

How much do you think this type of service is used by players in your community to get some data they need?

A significant portion of players use Discord as supplementary info to the wiki

In case the above is not zero, what do you feel is the most searched thing regarding the game and why?

Fleet compositions for events and quests as fleet building is the most complex part of the game, as minor changes could have big impacts on the performance of the fleet

How good is your solution of this type with searching data, is it easy for your players to find the data they need?

Discords search functions sucks, but important resources are pinned in relevant channels. Otherwise they would have to rely on the active members of the community to point them where they need be.

How much do you or your moderators need to oversee the data that flows around the community?

There is a lot of noise (chit-chat of no value), but moderation is taken very rarely

Have you ever met intentional spreading of fake data (misinforming new players, fake leaks, etc.) on this type of solution? If yes, how well could it be solved with your solution's tools?

Yes, sometimes it stems from incompetence, sometimes from malice. The Solution is warning the infringing member and if they keep going, banning them. Unfortunately making alternative Accounts for Discord is too easy.

A.7.4 Collective permanent solution (en Kancolle Wiki)**Which solution do you use to keep this type of data?**

A Website

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

n/A, you would need to ask gensui or Ravien for technical details. For displaying Content, we are using Mediawiki.

If you picked this solution yourself, why did you pick it over the competition (Discord over IRC, own wiki over Fandom wiki, etc.)? If someone else picked the solution and not you, why do you think they picked it?

Mediawiki on an own Server was chosen because of ads on the competing platform (Fandom). You might need to ask Totaku or Ravien for more details on the Wiki Lore

Have you or people who managed this solution before you ever considered switching to a different solution? If yes, to which one and why didn't you make the switch? If not, why not?

Because of ads and the global developers ruining the back and frontend of the Fandom wiki, they migrated to the en Kancolle wiki. I dont see a reason to switch from the current platform

What would make you switch to another solution (aside of the current one becoming unusable for any reason)?

n/A

What could be most improved on your solution (in case it's more things pick two or three most desired)?

internal structure and discoverability

Is there some feature you'd like to see that the competition of your solution has and your solution not?

n/A

If your solution fails, how will it deal with potential loss of accumulated data, is there any backup of sorts?

There should be a backup, see Ravien/gensui

If a backup is possible, how long do you think it would take to deploy and how much would it affect the community?

There should be a backup, see Ravien/gensui

Why do you think players use your solution to search for data to begin with?

Permanent documentation in a resource-rich environment, which has been fact checked and iterated over by multiple individuals

How much do you think this type of service is used by players in your community to get some data they need?

Almost all players are using the wiki(s) in one form or another during Events or to look up general gameplay

What do you feel is the most searched thing regarding the game and why?

Fleet compositions for events and quests as fleet building is the most complex part of the game, as minor changes could have big impacts on the performance of the fleet

This type of solution is usually aimed at searching to begin with, but do you think the way players get to the data they want can be improved?

Categorisation at the bottom of the page and linking all relevant mechanics when they are mentioned in the article.

How much do you or your moderators need to oversee the data that flows in?

Not much

How much do you or your moderators need to oversee the editing of the data that is already in?

Not much

Have you ever met intentional vandalism on this type of solution? If yes, how well could it be solved with your solution's tools?

No, soloblast might know some lore here though

A.8 Kantai Collection API

A.8.1 General part

Briefly introduce yourself (name or nickname) and what gaming community you (help) manage

I am Dewy, I help to research the game mechanics of Kantai Collection (KC). KC is a simulation game that requires very little player input after the player decides what strategy to take.

Write down which data types (from the four groups further below) you manage

I manage data mining / reverse engineering and data from game API to a limited extent.

A.8.2 General part regarding data

What technical background is behind your solution (cloud vs own hardware, what hardware, what version of own wiki/CMS, etc.)

My team manages a Chrome extension (KC3Kai) that parses data from the game API as the user plays and presents extra information that the game does not present. Some of this data is submitted to a database (TsunDB) for research on game mechanics by user consent.

