

Gestione dati della validazione clinica di INSIDE

12/12/2017

Meeting INSIDE

Giancarlo Sportelli



Validazione clinica di INSIDE

- Stato dello sviluppo di INSIDE
 - 2015 -> primi test con fasci monoenergetici
 - 2016 -> primo test clinico
 - 2017 -> primo test combinato PET/tracker
- È necessario ora dimostrarne l'efficacia clinica
- Le applicazioni di maggior interesse per i clinici sono:
 - Monitoraggio della risposta precoce di malattia in corso di trattamento
 - Monitoraggio sistematico di modifiche anatomiche d'impatto sfavorevole sulla distribuzione di dose
- Per il 2018 è programmata una serie di prese dati "longitudinali" in condizioni di trattamento clinico convenzionale

Caratteristiche dello studio longitudinale

- Non è il primo
 - “Clinical implementation and range evaluation of in vivo PET dosimetry for particle irradiation in patients with primary glioma”, S.P. Nischwitz et al., Radiotherapy and Oncology, 2015 (MIRANDA study)
- Ma presenta delle unicità:
 - Acquisizione dati in-beam
 - Cross-validazione con il particle tracker
 - Possibilità di introdurre imaging TOF
- Impegno notevole per l'acquisizione dati
 - Necessità di personale sia tecnico che clinico durante l'acquisizione PET sui 20 pazienti
- Si prevede la generazione di una considerevole quantità di dati

Data taking e data sharing

- Necessità di un **protocollo prestabilito** per ridurre la probabilità di errore e/o incertezza durante l'acquisizione dati e l'analisi a posteriori
- L'operatore tecnico dovrà sapere in anticipo cosa annotare del setup sperimentale e dovrà avere a disposizione un **software di supporto** per l'esecuzione del protocollo di acquisizione
- Il software dovrà permettere (richiedere?) l'archiviazione di tutte le informazioni (**metadati**) che possano essere utili in fase di analisi dati
- Dati e metadati possono essere **condivisi da remoto** con membri della collaborazione e, successivamente, con ricercatori abilitati (proposta)

Software di supporto per INSIDE?

MMND&ITRO2016

IOP Publishing

IOP Conf. Series: Journal of Physics: Conf. Series 777 (2017) 012021

doi:10.1088/1742-6596/777/1/012021

A dedicated software application for treatment verification with off-line PET/CT imaging at the Heidelberg Ion Beam Therapy Center

**W Chen², J Bauer^{1,2}, C Kurz¹⁻³, T Tessonier^{1,2}, J Handrack^{1,2}, T Haberer^{1,2},
J Debus^{1,2}, K Parodi^{1,3}**

¹Heidelberg Ion Beam Therapy Center (HIT), Im Neuenheimer Feld 450, Heidelberg 69221, Germany

²Department of Radiation Oncology, Heidelberg University Hospital, Im Neuenheimer Feld 400, Heidelberg 69221, Germany

³Ludwig-Maximilians-Universität München (LMU Munich), Department of Experimental Physics - Medical Physics, Am Coulombwall 1 (R127), Garching b. München, 85748, Germany

E-mail: wenjing.chen@med.uni-heidelberg.de

Abstract. We present the workflow of the offline-PET based range verification method used at the Heidelberg Ion Beam Therapy Center, detailing the functionalities of an in-house developed software application, SimInterface14, with which range analysis is performed. Moreover, we introduce the design of a decision support system assessing uncertainties and facilitating physicians in decisions making for plan adaptation.

Gestione dei dati

- Lo sviluppo di un software di supporto alla raccolta e gestione dei dati richiede che siano completati e definiti una serie di aspetti tecnologico/gestionali:
 - Protocollo di raccolta dati
 - Architettura del sistema informativo
 - Tecnologia del database
 - Schema dati / modello relazionale
 - Applicazione d'utilizzo (CRUD) e interfaccia utente
 - Ethics, privacy, regulations
 - Controllo dei diritti d'accesso

