

Článek 1

Generování výzkumných dat v Centrální laboratoři RTG difrakce a Bio-SAXS

- (1) Výzkumná data v Centrální laboratoři RTG difrakce a Bio-SAXS (dále jen **X-ray CF**) vznikají za účelem experimentálního studia atomární struktury látek pomocí RTG záření
- (2) Za účelem vzniku dat používá X-ray CF dvě experimentální metodiky:
 - Difrakci RTG záření na monokrystalech (dále jen **difrakce**). Difrakční experimenty se provádí na **monokrystalových difraktometrech**
 - Maloúhlý rozptyl RTG záření (dále jen **SAXS**). SAXS experimenty se v X-ray CF provádí pomocí **SAXS kamery**
- (3) Primárním zdrojem výzkumných dat v X-ray CF jsou **detektory RTG záření**, aktuálně detektory typu **HPC** (zkratka z anglického názvu Hybrid Photon Counting detector, <https://doi.org/10.1098/rsta.2018.0241>)

Článek 2

Typy HW úložišť pro výzkumná data používané v X-ray CF a jejich fyzické umístění

- (1) **Framegrabber** je počítač (typicky průmyslové PC provedení RACK server) řídící činnost (=mj. i sběr výzkumných dat pomocí) detektoru RTG záření. Frame grabbery jsou provozovány v RTG laboratoři 115/C12.
- (2) **Řídící počítač** je počítač (typicky PC kategorie workstation) ovládající průběh vědeckého experimentu realizovaného s monokrystalovým difraktometrem anebo SAXS kamerou. Řídící počítače jsou provozovány v RTG laboratoři 115/C12.
- (3) **Datový server** je počítač (typicky výkonné PC kategorie RACK server s velkým diskovým úložištěm RAID, Redundant Array of Independent Disks) sloužící jako dlouhodobé datové úložiště pro výzkumná data skupiny (event. více skupin) uživatelů. Datový server je možné považovat za **repozitář** dle definice čl. 3.4 Opatření ředitele CEITECu MU 4/2020. Datový server X-ray CF je provozován v serverovně CEITECu MU, místnost C35/2S027
- (4) **Osobní počítač** je počítač (typicky PC kategorie workstation nebo notebook) používaný operátory X-ray CF k analýze difrakčních nebo SAXS dat. Osobní počítače X-ray CF jsou provozovány v pracovně 3.13/C02

Článek 3

Uživatelé a skupiny uživatelů X-ray CF

- (1) Za **uživatelskou skupinu** (dále jen **UG, user group**) je v tomto dokumentu označována skupina uživatelů X-ray CF s jedním společným vedoucím (dále jen **PI, principal investigator**).
- (2) K příslušnosti do UG se uživatel obvykle přihlašuje při své registraci do Rezervačního systému X-ray CF uvedením jména PI (typicky svého vedoucího, školitele anebo supervizora). Pravdivost přihlášení se uživatele k UG následně potvrdí e-mailová korespondence X-ray CF s PI.
- (3) Jako alternativu k postupu dle 3.2 může PI sestavit (a zaslat e-mailem do X-ray CF) seznam uživatelů X-ray CF, kteří jsou/budou členy jím vedené UG
- (4) Délka přiřazení uživatele do UG je časově omezena na maximálně 21 následujících měsíců (detaily viz článek 8.6)

- (5) Pokud PI neprojeví explicitní nesouhlas s prodloužením příslušnosti uživatele X-ray CF ve své UG, bude vždy po každém prvním uživatelském difrakčním či SAXS experimentu v kalendářním roce doba příslušnosti uživatele do UG prodlužována o dalších 12 měsíců.
- (6) Pokud PI vedoucí UG projeví explicitní nesouhlas s další příslušností uživatele X-ray CF ve své UG, bude uživatel z UG bez zbytečného odkladu vyřazen.

