

PrInt: a Provenance model to support Integration processes

Disciplina: Procedência de Dados e Data Warehousing

Aluna: Dayse Silveira de Almeida

Professora: Cristina D. A. Ciferri



Roteiro

- Modelo PrInt
 - Operações transitivas
 - Operação que altera a chave
 - Operações de sobreposição
 - Política *Redo*
 - Validação
 - Reaplicação das operações

Fundamentação Teórica

Modelo PrInt

- PrInt (*Provenance model to support Integration process* - Modelo de Procedência para subsidiar Processos de Integração) [Tomazela *et al.*, 2010];
- Integração não retifica conflitos de dados nas fontes;
- Decisões de integração tomadas em processos anteriores são **reaplicadas** automaticamente em processos subsequentes.

Fundamentação Teórica

Modelo Print

- Usuário interage com o processo de integração para tomada de decisões sobre possíveis inconsistências;
- Operações:
 - Cópia, edição, inserção, remoção.
- Repositório de operações;
- Cópias locais integradas das fontes.

Fundamentação Teórica

Modelo Print

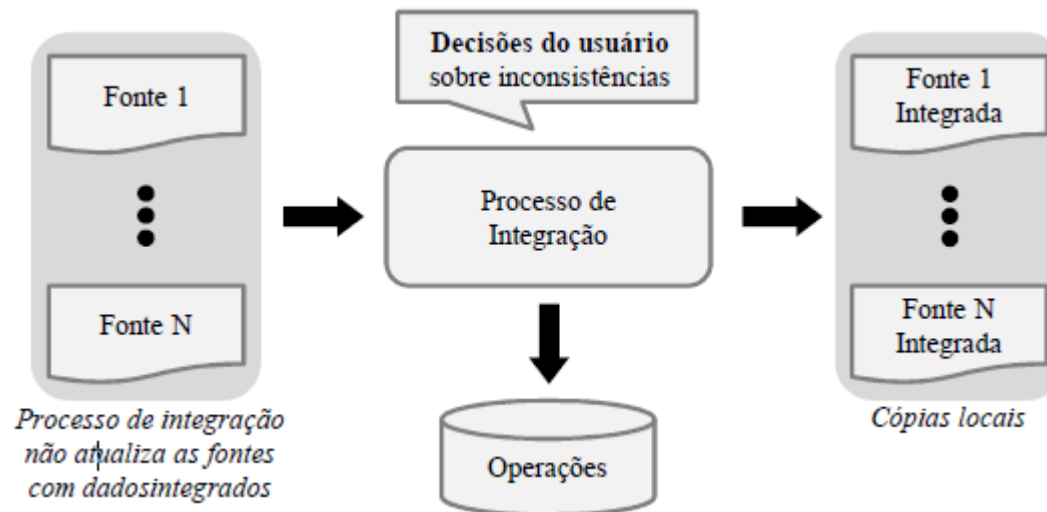


Figura 1: Integração de fontes de dados no tempo t1 [Tomazela, 2010].