Which solution do you use to keep this type of data?

KC3Kai stores user data in the native Chrome IndexedDB. TsunDB is built on PostgreSQL.

A.8.3 Data from developers (game API, cooperation, etc.)

How much do you think this type of service is used by players in your community to get some data they need?

I think that these services are very much used by players. The game provides very little information to the player and so they fill the missing gap.

What type of data does your solution provide and why do you think it's not in the game or - in case it is in the game - it is duplicated outside of the game and possibly more searched there?

KC3Kai helps users calculate how changes they make might affect gameplay. For example, it helps calculate potential damage or hidden variables such as accuracy and evasion per strategy setup. I think that the reason why this data is not provided by the game is because of the initial appeal of the game: a community driven approach to find how the mechanics work.

How easy it is to get the data you want and/or need from the developers?

The data obtained is intercepted from the game as the user plays. The developers have recently implemented an IP ban for users outside Japan, and so obtaining metadata from the developers have become more difficult.

A.8.4 Data mining (data mining, reverse engineering, modding, etc.)

How much do you think this type of service is used by players (or possibly a subset like speedrunners) in your community to get some data they need?

With the lack of information from the developers, this is the remaining method for players to gain information they need.

How easy it is to get the data you want from the game?

Tools have been developed for logging data from the game. Initially, I used these to personally collect data and form my own analysis. It was extremely time consuming to obtain enough data for meaningful analysis. Later on, I collaborated with Tsubakura to implement a community submission feature in KC3Kai to submit data to TsunDB. This has made it a lot easier to collect the data I need.

Since it's usually harder to get this type of data than the others, do you feel that it tends to be kept more secret or do you feel that people share the data as soon and openly as possible?

I feel that discretion has to be used. These results have a strong impact on player choice especially in games with limited information. Given the challenge of obtaining data, sometimes a wrong conclusion is reached from insufficient data and creates a false narrative. Misinformation is often hard to correct, as seen from recent history. Hence, I do feel it is important to share this data, but once enough confidence has been built.

Bibliografie

1. STEINBERG, S. *The definitive Guide: Videogame Marketing and PR (1st ed.)*. iUniverse, 2007. ISBN 978-0-59543-371-1.
2. SATORU, I. *Satoru Iwata's „Heart of a Gamer“ speech from the 2005 Game Developers Conference* [Online]. 2005 [cit. 2015-04-25]. Dostupné z: <https://www.youtube.com/watch?v=cYcBq8TN0gg>.
3. SMITH, A. *One, two, three, four I declare a space war* [Online]. 2014 [cit. 2015-04-25]. Dostupné z: <https://videogamehistorian.wordpress.com/2014/08/07/one-two-three-four-i-declare-a-space-war/>.
4. WALLACH, O. *50 Years of Gaming History, by Revenue Stream (1970-2020)* [Online]. 2020 [cit. 2015-04-25]. Dostupné z: <https://www.visualcapitalist.com/50-years-gaming-history-revenue-stream/>.
5. INTERNATIONAL TELECOMMUNICATION UNION. *ICT Data and Statistics (IDS)* [Online]. 2007 [cit. 2015-04-25]. Dostupné z: <http://www.itu.int/ITU-D/ict/statistics/ict/>.
6. INTERNATIONAL TELECOMMUNICATION UNION. *Measuring digital development, Facts and figures, 2019*. Place des Nations, CH-1211 Geneva 20, Switzerland: ITU Publications, 2019. ISBN 978-92-61-29521-9.
7. EDWARDS, C. *Valve Lines Up Console Partners in Challenge to Microsoft, Sony* [Online]. 2013 [cit. 2015-04-25]. Dostupné z: <https://www.bloomberg.com/news/articles/2013-11-04/valve-lines-up-console-partners-in-challenge-to-microsoft-sony>.
8. EPIC GAMES. *Epic Games Store 2020 Year in Review* [Online]. 2021 [cit. 2015-04-25]. Dostupné z: <https://www.epicgames.com/store/en-US/news/epic-games-store-2020-year-in-review>.
9. ORLAND, K. *The return of the \$70 video game has been a long time coming* [Online]. 2021 [cit. 2015-05-07]. Dostupné z: <https://arstechnica.com/gaming/2020/07/the-return-of-the-70-video-game-has-been-a-long-time-coming/>.
10. JENKINS, H. Educational researcher. 2011, roč. 11, č. 1. ISSN 1604-7982.
11. ROLFE, J. *Big Rigs: Over the Road Racing (PC) - Angry Video Game Nerd (AVGN)* [Online]. 2014 [cit. 2015-05-08]. Dostupné z: <https://www.youtube.com/watch?v=h6DtVHqyYts>.
12. ASHEN, S. *Hareraiser (The Worst Game Ever) - Stuart Ashen - Norwich Gaming Festival 2017* [Online]. 2017 [cit. 2015-05-08]. Dostupné z: <https://www.youtube.com/watch?v=ouvi-fwrfIY>.
13. GODIN, S. *Learning by analogy* [Online]. 2009 [cit. 2015-05-09]. Dostupné z: <https://seths.blog/2009/11/learning-by-analogy-2/>.