Článek 4

Toky výzkumných dat mezi počítači X-ray CF a mezi X-ray CF a uživatelem/UG

- (1) **Primární výzkumná data** v X-ray CF jsou generována fyzikálním dějem spojeným s průchodem fotonů RTG záření detektorem RTG záření
- (2) Signály z detektoru RTG záření jsou framegrabberem ukládány do počítačových souborů, tzv. **dataframes**. Dataframes je možné považovat za nezpracovaná (=raw) výzkumná data
- (3) Dataframes jsou v průběhu difrakčního nebo SAXS experimentu přenášeny z framegrabberu do řídicího počítače, kde je prováděna (automaticky, anebo na základě příkazů zadávaných operátorem X-ray CF) jejich primární analýza.
- (4) V průběhu difrakčního experimentu se na řídicím počítači provádí **redukce difrakčních dat** (z dataframes = binární/RAW formy do formy integrálních difrakčních intenzit = ASCII/HKL souborů).
- (5) Po ukončení SAXS experimentu se na řídicím počítači provádí **redukce SAXS dat** (z dataframes = binární/RAW formy do formy ASCII/DAT souborů)
- (6) Veškerá data vygenerovaná na řídicím počítači v průběhu difrakčního nebo SAXS experimentu jsou kopírována do zvláštního archivního podadresáře/podadresářů na datovém serveru X-ray CF
- (7) Má-li UG zřízen vlastní repozitář dat a udělí-li operátorům X-ray CF do tohoto repozitáře přístupová práva pro zápis, mohou být veškerá výzkumná data získávaná v průběhu difrakčního nebo SAXS experimentu s uživatelským vzorkem po ukončení experimentu kopírována rovněž do vlastního repozitáře UG.
- (8) Nemá-li UG vlastní repozitář dat s přístupovými právy pro zápis pro operátory X-ray CF, mohou být veškerá výzkumná data získaná v průběhu difrakčního nebo SAXS experimentu s uživatelským vzorkem po ukončení experimentu kopírována rovněž do speciálního podadresáře datového serveru X-ray CF přístupného uživatelům z UG (viz i článek 8.6)
- (9) Povinnou specifikaci místa, kam mají být nahrávána výzkumná data získaná v průběhu difrakčního nebo SAXS experimentu s uživatelským vzorkem (= výběr z variant postupů specifikovaných ve člancích 4.6, 4.7 anebo 4.8) uživatel obvykle provádí při zakládání požadavku na experiment v X-ray CF ve formuláři Rezervačního systému X-ray CF.
- (10) Alternativou nadřazenou k postupu dle 4.9 je volba jedné z variant (= možností dle článků 4.6, 4.7 anebo 4.8) učiněná PI v e-mailu odeslaném do X-ray CF.
- (11) **Analýza redukováných difrakčních dat** (= řešení tzv. fázového problému a následné tzv. zpřesňování strukturního modelu) se provádí na osobním počítači operátora X-ray CF.
- (12) Informace o výsledku difrakčního experimentu a/nebo výsledek následné analýzy difrakčních dat (= např. tzv. předběžný strukturní model ve formě CIF, Crystallographic Information File, <https://www.iucr.org/resources/cif>) zasílá operátor CF uživateli e-mailem.
- (13) Komunikace mezi uživatelem, PI a operátorem X-ray CF spojená s finální analýzou předběžného strukturního modelu se vede pomocí e-mailu.
- (14) **Analýza redukováných SAXS dat** se provádí na osobním počítači operátora X-ray CF.
- (15) Informace o výsledku SAXS experimentu (= textový soubor s redukovánými SAXS daty) a/nebo výsledky následných analýz SAXS dat (=např. simulovaná SAXS data ve formě textových souborů, event. strukturní modely ve formátu PDB, <http://mmcif.wwpdb.org/>) zasílá operátor CF uživateli e-mailem. Komunikace mezi uživatelem, PI a operátorem X-ray CF spojená s analýzou SAXS dat se vede pomocí e-mailu

