

NÁVRH A PLÁNOVÁNÍ DIGITÁLNÍHO REPOZITÁŘE

MICHAL KONEČNÝ

NÁRODNÍ TECHNICKÁ KNIHOVNA

29. ŘÍJNA 2019



CO NÁS ČEKÁ?

1. Představení
2. Jak to začíná?
3. Plánujeme archiv
4. Řídíme jeho vznik
5. Workshop
6. Otázky a odpovědi



O MNĚ

- Pracuji jako Senior Solution Architect v mezinárodní IT firmě
- Vystudoval jsem KISK
- 25 let vytvářím digitální obsah
- Digitální kurátorství je mým koníčkem

UDĚLÁME ARCHIV!

Kdo? Jak? Proč???

- Digital Preservation Handbook

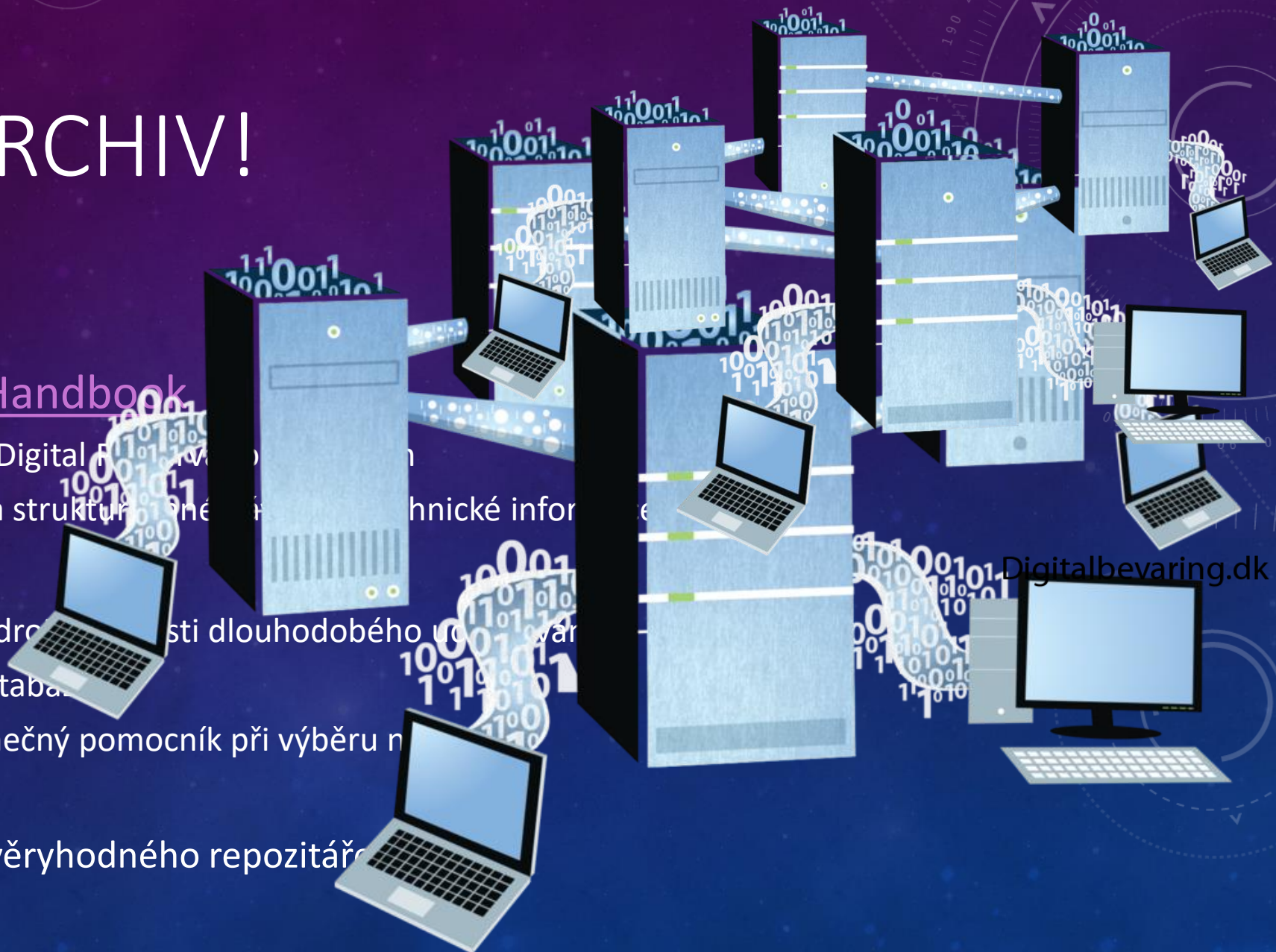
- Vytvářený a spravovaný Digital Preservation Handbook
- Obsahuje srozumitelné a strukturované technické informace

- DigiPres Commons

- Rozcestníků důležitých zdrojů pro dlouhodobé uchování
- Zahrnuje formátovou databázi
- POWRR Tool Grid – jedinečný pomocník při výběru nástrojů

- Platter

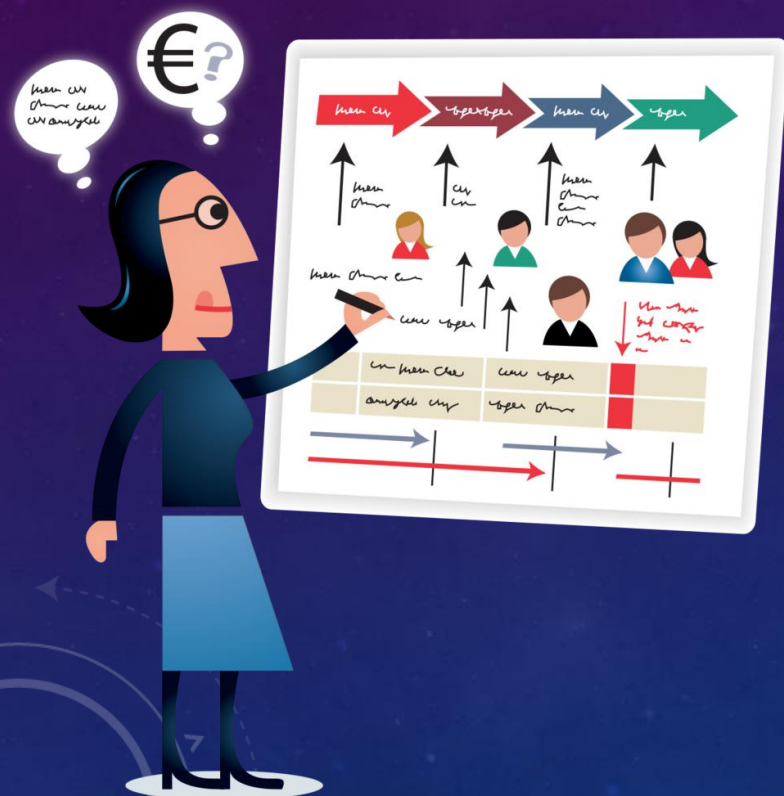
- Průvodce tvorbou důvěryhodného repozitáře



