

dotazování ředitelů, učitelů, či zda se tázal inspektor). Tyto a další podobné informace je třeba odvodit až po prozkoumání samotných datových souborů, přičemž jejich identifikace není vždy jednoznačná. Povahu dat pak podrobně znají ti zaměstnanci ČŠI, kteří jsou odpovědní za jejich pořizování, např. školní inspektoři.

Po zvolení jednoho či více formulářů je uživateli nabídnuta možnost vybrání konkrétních otázek, které budou následnou součástí výstupu. Kromě standardní možnosti, kdy uživatel vyplní první řádek excelového výstupu názvy otázek a zahrnutí či vyjmutí nezodpovězených formulářů, je možné zvolit si podobu výstupů jednotlivých odpovědí – **název odpovědi** (celé slovní znění odpovědi), **PID** (číselný identifikátor otázky/odpovědi) a **STAT ID** (kódovaný identifikátor odpovědi). Pro analytickou činnost je nejvhodnější forma STAT ID pracující s již nakódovanými soubory. V praxi však STAT ID není inspektory běžně vyplňováno, což znesnadňuje pozdější analytickou práci se soubory, které je nutné manuálně nakódovat. Soubory bez vyplněného STAT ID či soubory se specifickými rozvětvenými odpověďmi nahradí odpovědi znakem „?“ , což znemožní i zpětné manuální kódování datového souboru přes funkci „najít a nahradit“.

Problémem pro následnou analýzu je taktéž nepropojitelnost většiny odpovědí v čase, kdy se odpovědi liší jak svým zněním, tak číselným identifikátorem (vhodnější by bylo zavedení „knihovny“ odpovědí s určenými číselnými identifikátory, které by zajistily kontinuitu ve znění odpovědí a taktéž porovnatelnost takových odpovědí v čase). Jedním z důvodů proměny znění otázek je nutná změna kritérií pro hodnocení v čase, přesto se objevují stejné odpovědi s rozdílnými identifikátory.

Poslední možností přípravy datového výstupu je shlukování dat dle nabízených kritérií, tj. podle čísla akce a identifikátoru SF, podle názvu instituce (resp. školy) a podle činnosti školy. Pro pokročilou nedeskriptivní analytickou činnost je takové shlukování irelevantní. Využitelnější je možnost filtrování dat pomocí dodatečných filtrů umožňujících vygenerování dat např. jen dle specifické odpovědi u konkrétní otázky. V praxi se však podobná filtrace dat provádí až přímo v analytických programech a je vhodnější pouze pro deskriptivní statistiku.

Systém InspIS DATA umožňuje zpracování konečných datových výstupů ve dvou formách: 1) vygenerování výstupu do GRIDu a 2) vygenerování výstupu do excelového souboru. Výstup v GRIDu umožňuje generování výsledků do online deskriptivní tabulky, která však není tolik přehledná jako běžné výstupy ze statistických programů a znesnadňuje další možnosti zpracování dat. Slouží tak spíše jako jednoduchý náhled do vybraných dat bez nutnosti stahovat celý datový soubor.

Analyticky užitečný je výstup do excelu, a to za předpokladu manuálního nakódování jednotlivých odpovědí, vyčištění souboru a převedení do formátu vhodného pro běžně používané statistické programy (jako např. SPSS, Stata a další). Výstup pro excel umožňuje pouze základní deskriptivní statistiku. Analyticky přívětivější by byla možnost generovat výstupy dat z InspIS s nakódovanými odpověďmi tak, aby nebylo nutné čištění a rekódování dat do použitelné podoby. Ideálně i přímo v dalších formátech analytických programů. To by výrazně zefektivnilo práci analytických týmů.

Pro účely párování datových souborů ze systému InspIS DATA jak mezi sebou, tak mezi externími datovými soubory, je nejvhodnější specifický číselný identifikátor **REDIZO** (identifikační číslo ředitelství školy). Párování pomocí názvu školy se ukázalo jako nevhodné pro rozdílný zápis názvu školy v různých datových souborech, takže by vyžadovalo manuální sjednocení názvů (kvůli potřebnosti totožných textových řetězců pro účely párování), což však není příliš efektivní a bez chyb snad ani možné.