Článek 5

Dostupnost výzkumných dat a HW úložišť X-ray CF

- (1) X-ray CF neumožňuje uživatelům X-ray CF fyzický přístup k HW úložištím svých výzkumných dat
- (2) Vedle přímého přístupu k počítačovému HW používá X-ray CF i vzdálený přístup k HW úložištím dat pomocí LAN/WAN a protokolu TCP/IP
- (3) Úložiště kategorie frame grabber (viz čl. 2.1) obvykle nemá ani klávesnici, ani monitor. Jediný možný přístup k souborům na Frame grabberu je vzdálený přístup z řídicího počítače (viz čl. 2.2) přes privátní LAN síť (maska=192.168.x.x) a aplikaci Remote desktop OS Windows spuštěnou na řídicím počítači
- (4) Pro vzdálený přístup do počítačů s OS Windows přes WAN se v X-ray CF používá SW TeamViewer (<https://www.teamviewer.com>).
- (5) Vzdálený přístup do datového serveru X-ray CF i přístupová práva k souborům na datového serveru uloženým jsou řízena prostředky OS Linux (aktuálně Debian verze 4.9.0-14-amd64)
- (6) Přístup do datového serveru a do osobních počítačů operátorů X-ray CF je chráněn heslem/hesly.

Článek 6

Velikosti datových souborů a datové kvóty

- (1) Na žádných počítačích provozovaných X-ray CF nejsou nastavovány žádné diskové kvóty
- (2) Typická velikost souboru obsahujícího jeden dataframe získaný metodikou difrakce (tj. jednoho souboru *.rodhypix) je cca 300 kB
- (3) Obvyklý počet dataframes změřených při jednom standardním difrakčním experimentu se pohybuje od několika stovek po (nízké) desítky tisíc
- (4) Velikost podadresáře s výzkumnými daty vygenerovaného na řídicím počítači v průběhu jednoho difrakčního experimentu a jeho vyhodnocování se pohybuje v intervalu od několika set MB po (nízké) jednotky GB
- (5) V kroku redukce dat (z formy dataframes do formy integrálních difrakčních intenzit, viz článek 4.4) se počet souborů s výstupy difrakčního experimentu snižuje cca 100x.
- (6) Výstupem z úspěšně ukončeného difrakčního experimentu po kroku redukce dat je necelá desítka souborů o souhrnné velikosti řádu (nízkých) jednotek MB, tyto soubory operátor X-ray CF následně zpracovává v průběhu analýzy difrakčních dat (viz článek 4.9)
- (7) Typická velikost CIF souboru obsahujícího strukturní model získaný analýzou difrakčních dat (viz článek 4.11 a 4.12) je řádu (nízkých) jednotek MB
- (8) Typická velikost souboru obsahujícího jeden dataframe získaný metodikou SAXS (tj. jednoho souboru *.img) je cca 1180 kB
- (9) Počet dataframes změřených při standardním SAXS experimentu se pohybuje od několika jednotek do několika desítek
- (10) Obvyklá velikost podadresáře se soubory vygenerovanými na řídicím počítači při jednom SAXS experimentu se pohybuje v intervalu od několika jednotek po (nízké) desítky MB
- (11) V kroku redukce dat (z formy dataframes do formy ASCII, resp. DAT/TXT) se souhrnná velikost souborů obsahujících výstupy z úspěšně provedeného SAXS snižuje cca 25x.
- (12) Výstupem z úspěšně ukončeného SAXS experimentu po kroku redukce dat (= vstup pro následné analýzy SAXS dat dle článku 4.14) je řádově několik jednotek textových souborů o souhrnné velikosti řádově desítky kB.
- (13) Výstupem z analýzy redukovaných SAXS dat jsou soubory se simulovanými SAXS daty, popř. strukturní modely ve formátu PDB, o typické velikosti od jednotek po nízké desítky MB