Digitalbevaring.dk

Digitalbevaring.dk

PLÁNUJEME ARCHIV



1. Poznejte dobře instituci a klíčové stakeholdery
2. Poznejte dobře digitální objekty a vytvořte si jejich registr
3. Vyhodnoťte připravenost instituce a zajistěte základní bezpečnost dat
4. Zdokumentujte procesy
5. Definujte strategii a pravidla dlouhodobé ochrany
6. Vyberte vhodné nástroje a navrhňte digitální repozitář
7. Naplánujte akce dlouhodobé ochrany
8. Rozvíjejte znalosti a kompetence
9. Sledujte změny a novinky

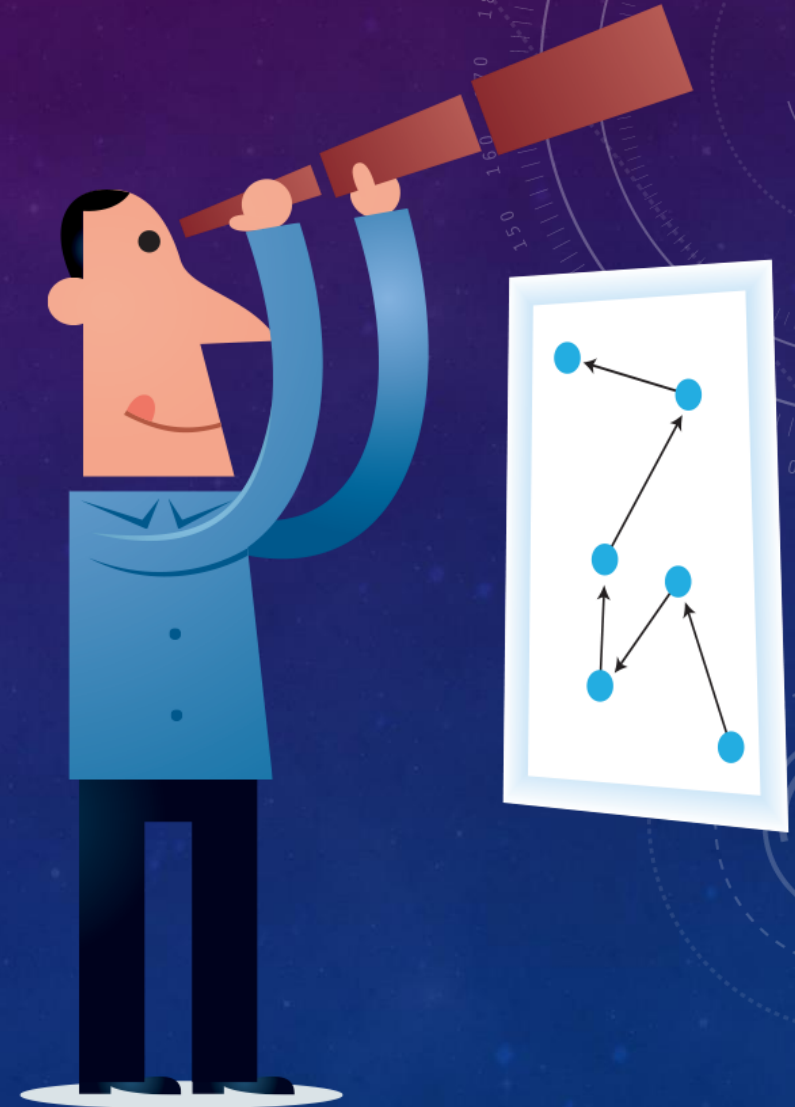
INSTITUTE A STAKEHOLDERĚŘI

1. Co je naše instituce zač?
2. Jaká je její vize a mise?
3. Jaká je její strategie digitální ochrany?
4. Jakou roli v této strategii hraje náš archiv?
5. Kdo jsou klíčoví stakeholderi?



VICE, MISE, STRATEGIE...

- **Vize**
 - Představa o budoucím stavu: odpověď na otázku “Co?”
- **Mise**
 - Poslání, smysl a účel organizace, důvod její existence. Odpověď na otázku “Proč?”
- **Strategie**
 - Prostředky a dílčí cíle pro naplnění vize. Odpověď na otázku “Jak?”



STAKEHOLDĚŘI: CO JSOU ZAČ?



Digitalbevaring.dk

User stories

- Používají se při (agilním) vývoji software a produktů
- Obsahují pouze takové množství informací, které dovolí odhadnout náročnost implementace
- Stručnější a jednodušší než *use cases* nebo *use scenarios*
- Formulují **kdo** chce **co** dělat a **proč**:

[As a anonymous user] [I want to download PDF document] so I could [read it on my mobile device.]

[Jako nepřihlášený uživatel] [chci stáhnout PDF dokument,] [abych si mohl přečíst na svém mobilním telefonu.]

- Pro práci s user stories existuje řada [online](#) i offline nástrojů a [šablon](#)