Fundamentação Teórica

Modelo PrInt

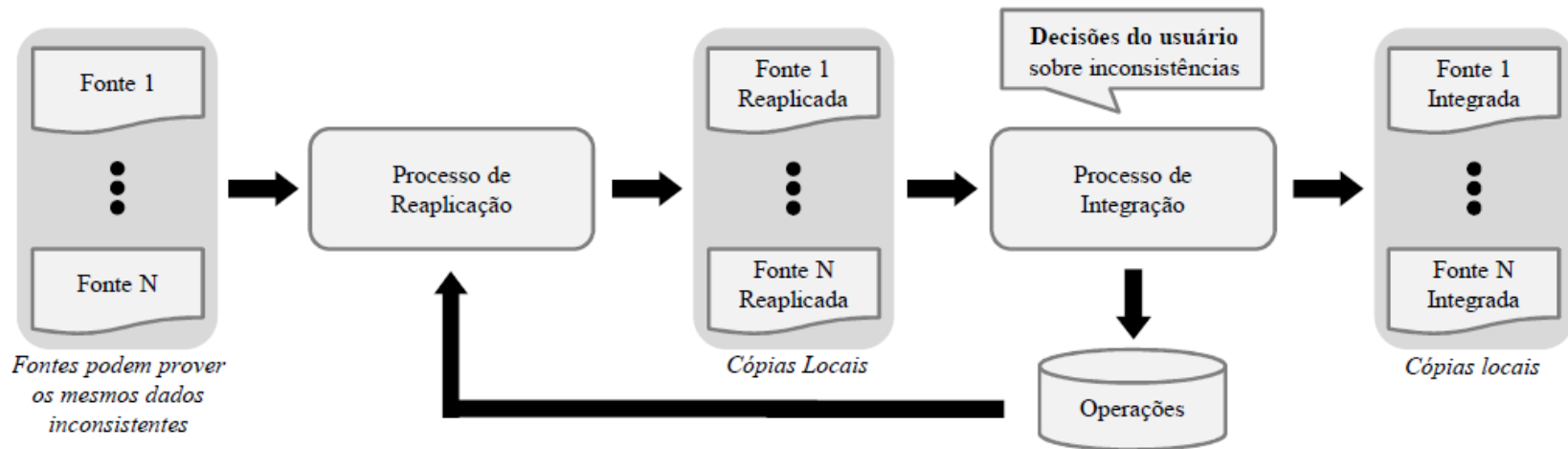


Figura 2: Integração de fontes de dados no tempo t_n [Tomazela, 2010].

Fundamentação Teórica

Exemplo Modelo Print

- O objeto *Artigo* possui os atributos:
 - *Título, Ano, Revista, Página inicial, Página final, ISSN, Local publicação;*
- E o subobjeto *Autores*, composto pelos atributos *Nome* e *Ordem citação;*
- Chave: atributo *Título;*
- Três fontes distintas são definidas, contendo dados curriculares de três autores: *Ana, Bruno e Carlos.*

Exemplo Modelo Print

Fonte: Ana

Artigo

Título: “MultiPrint: Multiuser Provenance Model for Data Integration”

Ano: “2004”

Revista: “Journal of the Brazilian Computer Society”

Página inicial: “5”

Página final: “20”

ISSN: “01046000”

Local publicação: “Porto Alegre, RS, Brasil”

Autores

Autor

Nome: “Ana”

Ordem citação: “1”

Autor

Nome: “Bruno”

Ordem citação: “2”

Fonte: Bruno

Artigo

Título: “MultiPrint: Multiuser Provenance Model for Data Integration”

Ano: “200”

Revista: “JBCS”

Página inicial: “5”

Página final: “24”

ISSN: “01046500”

Local publicação: “Porto Alegre, RS, Brasil”

Autores

Autor

Nome: “Ana”

Ordem citação: “1”

Autor

Nome: “Bruno”

Ordem citação: “2”

Autor

Nome: “Carlos”

Ordem citação: “3”

Fonte: Carlos

Artigo

Título: “Multiprint: multiuser provenance model for data integration”

Ano: “2005”

Revista: “Journal of the Brazilian Comp. Society”

Página inicial: “1”

Página final: “20”

ISSN: “01000000”

Local publicação: “Porto Alegre-RS”

Autores

Autor

Nome: “Ana”

Ordem citação: “1”

Autor

Nome: “Carlos”

Ordem citação: “2”

Autor

Nome: “Bruno”

Ordem citação: “3”

Autor

Nome: “Daniel”

Ordem citação: “4”

Fundamentação Teórica

Operações Print

- Cópia, edição, inserção, remoção;
- Cópia – da fonte correta (*origem*) para a fonte considerada incorreta (*destino*);
- Edição – fontes com valores incorretos (consistentes ou não);
 - Edição atua sobre um *destino*, e o atributo *origem* não é utilizado;

Fundamentação Teórica

Operações Print

- Cópia, edição, inserção, remoção;
- Inserção – objeto presente em apenas uma fonte;
 - Para inserir o objeto faltante na fonte:
 - In – cria um objeto no *destino* contendo valores apenas para o atributo chave (*idObjeto*);
 - Os atributos *atributo*, *valorOrigem* e *valorDestino* não são utilizados;
 - Cp – uma para cada atributo não chave do objeto;
- Remoção – objeto incorretamente inserido em uma fonte;
 - Para removê-lo:
 - Ed – modifica o valor de cada atributo não chave para nulo;
 - Rm – permitida apenas se todos os atributos não chaves forem nulos;
 - Remove um objeto de um destino sem precisar de uma origem;
 - Atributos *atributo*, *valorOrigem* e *valorDestino* não são utilizados.

