

# POUŽITÍ DOTAZOVACÍHO JAZYKU PQF V INTERPI

INTERPI – Interoperabilita v paměťových institucích

Program aplikovaného výzkumu a vývoje národní  
kulturní identity (NAKI)  
(DF11P010VV023)

Zpracovali:

Jana Šubová (Cosmotron Bohemia, s. r. o.)



## INTERPI

aktualizované 30. 3. 2017

## Struktura PQF dotazu

Dotazovací jazyk PQF je pro potřeby INTERPI implementován v následujícím rozsahu:

- atribut 1 (use - vyhledávací kritérium),
- atribut 4 (fráze),
- atribut 5 (rozšíření),
- atribut 6 (úplnost - completeness).

Vyhledávací termín je vždy uzavřen v uvozovkách a nejsou rozlišována malá/velká písmena.

## Atribut 1 (vyhledávací kritérium)

Pro INTERPI jsou využívány převážně proprietární use atributy, které zohledňují specifika datové struktury. Atributy mohou být doplňovány nebo jejich obsah může být mírně modifikován na základě potřeb uživatelů.

Přehled implementovaných vyhledávacích kritérií:

- **2050** - stav záznamu - obsahuje kód pro stav záznamu vzhledem na jeho životný cyklus:
  - nr - rozpracovaný návrh
  - nd - dokončený návrh
  - rr - rozpracovaná revize
  - rd - dokončená revize
  - sz - schválený záznam
  - zz - zrušený záznam
  - ir - bez revize INTERPI (záznamy získané importem z jiných systémů)
- **2051** - třída - obsahuje jednoznačný kód pro označení třídy entit:
  - o - třída osoba/bytost
  - r - třída rod/rodina
  - k - třída korporace
  - g - třída geografický objekt
  - u - třída událost
  - d - třída dílo/výtvar
  - p - třída obecný pojem
- **2054** - preferované označení entity - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla a bez ohledu na zařazení do třídy - při použití jako kompletního termínu obsahuje hlavní, vedlejší část označení a doplňky
- **2055** - všechna označení entity - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení a zařazení do třídy - použití jako kompletního termínu viz atribut 2054
- **2056** - všechna označení entity třídy osoba/bytost - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054

- **2057** - preferované označení entit třídy osoba/bytost - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2058** - všechna označení entity třídy rod/rodina - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054
- **2059** - preferované označení entit třídy rod/rodina - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2060** - všechna označení entity třídy korporace - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054
- **2061** - preferované označení entit třídy korporace - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2062** - všechna označení entity třídy událost - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054
- **2063** - preferované označení entit třídy událost - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2064** - všechna označení entity třídy geografický objekt - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054
- **2065** - preferované označení entit třídy geografický objekt - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2066** - všechna označení entity třídy dílo/výtvar - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054
- **2067** - preferované označení entit třídy dílo/výtvar - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2068** - všechna označení entity třídy obecný pojem - obsahuje všechna označení entity bez ohledu na aplikovaná pravidla, jazyk nebo preferované/variantní označení - použití jako kompletního termínu viz atribut 2054
- **2069** - preferované označení entit třídy obecný pojem - obsahuje všechna preferované označení entity bez ohledu na aplikovaná pravidla označení - použití jako kompletního termínu viz atribut 2054
- **2008** - obecný doplněk - obsahuje obecný doplněk v označení entit, kde je obecný doplněk využíván
- **2009** - chronologický doplněk - obsahuje obecný doplněk v označení entit, kde je chronologický doplněk využíván
- **2010** - geografický doplněk - obsahuje obecný doplněk v označení entit, kde je geografický doplněk využíván