REGISTR DIGITÁLNÍCH OBJEKTŮ



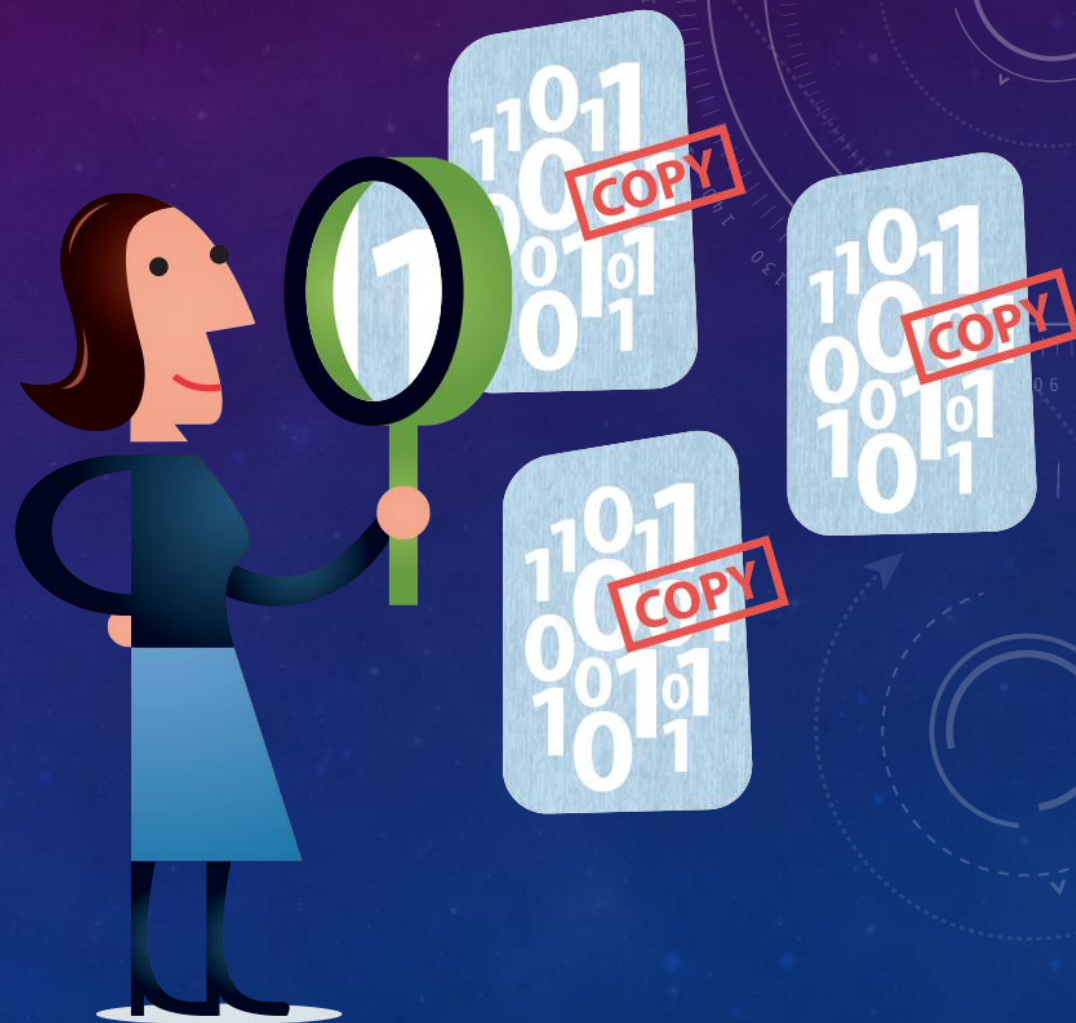
Digitalbevaring.dk

1. Co je předmětem kolekce?
2. Odkud se to vzalo a jaký to má účel?
3. Kde a jakým způsobem je to uložené?
4. Proč se to uchovává?
5. Kdo je zodpovědný? Kdo to používá?
Koho se to týká?
6. Jak jsou data zpřístupněna?
7. Budou se data měnit a růst v
budoucnosti?

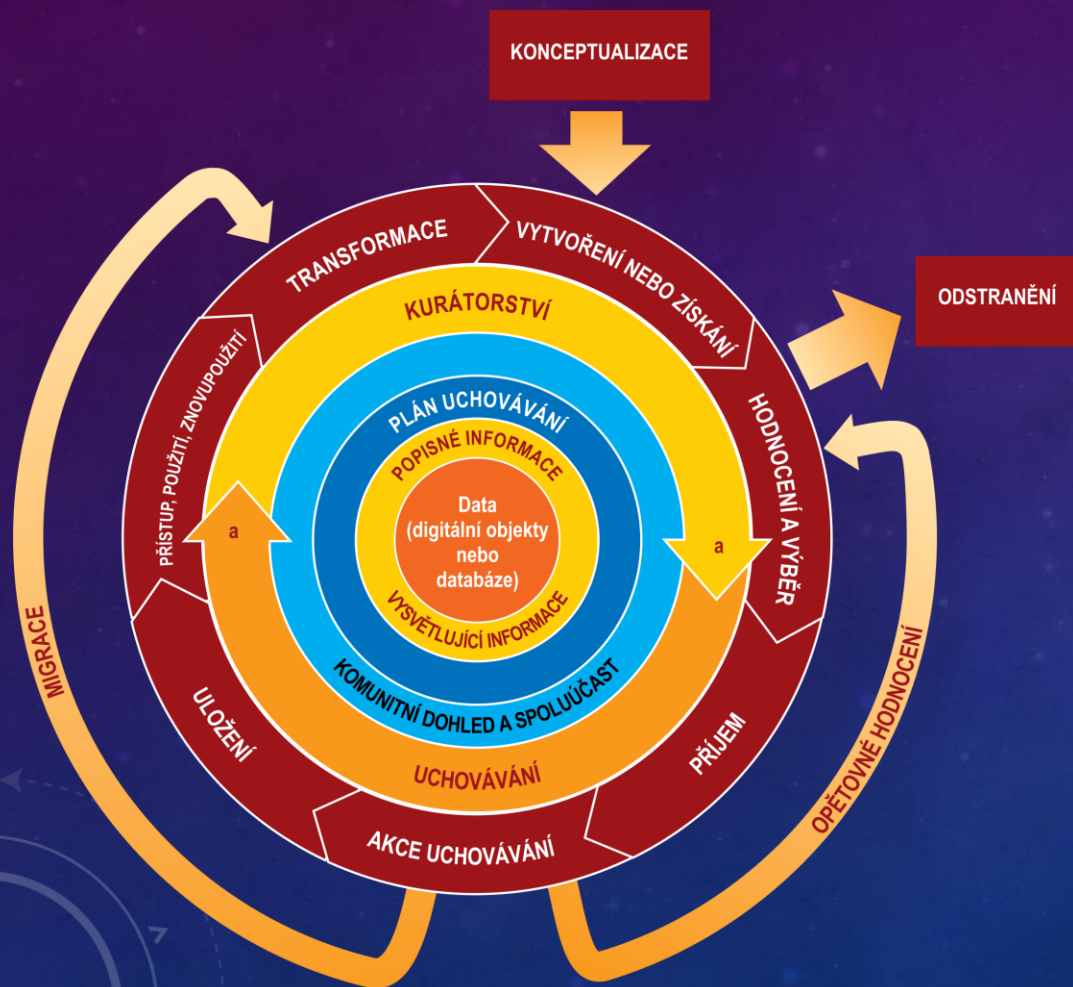
JSME PŘIPRAVENÍ?

Připravenost organizace podle
NDSA:

- Úroveň 1: Chrání svoje data
- Úroveň 2: Zná svoje data
- Úroveň 3: Sleduje svoje data
- Úroveň 4: Opravuje svoje data