Protocollo di raccolta dati

> prima dell'acquisizione

- Stato hw
 - Stato fascio
 - Impostazioni (energia, dose, durata)
 - Linea utilizzata
 - Stato del gantry
 - Tipo di collimazione
 - Dati di calibrazione
 - Tempo
 - Energia
 - Geometria sistema
 - Posizione delle teste PET
 - Posizione del tracker
- Stato sw
 - Versioni dei sw
 - Impostazioni sw di acquisizione
 - Correzioni applicate
- Target irraggiamento
 - Tipo di target (fantoccio/paziente)
 - Riferimenti clinici
- Persone coinvolte
 - Ricercatori
 - Medici
 - Tecnici
- Ruoli
 - Controllo fascio
 - Controllo PET
 - Controllo Tracker
 - Gestione del target

Protocollo di raccolta dati

> durante l'acquisizione

> dopo l'acquisizione

- Dati PET

- Log di acquisizione
- Event rates
- Ricostruzioni online
- Dati di acquisizione

- Dati derivati

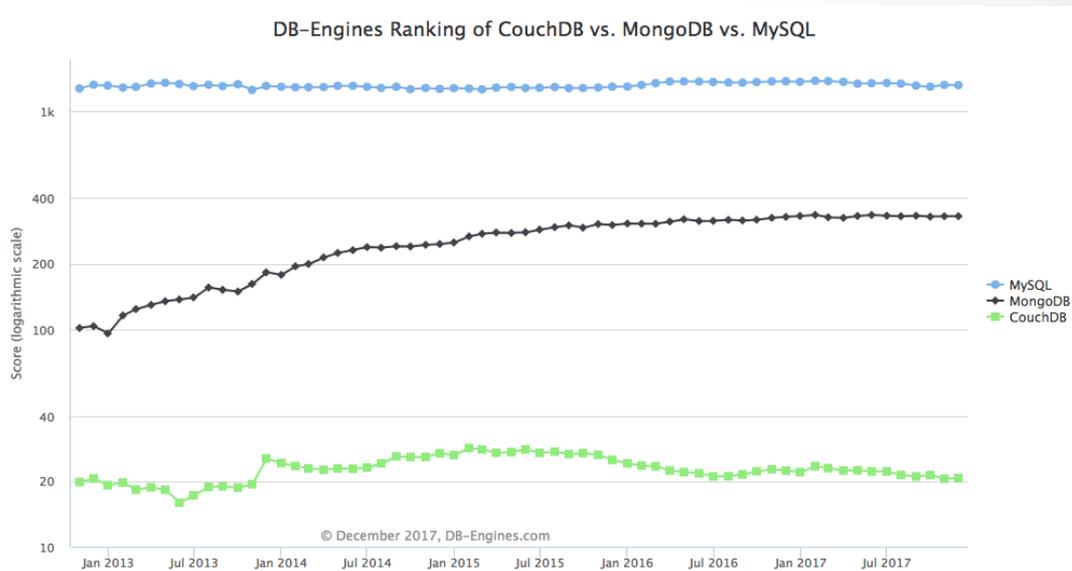
- Simulazioni
- Ricostruzioni
- Analisi dati
- Filtraggio dati

- Dati trattamento

- Monitor energia
- Monitor dose
- Monitor pencil beam
- Distribuzione temporanea spill

Tecnologia del database

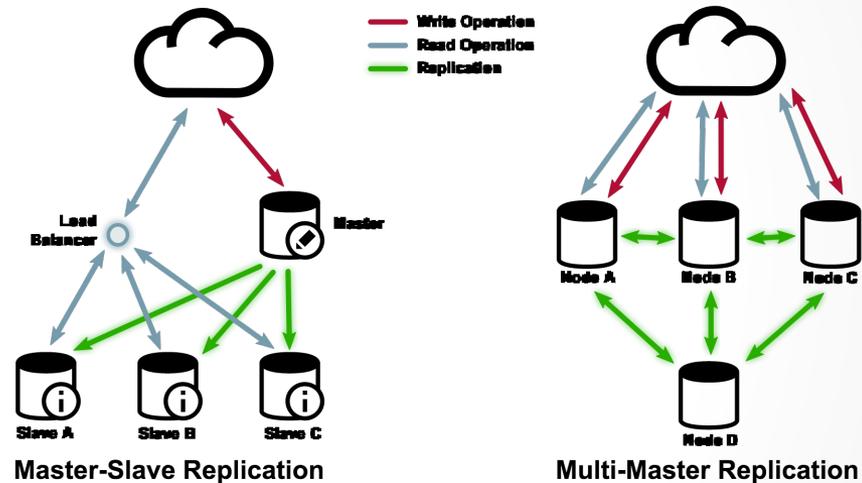
- Esistono diversi tipi di database, divisi in due grandi gruppi: strutturati (SQL) e non strutturati (NoSQL/document oriented)
- Vantaggi dei database strutturati
 - Più consolidati nel tempo (le ditte li conoscono meglio)
 - Assicurano ACID compliancy (atomicity, consistency, isolation, durability)
- Vantaggi dei database non strutturati
 - Più flessibili
 - Prestazioni migliori
 - Big data friendly
 - Sviluppo rapido
- Il tipo di tecnologia del database può influenzare pesantemente la scelta dell'architettura del sistema informativo e i costi di implementazione