14. MEDLER B. *Game Studies*. 2005, roč. 62, č. 7, s. 48–51. ISSN 0013-1784.
15. STONE, T. *Scripting API now in public beta!* [Online]. 2018 [cit. 2015-05-13]. Dostupné z: <https://www.minecraft.net/sv-se/article/scripting-api-now-public-beta>.
16. JOBST, K. *Can You Spot The Fake Speedrun?* [Online]. 2020 [cit. 2015-05-13]. Dostupné z: <https://youtu.be/Nq-y-TJ9PFU?t=1026>.
17. MODSERN VINTAGE GAMER. *How Diablo was completely Reverse Engineered without Source Code — MVG* [Online]. 2019 [cit. 2015-05-13]. Dostupné z: https://www.youtube.com/watch?v=5tADL_fmSHQ.
18. CD PROJEKT RED. *Important Update* [Online]. 2021 [cit. 2015-05-13]. Dostupné z: <https://twitter.com/CDPROJEKTRED/status/1359048125403590660>.
19. BBC CLICK. *Elite: Dangerous - Unknown Probes - BBC Click Segment* [Online]. 2016 [cit. 2015-05-13]. Dostupné z: <https://www.youtube.com/watch?v=Y7irrU85Nz0>.
20. SATO. *NieR: Automata Spinoff And Prequel To Be Told Through A Stage Play And Musical Next Year In Japan* [Online]. 2017 [cit. 2015-05-13]. Dostupné z: <https://www.siliconera.com/nier-automata-spinoff-prequel-told-stage-play-musical-next-year-japan/>.
21. TOTAKU. *Live Event: „The Zuiun Festival begins!“* [Online]. 2017 [cit. 2015-05-13]. Dostupné z: https://en.kancollewiki.net/Zuiun_Festival#Event_Detail_Summary.
22. MORIN, R. *Inside the Artificial Universe That Creates Itself* [Online]. 2016 [cit. 2015-05-13]. Dostupné z: <https://www.theatlantic.com/technology/archive/2016/02/artificial-universe-no-mans-sky/463308/>.
23. ROBERTS, C. *Inside Star Citizen - BBC Click* [Online]. 2019 [cit. 2015-05-13]. Dostupné z: https://www.youtube.com/watch?v=W_KUzkJoEhU&t=632s.
24. KAPOR, M. *Internet Quotation Appendix* [Online]. 2002 [cit. 2015-05-08]. Dostupné z: https://cyber.harvard.edu/archived_content/people/reagle/inet-quotations-19990709.html.

Obsah přiloženého média

readme.txt.....	stručný popis obsahu média
src	
├ thesis.....	zdrojová forma práce ve formátu L ^A T _E X
│├ caseStudies.....	případové studie
│└ images.....	obrázky importované do práce
text.....	text práce
└ thesis.pdf.....	text práce ve formátu PDF