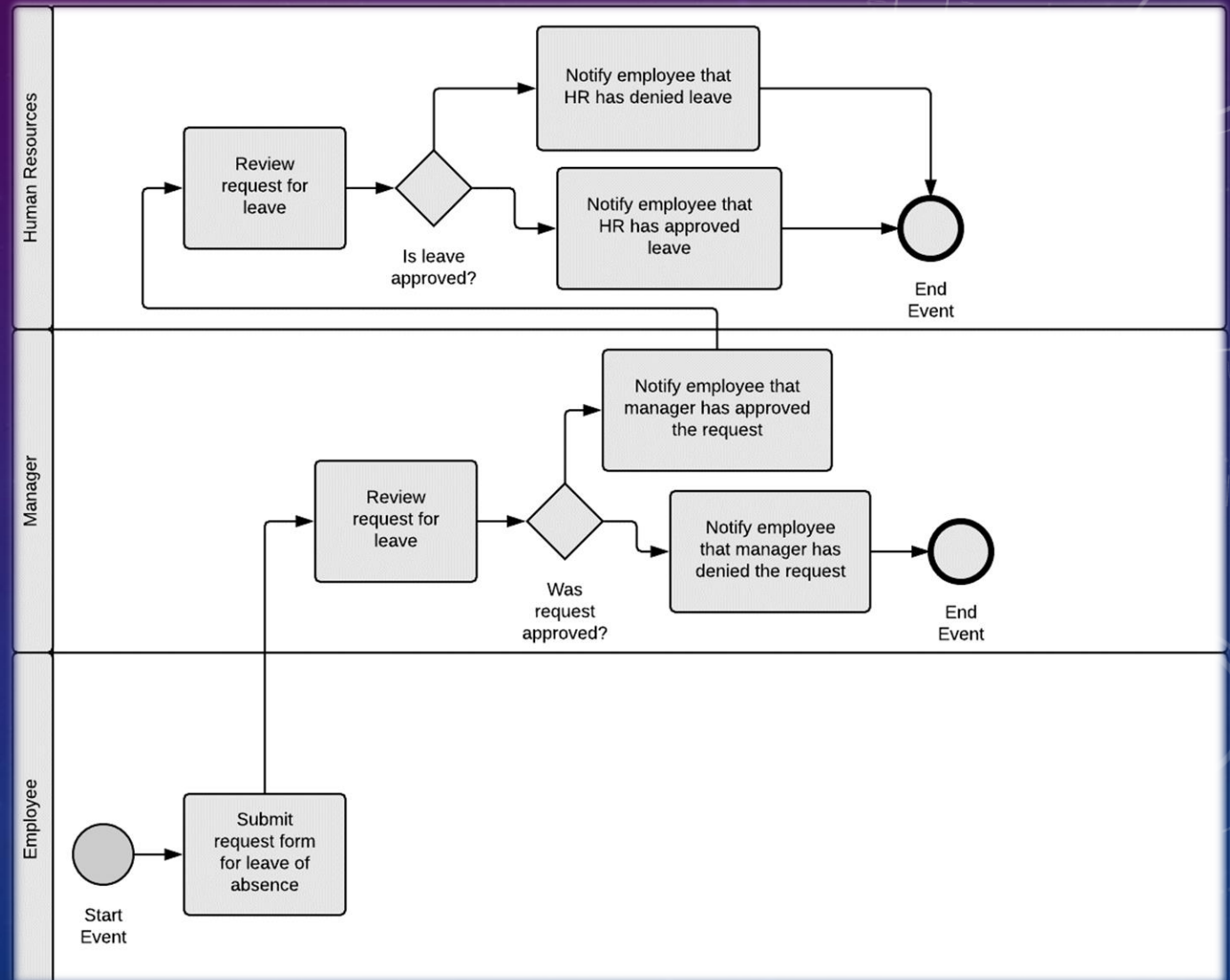
DLOUHODOBÁ OCHRANA JAKO PROCES



- Všechny aktivity spojené s archivem lze popsat jako procesy
- Proces je souhrn činností, které se opakují s relativně malými odchylkami
- Procesem je i provádění postupných akcí podle modelu životního cyklu DCC
- Proces můžeme popsat slovně
- Nebo ilustrovat pomocí diagramů BPMN (Business Process Model and Notation)

BPMN

- BPMN je standard pro modelování procesů formou diagramů, které jsou podobné vývojovým diagramům
- Umožňuje vytvořit univerzálně srozumitelný zápis procesu
- Smyslem je usnadnit pochopení komplexních procesů.



STRATEGIE A PRAVIDLA

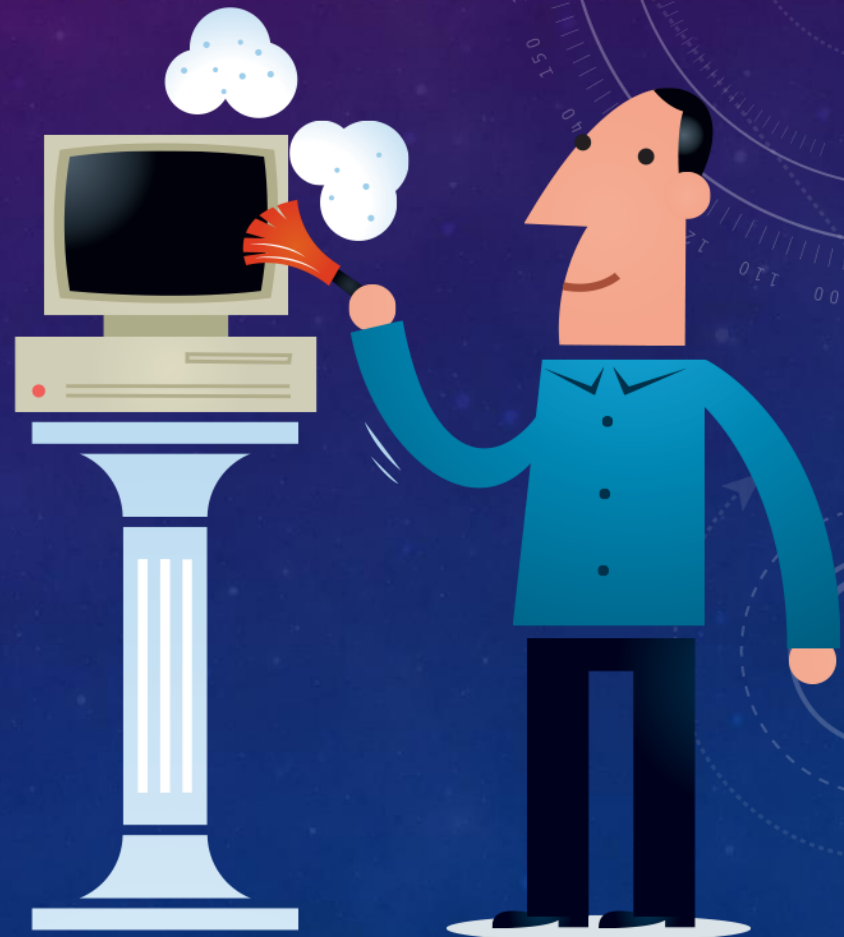


Digitalbevaring.dk

1. Jaké zaujímá archiv místo v rámci instituce?
2. Co do archivu patří a co ne?
3. Kdo jsou stakeholderi?
4. Jaká bude strategie dlouhodobé ochrany?
5. Kdo schvaluje plán archivu?
6. Kdo kontroluje činnost archivu?

PŘÍSTUPY A METODY K OCHRANĚ DAT

- Formátová migrace
- Normalizace a spoléhání na standardy
- Fyzická migrace nebo obnovování nosičů
- Emulace
- Digitální archeologie
- Uchování technologií – „muzeum“ hardware
- Přenos na analogový nosič
- Zapouzdření



STRATEGIE DLOUHODOBÉ OCHRANY



- Ochrana technologií
- Emulace technologií
- Migrace informací
- Ostatní
 - Digitální archeologie
 - Přenos na analogový nosič
 - Zapouzdření

VOLBA FORMÁTŮ A NÁSTROJŮ



Digitalbevaring.dk

- Do archivu mohou vstupovat pouze takové formáty, které umíme zpracovat
 - Formátový registr nám poskytne informace o souborech
 - POWRR Tool Grid nám pomůže vybrat nástroje
- Pro uchovávání používáme pouze takové formáty, které jsou perspektivní z hlediska dlouhodobé ochrany
 - Formátová databáze nám pomůže rozhodnout

JAK SPRÁVNĚ VYBÍRAT?