Doby skladování dat na HW X-ray CF

- (1) Experimentální data odpovídající jednomu dataframe jsou framegrabberem akumulovány během tzv. **expoziční doby**. Pro difrakční experimenty X-ray CF se používají expoziční doby v rozsahu jednotek až desítek sekund, pro SAXS experimenty expoziční doby délky stovek až tisíců sekund.
- (2) Doba uložení individuálního dataframe na framegrabberu je řádu max. jednotek sekund (=jen tak dlouho, než je po ukončení expozice soubor s dataframe úspěšně přenesen do řídicího počítače).
- (3) Experimentální data generovaná na řídicích počítačích v průběhu jednotlivých difrakčních/SAXS experimentů se z řídicích počítačů programově neodstraňují (=nemažou se), doba jejich skladování je typicky limitovaná jen fyzickou životností datového nosiče (=HDD/SSD) či jiných HW/SW komponent řídicího počítače (=RAW aj. data jsou na řídicím počítači skladována typicky několik let)
- (4) Doba skladování experimentálních difrakčních/SAXS dat na datovém serveru není omezena časově, ale velikostí disponibilního místa na jeho diskových úložištích. Kapacita, resp. velikost HW aktuálně provozovaného jako repozitář X-ray CF umožňuje skladovat kompletní experimentální data získaná na instrumentaci X-ray CF přinejmenším 8 let.
- (5) Doba skladování výsledků získaných analýzou experimentálních difrakčních a SAXS dat operátory X-ray CF na osobních počítačích není časově nijak omezena. Na HW X-ray CF jsou k dispozici strukturní výsledky získané pomocí difrakční instrumentace MU počínajíc rokem 1999 a výstupy SAXS experimentů prováděných na instrumentaci CEITECu MU počínajíc rokem 2013.

Archivace a indexace dat

- (1) Data získávaná v průběhu difrakčních nebo SAXS experimentů jsou zapisována do unikátních podadresářů vytvářených na discích řídicích počítačů
- (2) **Název difrakčního či SAXS experimentu** a současně i název podadresáře s daty vzniklými v průběhu difrakčního nebo SAXS experimentu slouží jako tzv. **primární klíč pro jednoznačné přiřazení** experimentu a **výzkumných dat při experimentu vytvořených k uživateli X-ray CF**, k UG a/nebo k záznamu s požadavkem na experiment v Rezervačním systému X-ray CF.
- (3) Názvy podadresářů s výsledky experimentů X-ray CF mají povinnou strukturu tvaru:
RokZahájeníExperimentu-MěsícZahájeníExperimentu-DenZahájeníExperimentu-JménoUživatele-NázevVzorkuPoužitýUživateléVRezervačnímSystému
- (4) Podadresáře s daty s názvem specifikovaným v článku 8.3 jsou archivovány na datovém serveru X-ray CF, eventuálně i na jiných repozitářích (viz i články 4.6-4.8).
- (5) Výsledky difrakčních nebo SAXS experimentů budou kopírovány/archivovány tak, jak rozhodne uživatel (event. PI řídící jeho UG) při zadávání požadavku na experiment v X-ray CF (viz i články 4.6-4.10).
- (6) Nemá-li UG vlastní repozitář, může být uživateli z takové UG vytvořen speciální účet na datovém serveru X-ray CF umožňující přístup do podadresáře dle článku 4.8. Primární platnost uživatelského účtu na datovém serveru X-ray CF bude nastavena do 30.9. roku následujícího po roce vytvoření tohoto účtu, a může být každoročně prodlužována dle postupu uvedeného v článku 3.5

Mazání dat

- (1) Vznik i zánik souborů s dataframes na framegrabberu plně kontroluje SW řídicího počítače

- (2) Experimentální data ukládaná na řídicím počítači se programově nemazou, limitací jejich dostupnosti na řídicím počítači je jen životnost HW tohoto počítače.
- (3) Experimentální data ukládaná na dataserveru se programově nemazou, limitací pro jejich dostupnost na diskovém úložišti (resp. jejich smazání z něj) je jen dostatečně velké prázdné místo pro výsledky nových difrakčních a SAXS experimentů
- (4) Výsledky analýz difrakčních a SAXS experimentů ukládané na osobních počítačích operátorů X-ray CF se nikdy nemazou

Článek 10

Zacházení s citlivými daty

- (1) X-ray CF neshromažďuje žádná citlivá data ani o svých uživateliích, ani o vzorcích v X-ray CF studovaných
- (2) Při práci instrumentace X-ray CF nevznikají žádná citlivá data

Článek 11

Zodpovědné osoby a jimi spravovaný HW/SW

- (1) HW a SW pro difrakční metodiky: Jaromír Marek
- (2) HW a SW pro SAXS metodiky: Tomáš Klumpler
- (3) Datový server X-ray CF: Marek Nečas