Architettura del sistema informativo

- Il sistema informativo può avere diversi tipi di architettura:

- Local only
- Remote only
- Master-Slave (tipo svn)
- Master-Master (tipo git)



- Scelta preferibile è un DB engine capace di replicazione master-slave o master-master
- Di conseguenza si avrà un server locale durante l'acquisizione dati e uno remoto sul cloud, entrambi sincronizzati regolarmente

Schema dati / modello relazionale

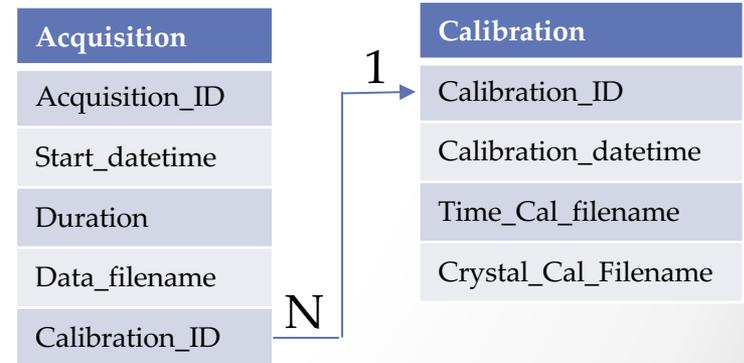
- Qualunque sia il tipo di DB, si definirà uno schema dati (NoSQL) o un modello relazionale (SQL)

Unstructured NoSQL (dictionaries)

```
// One-to-few
> db.acquisition.findOne()
{
  start_datetime: '2017-12-12 15:30 CET',
  duration: '3600',
  data_filename: '/storage/acquisitions/pet_20171212_1530.dat',
  calibration : {
    datetime: '2017-12-11 09:30 CET',
    time_cal_filename: '/storage/calibrations/timecal_20171211_001.dat',
    crystal_cal_filename: '/storage/calibrations/cycal_20171211_001.dat',
  }
}
```

```
// One-to-many
> db.calibration.findOne()
{
  _id : ObjectID('0001'),
  datetime: '2017-12-11 09:30 CET',
  time_cal_filename: '/storage/calibrations/timecal_20171211_001.dat',
  crystal_cal_filename: '/storage/calibrations/cycal_20171211_001.dat',
}
> db.acquisition.findOne()
{
  start_datetime: '2017-12-12 15:30 CET',
  duration: '3600',
  data_filename: '/storage/acquisitions/pet_20171212_1530.dat',
  calibration : ObjectID('0001')
}
```

Structured SQL (tables)