„Každé rozhodnutí něco uchovat musí někdo zaplatit“

([UNESCO](#))

- Ocenění a výběr musí vycházet z pochopení účelu archivu
- Nelze uchovávat vše a do nekonečna
- Nelze ochránit vše, co by se mělo uchovávat
- Uchovávat musíme víc, než jen data
- Při výběru nám může pomoci aplikace [Decision Tree](#)



Digitalbevaring.dk

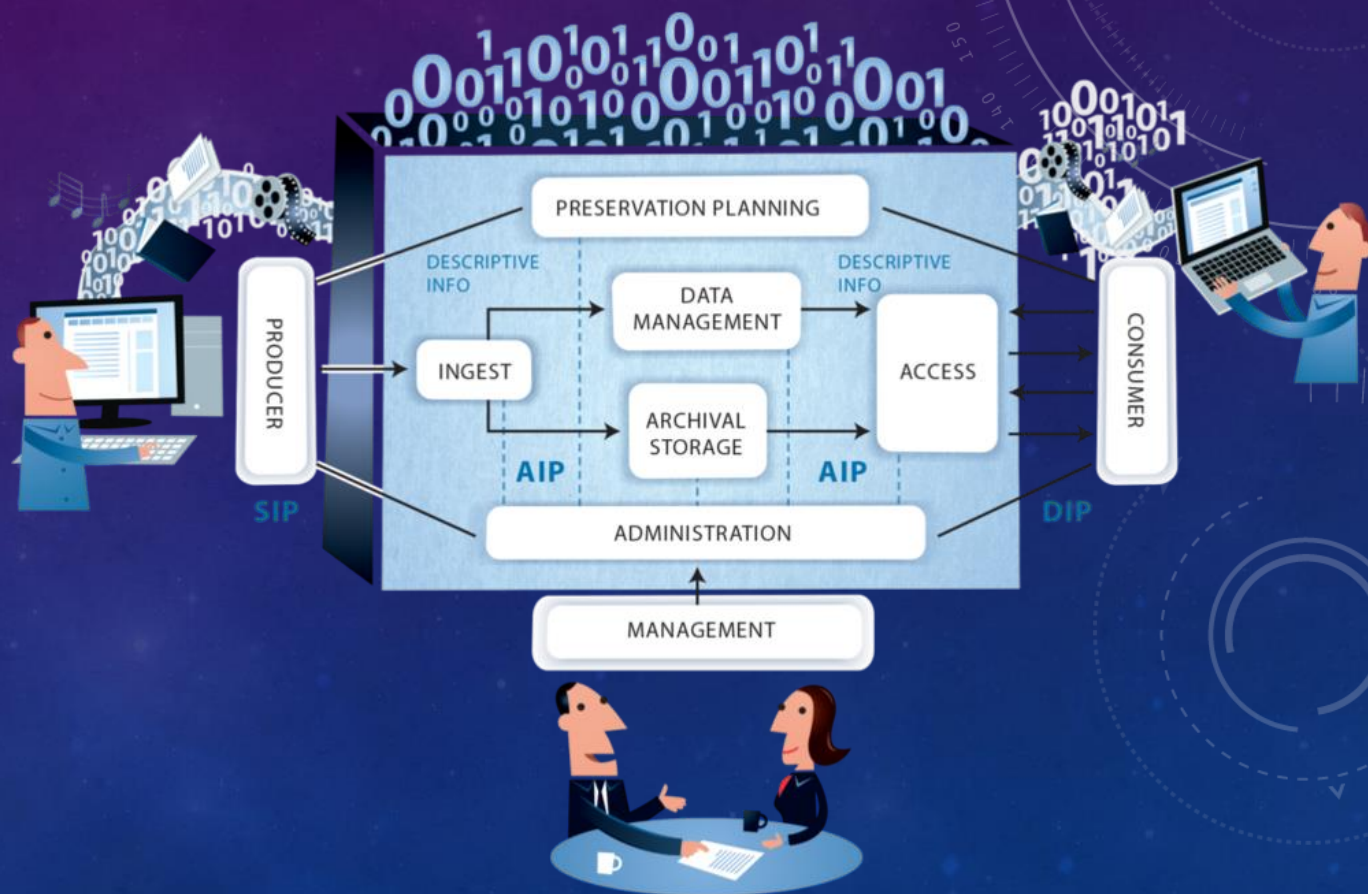
STANDARDS, CERTIFIKACE A NÁSTROJE

- **Standardy** definují závazné nebo doporučené parametry pro
 - Návrh digitálního archivu
 - Technickou realizaci
 - Správu a provoz
 - Kontrolu
- **Nástroje** umožňují vše výše uvedené realizovat
- **Certifikace** slouží jako důkaz, že to děláme správně



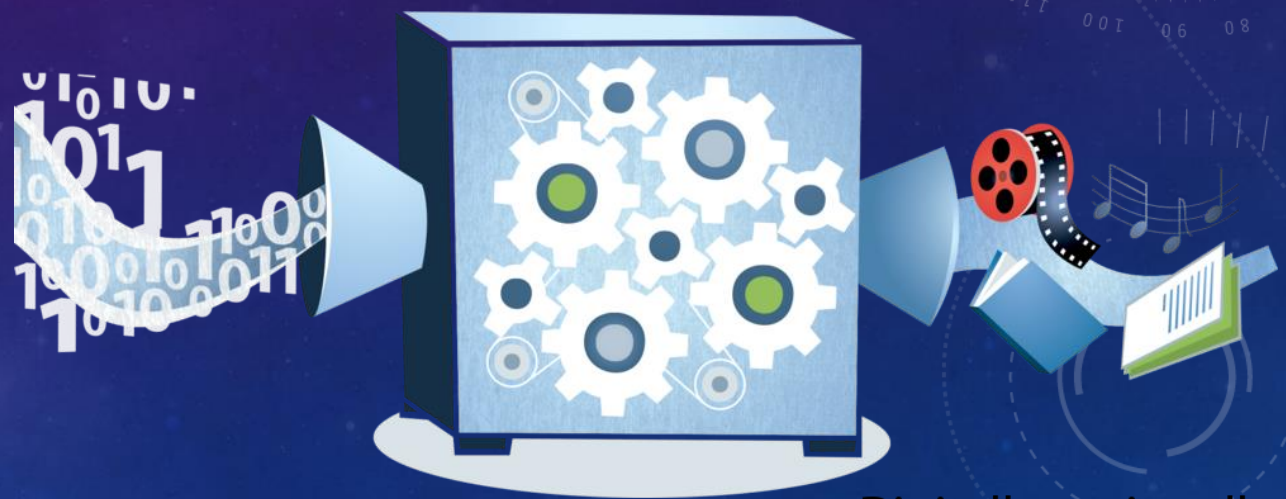
OPEN ARCHIVAL INFORMATION SYSTEM

- Otevřený Archivační Informační Systém
- Terminologický a konceptuální rámec pro:
 - popis a porovnávání architektury
 - volbu strategií LTP
 - porovnání datových modelů digitálních archivů
 - odhadování změny v průběhu času



CO BY NEMĚLO CHYBĚT V SIP BALÍČKU?