- 2011 - příslušnost k zemi původu - obsahuje hlavní část označení (preferovaných i variantních) entity geografický objekt, která je využita pro vyjádření příslušnosti k zemi původu
- 2012 - obor působnosti - obsahuje hlavní část označení (preferovaných i variantních) entity obecný pojem, která je využita pro vyjádření oboru působnosti
- 2013 - kategorie - obsahuje hlavní část označení (preferovaných i variantních) entity obecný pojem, která je využita pro vyjádření kategorie entity
- 2014 - jazyková oblast - obsahuje kód jazykové oblasti entity
- 2015 - datum vzniku - obsahuje datum vzniku entity
- 2016 - datum zániku - obsahuje datum zániku entity
- 2017 - místo vzniku - obsahuje hlavní část označení (preferovaných i variantních) entity geografický objekt, která je využita pro vyjádření místa vzniku entity
- 2018 - místo zániku - obsahuje hlavní část označení (preferovaných i variantních) entity geografický objekt, která je využita pro vyjádření místa zániku entity
- 2019 - sídlo - obsahuje hlavní část označení (preferovaných i variantních) entity geografický objekt, která je využita pro vyjádření sídla entity
- 2020 - geografická působnost - obsahuje hlavní část označení (preferovaných i variantních) entity geografický objekt, která je využita pro vyjádření geografické působnosti entity
- 2070 - hierarchická struktura geografické entity - všechny úrovně
- 2071 - hierarchická struktura geografické entity - stát
- 2072 - hierarchická struktura geografické entity - oblast/země
- 2073 - hierarchická struktura geografické entity - kraj
- 2074 - hierarchická struktura geografické entity - okres
- 2075 - hierarchická struktura geografické entity - obec/město
- 2076 - hierarchická struktura geografické entity - část obce
- 1016 - vše - obsahuje údaje ze všech částí záznamu
- 12 - INTERPI identifikační číslo záznamu

## Atribut 4 (fráze)

Implementovaná pouze hodnota 1 - fráze. Při použití tohoto atributu je možné zadat jako vyhledávací termín více slov - oddělených mezerami. Výsledkem budou ty záznamy, které mají ve vyhledávacím kritériu uvedena všechna zadaná slova ve stejném pořadí a bez výskytu jiného slova mezi nimi. Může se před nimi a za nimi vyskytovat jakýkoliv počet jiných slov.

## Atribut 5 (rozšíření)

Implementovaná pouze hodnota 1 - pravostranné rozšíření. Při použití tohoto atributu je zadán vyhledávací termín rozšířen o jakýkoliv počet libovolných znaků. Atribut je možné využít pouze s jednoslovnými termíny - tj. není implementován pro použití s frázemi.

## Atribut 6 (úplnost - completeness)

Implementovaná pouze hodnota 3 - kompletní pole (resp. hodnota 2 - kompletní podpole, která je používána ve stejném významu). Při použití tohoto atributu jsou výsledkem pouze záznamy, které obsahují v příslušném vyhledávacím kritériu uvedenou hodnotu přesně a úplně.

## Syntaxe

Pro zadávání dotazu platí „dědičnost“ atributů. Atributy uvedené na začátku dotazu platí pro všechny termíny.

```
@attr 1=1016 @and 'xx' 'yy'
```

Atribut 1 s hodnotou "1016" platí pro oba termíny 'xx' i 'yy'.

```
@attr 6=3 @attr 1=1016 @and @and 'xx' 'yy' 'zz'
```

Atribut 6 s hodnotou "3" a atribut 1 s hodnotou "1016" platí pro všechny tři termíny 'xx' i 'yy' i 'zz'.

Pro některý z vyhledávacích termínů je možné zděděný atribut libovolně změnit, nebo doplnit jiný.

```
@attr 1=1016 @and 'xx' @attr 5=1 'yy'
```

Atribut 1 s hodnotou "1016" platí pro oba termíny, pro druhý termín platí i atribut 5 s hodnotou "1".

Lze kombinovat i více různých hodnot atributů.

```
@attr 1=1016 @and 'xx' @attr 1=2050 'yy'
```

Atribut 1 s hodnotou "1016" platí pro oba termíny, pro druhý termín platí i atribut 5 s hodnotou "1".

Pro kombinaci různých vyhledávacích termínů je možné využít operátory AND, OR, NOT.