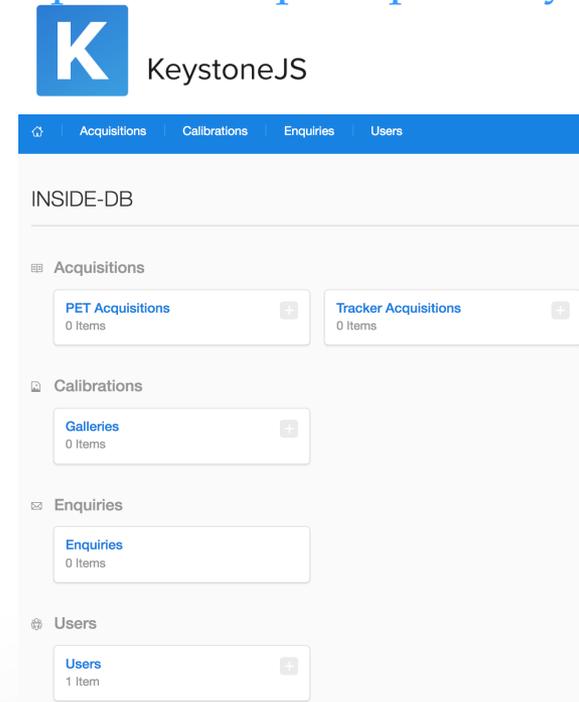
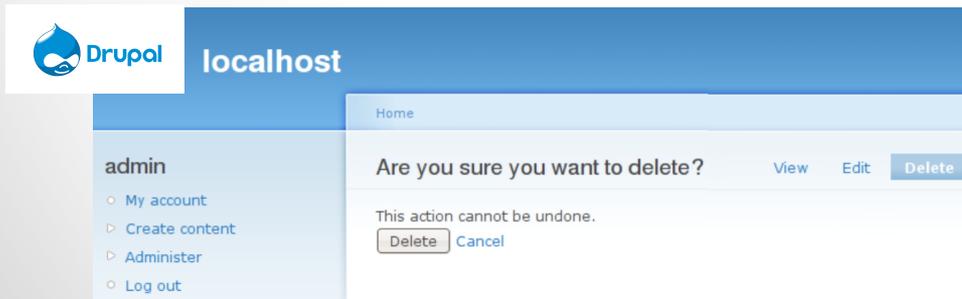


Applicazione d'utilizzo e interfaccia utente

- Per la raccolta, modifica e utilizzo dei dati nel database si svilupperà un'applicazione CRUD (create, read, update, delete) con interfaccia web.

Example mockup: <http://bit.ly/2jBiY3I>

- Due framework candidati:
 - Drupal
 - Keystone.js



Ethics, privacy, regulations

- La raccolta dati dovrà essere adeguata secondo le norme vigenti in materie di etica e privacy
- Parte dei dati dovrà essere anonimizzata:
 - Soggetti identificati tramite un codice alfanumerico generato automaticamente in fase di uploading
 - Solo il detentore del dato che effettua il caricamento è in grado di risalire dal codice al soggetto
 - L'anonimizzazione delle immagini viene richiesta a chi carica il dato
 - Un ulteriore processo automatico di controllo verifica l'esistenza, ed eventualmente cancella informazioni riguardanti il nome del paziente, la sua data di nascita, ecc.

- Know-how dal progetto Arianna

<https://arianna.pi.infn.it/en>



Controllo accesso e gestione dei diritti di utilizzo

- Accesso tramite username e password
- Indirizzo sito certificato
- Tutte le comunicazioni cifrate con protocollo HTTPS
- Implementazione nella sala calcolo INFN-Pisa, su piattaforma appartenente alla rete GARR (Gruppo per l'Armonizzazione della Rete della Ricerca), sottoposto al controllo di GARR-CERT:
 1. verifica tentativi di accesso fraudolento e avvisa immediatamente i responsabili del centro di calcolo
 2. verifica costantemente le configurazioni di tutte le macchine che afferiscono alla rete e segnala eventuali vulnerabilità (porte aperte, programmi installati con falle di sicurezza, scorrette configurazioni)
- Audit interno INFN
- Sistema granulare di gestione dei ruoli e dei diritti d'accesso

Timeline di sviluppo

- Dicembre 2017: survey ditte di sviluppo SW (tecnologie disponibili e costi)
- Gennaio 2018: definizione specifiche di sviluppo e inizio lavori
- Marzo 2018: produzione primo prototipo software
- Aprile 2018: test e validazione del software

Conclusioni

- Si propone lo sviluppo di un nuovo framework software di gestione dei dati acquisiti con INSIDE
- Il software sarà sviluppato per supportare l'acquisizione, l'analisi e, a lungo termine, la produzione di feedback clinician-oriented
- Si useranno tecnologie di gestione dati e metadati web-based, e database distribuiti
- La protezione dei dati, il rispetto delle normative vigenti sulla privacy e il controllo degli accessi verranno affrontati sulla base del lavoro già fatto per il progetto Arianna