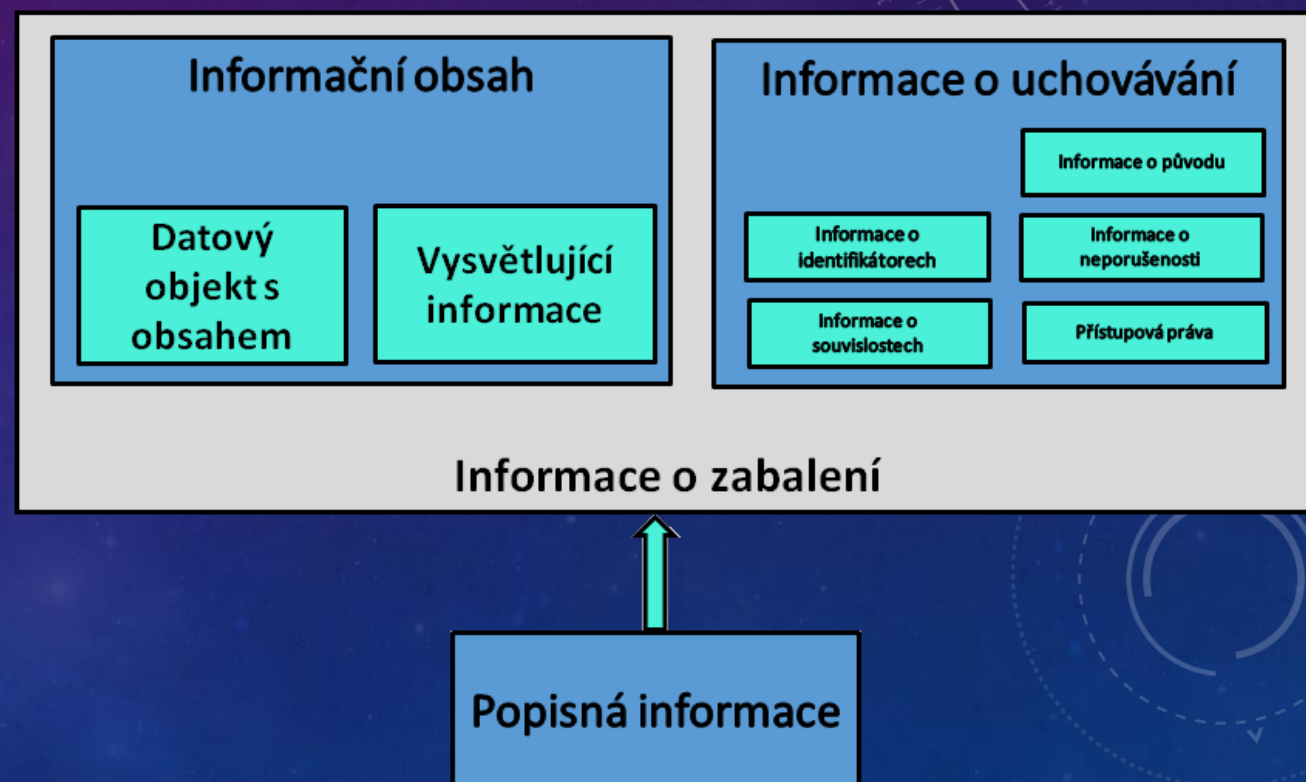
- Data v dohodnutém formátu
- Metadata a dokumentace nezbytné pro dlouhodobé uchování
 - Vysvětlující informace
 - Informace o původu
- Způsob „zabalení“
 - Měl by být vhodný pro náš repozitář
 - Závisí na repozitáři / systému tvůrce dat
 - Bagit (+ METS)
 - Projekt E-Ark (2014 - 2017)



JAK SPRÁVNĚ VYTVOŘIT AIP BALÍČEK?

V závislosti na strategii, plánech a akcích

- Normalizujeme
- Transformujeme (nejen data, ale i metadata)
- Zapouzdřujeme
- Doplnujeme technická metadata
- Doplnujeme o metadata uchování ([PREMIS](#))
- Balíme
 - Soubory (Bagit, [Archivematica](#), [E-Ark](#))
 - Strukturální metadata (METS, [MPEG21](#) [DIDL](#))



JAK MŮŽE VYPADAT DIP BALÍČEK?

- Závisí na cílové skupině a způsobu použití
- Rozlišujeme mezi
 - Přímým výstupem („renderováním“ obsahu)
 - Exportem („náš“ DIP = „cizí“ SIP)



AKCE DLOUHODOBÉ OCHRANY

- Monitorování změn v technologiích i určené skupině
- Doporučení, jak na tyto změny reagovat
- Identifikace „spouštěčů“ pro naplánované akce
- Vybudování znalostní báze



PŘÍKLADY PLÁNOVÁNÍ REPOZITÁŘE

- [Technické řešení Národního digitálního archivu](#)
Informace o architektuře, plánech uchovávání, formátech
- [Archaeology Data Service / Digital Antiquity](#)
Good practices pro dlouhodobou ochranu v oblasti archeologických výzkumů
- [McMaster University Libray Digital Archive](#)
Strategie, implementace a plán digitálního archivu
- [IASA Guidelines](#)
Pokyny pro dlouhodobou ochranu zvukových a audiovizuálních záznamů

DIGITÁLNÍ ARCHIV VS. PROJEKTOVÝ MANAGEMENT

Projekt je jedinečný časově, nákladově a zdrojově omezený proces realizovaný za účelem vytvoření definovaných výstupů (naplnění projektových cílů) v požadované kvalitě a v souladu s platnými standardy a odsouhlasenými požadavky.

(SPPR)

Vybudování digitálního repozitáře je projekt.

STANDARDY PROJEKTOVÉHO ŘÍZENÍ

- IPMA – International Project Management Association
- PMI – Project Management Institute
- PRINCE2 – PROjects IN Controlled Environments 2nd Version

CÍLE A PŘÍNOSY

Cíle projektu

- Stav, do kterého se chceme realizací projektu dostat. Je to výsledek realizace projektu (např. popis služeb, produktů, díla, ...).

Přínosy projektu

- Užitky, které se mohou dostavit používáním dosažených cílů projektu (užíváním stavu projektem dosaženého).

Produkty projektu

- Konkrétní, pojmenovaný výsledek činnosti, etapy, projektu. Produktem může být i provedení služby.

FÁZE ŘÍZENÍ PROJEKTU

Iniciace	• Proč projekt vznikl?
Strategie	• Jaké jsou hranice projektu?
Plánování	• Co, jak, s kým, kdy a za kolik?
Realizace	• Provedení plánu
Předání	• Předání produktu
Vyhodnocení	• Co se povedlo a co ne...
Uzavření	• Přechod do režimu používání

INICIACE

Projektový trojimperativ

Cíl projektu vyjádřený v dimenzích

- Výsledek
- Čas
- Náklady



STRATEGIE

- V praxi je prakticky nemožné dosáhnout ideálního výsledku ve všech osách současně.
- Při iniciaci projektu je nutné rozhodnout, která osa bude určující.