## Příklady použití

Pro následující příklad předpokládejme následující hodnoty vyhledávacího kritéria 2056 (kompletní termíny):

jindrák, františek 1915-	(ID 001)
jindrák, jindřich 1931-1993	(ID 002)
jindra, jan nepomucký 1863-1930	(ID 003)
jindra, jan 1877-	(ID 004)
jindra, jan 1922-1979	(ID 005)
jindra, jan 1962-	(ID 006)
jindr, josef	(ID 007)
jindrová, zdeňka 1910-	(ID 008)
nepomucký, jan 1895-1948	(ID 009)

Vyhledávací kritérium 2056 (jednotlivé slova):

františek	(ID 001)
jan	(ID 003, 004, 005, 006, 009)
jindr	(ID 007)
jindra	(ID 003, 004, 005, 006)
jindrák	(ID 001, 002)
jindrová	(ID 008)

jindřich	(ID 002)
josef	(ID 007)
nepomucký	(ID 003, 009)
zdeňka	(ID 008)

#### A hodnoty kritéria 2012:

česká literatura (ID 006, ID 008)

#### Příklady použití operátorů:

@attr 1=2056 @and 'jindra' 'jan'  
výsledkem jsou ID: 003, 004, 005, 006  
@attr 1=2056 @not 'jindrak' 'františek'  
výsledkem je ID: 002  
@attr 1=2056 @or 'jindra' 'jan'  
výsledkem jsou ID: 003, 004, 005, 006, 009

#### Příklady atributu fráze:

@attr 1=2056 @attr 4=1 'jindra, jan'  
výsledkem jsou ID: 003, 004, 005, 006  
@attr 1=2056 @attr 4=1 'jan nepomucký'  
výsledkem je ID: 003

#### Příklady atributu úplnost:

@attr 1=2056 @attr 6=3 'jindra, jan'  
výsledkem není žádný záznam  
@attr 1=2056 @attr 6=3 'jindra, jan 1962-'  
výsledkem je ID: 006

#### Příklady atributu rozšíření:

@attr 1=2056 @attr 5=1 'jindr'  
výsledkem jsou ID: 001, 002, 003, 004, 005, 006, 007, 008  
@attr 1=2056 @attr 5=1 'jindr' 'j'  
výsledkem jsou ID: 002, 003, 004, 005, 006, 007

#### Příklady použití více atributů:

@and @attr 1=2056 'jindra' @attr 1=2012 @attr 4=1 'česká literatura'  
výsledkem je ID: 006

## Třídící sekvence

Při vyhledávání záznamů například webovou službou findData je možné využít parametr sQuerySort pro třídění. Bez zadání třídění budou záznamy vráceny v pořadí, v jakém byly vkládány do databáze.

Třídící sekvence se skládá z maximálně 4 skupin výrazů pro výběr třídícího kritéria a směru třídění. Každá skupina má následující části:

- výběr třídícího kritéria - pro třídící kritérium se využívá číslo vhodného vyhledávacího kritéria ve variantě kompletní termín (vyhledávací kritéria vhodné pro třídění jsou uvedeny níže). Jako třídící kritérium jsou vhodné především ty vyhledávací kritéria, které obsahují jednu hodnotu, v jiném případě bude záznam zatříděn pouze podle jedné z hodnot vyhledávacího kritéria. Číslo vyhledávacího kritéria je uvozeno termínem „1=“.  
1=2054c
- směr třídění - možné je třídít záznam vzestupně (znak <) nebo sestupně (znak >). Znak třídění je uvozen písmenem „i“.

### Příklady třídících sekvencí

1=2054c i<

- třídění vzestupně podle preferovaného označení entity

1=2051c i< 1=2054c i<

- třídění vzestupně podle třídy entit a v rámci třídy podle preferovaného označení entity

### Doporučené třídící kritéria:

- **2051c** - třída
- **2052c** - podtřída
- **2054c** - preferované označení entity
- **1012c** - datum a čas poslední modifikace záznamu
- **30c** - datum vytvoření záznamu