PLÁNOVÁNÍ



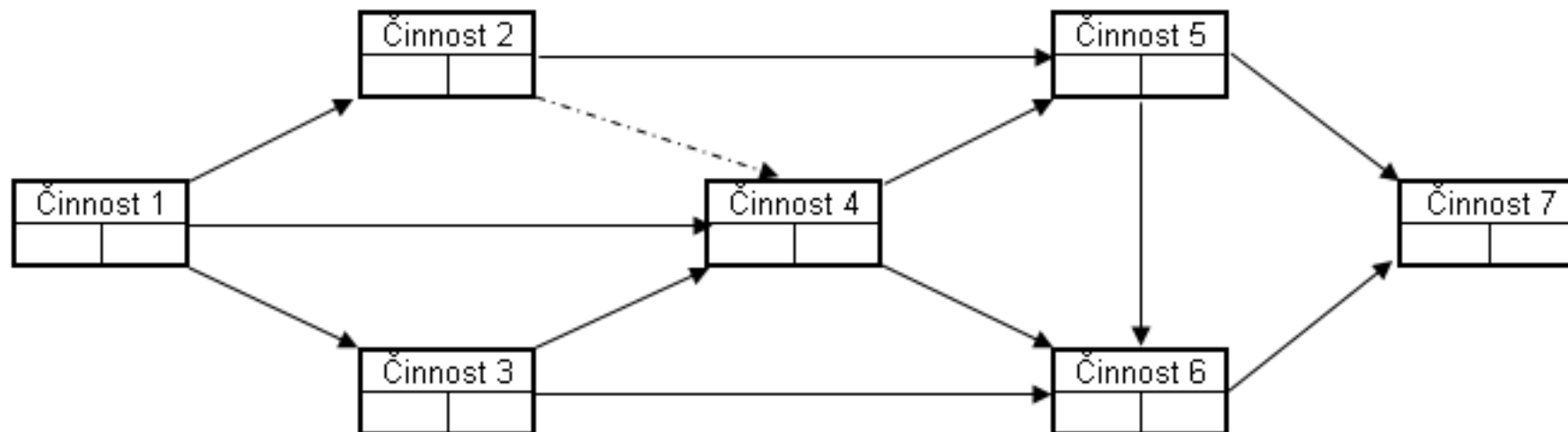
WORK BREAKDOWN STRUCTURE – PLÁN „CO“

- Pořadí není důležité
- Graf definuje pouze produkty



SÍŤOVÝ GRAF - PLÁN „JAK“

- Určuje posloupnost činností a jejich závislosti.



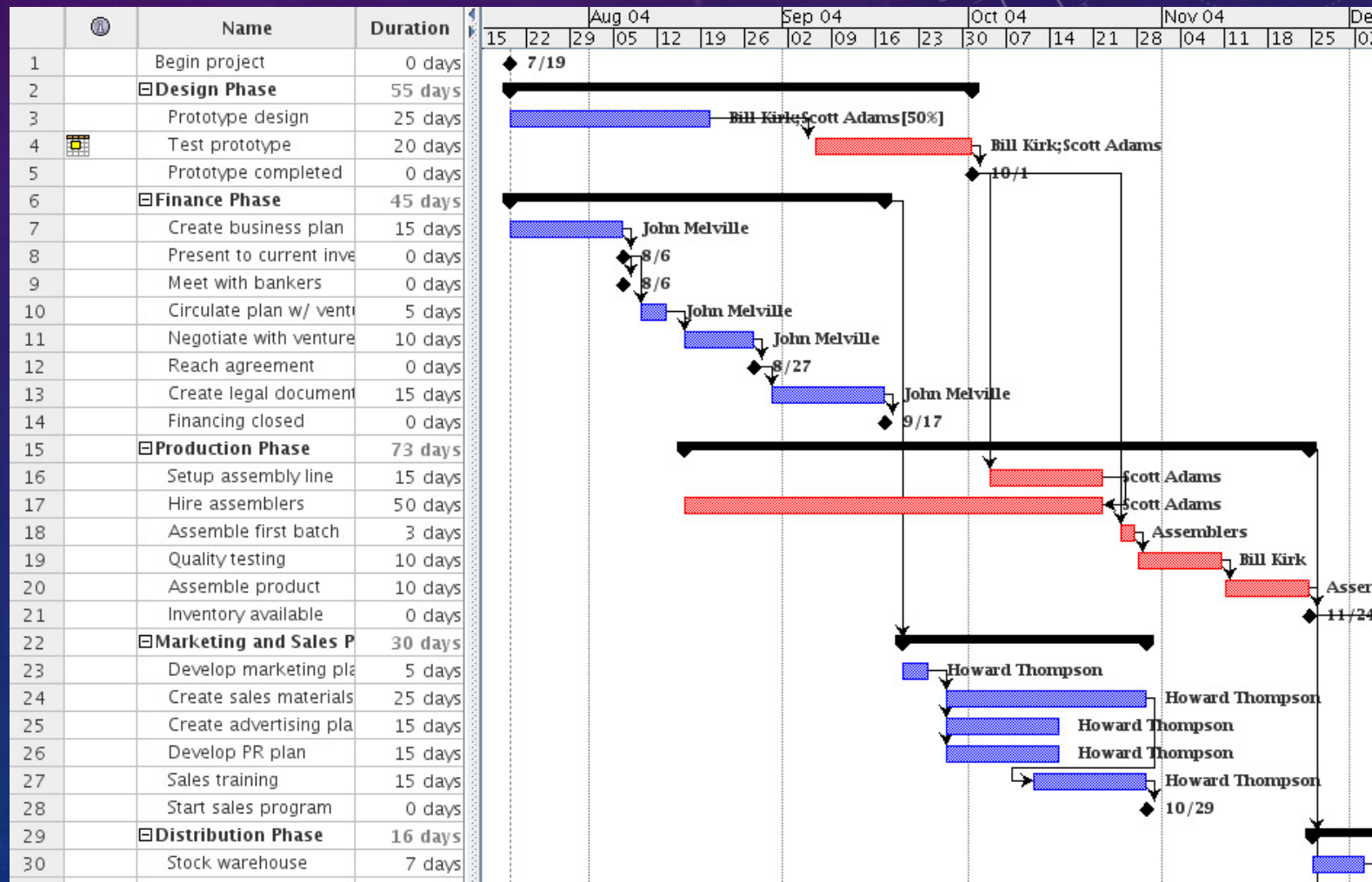
ZDROJE – PLÁN „S KÝM“

- Jaké zdroje budeme potřebovat?
- Jak tyto zdroje obsadíme lidmi?
- Je nutné zvážit zastupitelnost



HARMONOGRAM – PLÁN „KDY“

- Často se používá tzn. Ganttův diagram
- Operuje s časem, závislostmi aktivit a dostupnými zdroji
- Lze využít šablony (obtížně se aktualizují) nebo specializovaný SW – Microsoft Project, Project Libre, Ganttter...)



PLATTER

- Průvodce plánováním důvěryhodného digitálního repozitáře
- Vznikl v rámci projektu Digital Preservation Europe
- Určen zejména pro manažery
- Klíčovým hlediskem je důvěryhodnost

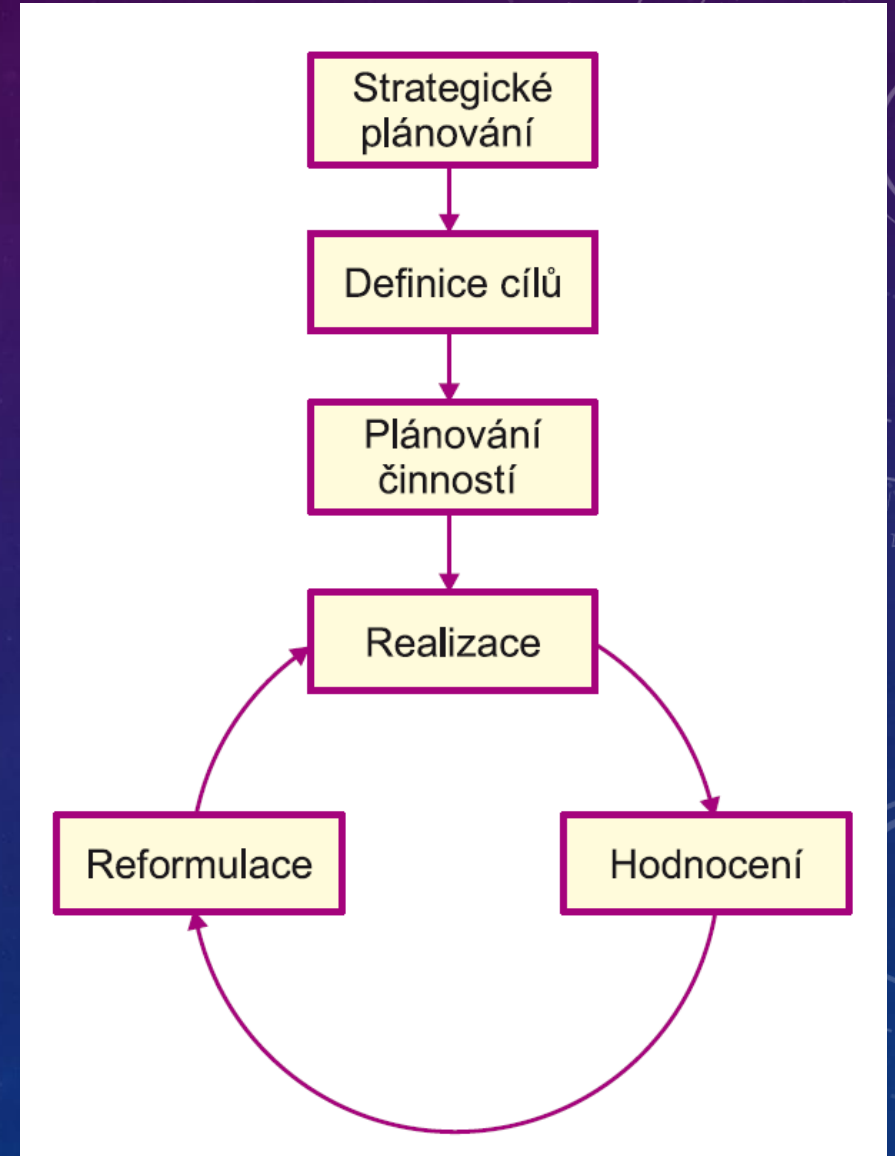
Repozitář je považován za důvěryhodný, pokud lze prokázat jeho schopnost plnit určité funkce a pokud tyto funkce splňují minimální dohodnutá kritéria, jež by měly splňovat všechny „důvěryhodné repozitáře“.

PLATTER – DESATERO DŮVĚRYHODNÉHO REPOZITÁŘE

1. Repozitář se zavazuje k trvalé správě/ochraně digitálních objektů pro definovanou komunitu.
2. Repozitář musí prokázat organizační způsobilost pro tento úkol (tj. vhodné financování, personální zajištění, vhodné procesy).
3. Repozitář dostojí smluvním a právním požadavkům a splňuje své povinnosti v této oblasti.
4. Repozitář má vypracován účelný a účinný strategický rámec.
5. Repozitář získává a zpracovává digitální objekty podle stanovených kritérií, která odpovídají jeho cílům a schopnostem.
6. Repozitář udržuje a zajišťuje dlouhodobou integritu, autenticitu a využitelnost spravovaných digitálních objektů.
7. Repozitář archivuje potřebná metadata o všech akcích, které byly s digitálními objekty v průběhu skladování provedeny.
8. Repozitář naplňuje potřebná kritéria pro zpřístupňování.
9. Repozitář má strategický program pro plánování ochrany.
10. Repozitář má odpovídající technickou infrastrukturu, potřebnou k trvalému udržování a zabezpečení spravovaných digitálních objektů

PLATTER PLÁNOVACÍ CYKLUS

- Formalizovaný soubor kroků pro definici a specifikaci cílů organizace
- Cyklický proces v mnoha částech odpovídající analýze rizik DRAMBORA



PLATTER – PLÁNY STRATEGICKÝCH CÍLŮ

1. Finanční plán,
2. Plán řízení lidských zdrojů,
3. Datový plán,
4. Akviziční plán,
5. Plán zpřístupňování,
6. Plán ochrany,
7. Technický plán,
8. Plán zajištění kontinuity,
9. Krizový plán.

DĚKUJI ZA VAŠI POZORNOST

... a rád zodpovím vaše dotazy

Michal Konečný

jsem@digitalnikurator.cz



POUŽITÉ ZDROJE

- DCC Curation Lifecycle Model. 2015. DCC: because good research needs good data [online]. Digital Curation Centre [cit. 2015-11-03]. Dostupné z: <http://www.dcc.ac.uk/resources/curation-lifecycle-model>
- HUTAŘ, Jan a Marek MELICHAR. 2015. *Dlouhodobá digitální archivace: Přehled a úvod do problematiky*. Výstup projektu LTP-Pilot. Brno. Dostupné také z: <https://drive.google.com/file/d/0BzOLuOh094X8d0REWV9idmEzXzA/>
- HOLDSWORTH, David, Seamus ROSS (ed.) a Michael DAY (ed.). 2007. *Digital Curation Manual: Preservation Strategies for Digital Libraries*. Version 1.0. Edinburgh: Digital Curation Centre. ISSN 1747-1524. Dostupné také z: <http://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/preservation-strategies/preservation-strategies.pdf>
- HUTAŘ, Jan a Marek MELICHAR. Nástroje pro digitální archivaci. *Knihovna: knihovnická revue*, 2015, 26(2), s. 69-82. ISSN 1801-3252. Dostupné také z: <http://knihovnarevue.nkp.cz/aktualni-cislo/knihovny-a-informace/nastroje-pro-digitalni-archivaci>
- Certifikace - LTP Portál. *LTP Portál* [online]. Brno: Moravská zemská knihovna [cit. 2019-03-21]. Dostupné z: <http://ltp-portal.mzk.cz/digital-preservation/certifikace>
- BROWN, Adrian. 2008. *Digital Preservation Guidance Note 1: Selecting file formats for long-term preservation*. Issue 2. The National Archives. Dostupné také z: <http://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf>
- Digital Preservation Handbook, 2019. Digital Preservation Coalition [online]. Digital Preservation Coalition [cit. 2019-10-28]. Dostupné z: <https://www.dpconline.org/handbook>
- LAVOIE, Brian F. a Marek MELICHAR, 2015. Úvod do referenčního modelu otevřeného archivačního informačního systému (OAIS). Výstup projektu LTP-Pilot. Brno. Dostupné také z: <https://drive.google.com/file/d/0BzOLuOh094X8ejZ0NHVOY3VuOTg/>
- ROSENTHAL, Colin, Asger BLEKINGE-RASMUSSEN a Jan HUTAŘ, 2009. Průvodce plánem důvěryhodného digitálního repozitáře (PLATTER). Praha: Národní knihovna České republiky, 51 s. ISBN 9788070505694. Dostupné také z: <http://www.ndk.cz/platter-cz/Platter.pdf>
- PITAŠ, Jaromír, 2012. *NÁRODNÍ STANDARD KOMPETENCÍ PROJEKTOVÉHO ŘÍZENÍ VERZE 3.2: PŘÍLOHA Č.3 VÝKLADOVÝ SLOVNÍK POJMŮ VERZE 3.2*. Brno: Společnost pro projektové řízení, o.s.
- *Řízení projektů v praxi: Klíčové principy projektového řízení*, 2008. Pozorice: Shine consulting.