Fundamentação Teórica

Operações Print

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
1	ed	null	Carlos	Artigo[Título=MultiPrint: multiuser...]	Título	MultiPrint: Multiuser...	MultiPrint: multiuser...
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200
3	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	2005
4	ed	null	Carlos	Artigo [Título=MultiPrint: Multiuser...] Autor[Nome=Carlos]	Ordem citação	3	2
5	ed	null	Carlos	Artigo [Título=MultiPrint: Multiuser...] Autor [Nome=Bruno]	Ordem citação	2	3
6	ed	null	Carlos	Artigo [Título=MultiPrint: Multiuser...] Autor [Nome=Daniel]	Ordem citação	null	4
7	rm	null	Carlos	Artigo [Título=MultiPrint: Multiuser...] Autor [Nome=Daniel]	null	null	null
8	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Página final	24	20
9	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046500	01000000
10	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Local publicação	Porto Alegre, RS, Brasil	Porto Alegre-RS
11	in	Bruno	Ana	Artigo [Título=MultiPrint: Multiuser...] Autor [Nome=Carlos]	null	null	null
12	cp	Bruno	Ana	Artigo [Título=MultiPrint: Multiuser...] Autor [Nome= Carlos]	Ordem citação	3	null
13	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	Página final	24	20
14	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046500	01046000

Figura 3: Repositório de operações.

Fundamentação Teórica

Operações Print

- Armazenada a procedência das transformações realizadas sobre os dados;
- Relacionamento:
 - Transitividade;
 - Alteração da chave.

Fundamentação Teórica

Transitividade

- Transitiva direta: $a(\mathbf{destino}, \text{idObjeto}, \text{atributo}) = b(\mathbf{origem}, \text{idObjeto}, \text{atributo})$

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
1	ed	null	Carlos	Artigo[Título=MultiPrint: multiuser...]	Título	MultiPrInt: Multiuser...	MultiPrint: multiuser...
2	cp	Ana	Bruno	Artigo [Título=MultiPrInt: Multiuser...]	Ano	2004	200
3	cp	Bruno	Carlos	Artigo [Título=MultiPrInt: Multiuser...]	Ano	2004	2005

Figura 4: Operação transitiva direta.

Fundamentação Teórica

Transitividade

- Transitiva direta: $a(\mathbf{destino}, \text{idObjeto}, \text{atributo}) = b(\mathbf{origem}, \text{idObjeto}, \text{atributo})$

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
8	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Página final	24	20
...							
13	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	Página final	24	20
...							

Figura 5: Operação transitiva direta.

Fundamentação Teórica

Transitividade

- Transitiva indireta: se $a \rightarrow b$ e $b \rightarrow c$, então c é transitiva indireta a a .
- Nova fonte:

Fonte: Daniel

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "2004"

Revista: "Journal of the Brazilian Computer Society"

Página inicial: "5"

Página final: "15"

ISSN: "01046500"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem prioridade: "1"

Autor

Nome: "Bruno"

Ordem prioridade: "2"

Autor

Nome: "Carlos"

Ordem prioridade: "3"

Autor

Nome: "Daniel"

Ordem citação: "4"

Fundamentação Teórica

Transitividade

- Transitiva indireta: se $a \rightarrow b$ e $b \rightarrow c$, então c é transitiva indireta a a .

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
8	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Página final	24	20
...							
13	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	Página final	24	20
...							
15	cp	Ana	Daniel	Artigo[Título=MultiPrint: multiuser...]	Página Final	24	15

Figura 6: Operação transitiva indireta.

Fundamentação Teórica

Operação que altera a chave

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
1	ed	null	Carlos	Artigo[Título=Multiprint: multiuser...]	Título	MultiPrint: Multiuser...	Multiprint: multiuser...
...							

Figura 7: Operação que altera a chave.

Fundamentação Teórica

Operações de Sobreposição

- Uma operação sobrescreve o resultado de outra e pode gerar inconsistências;

Fundamentação Teórica

Operações de sobreposição

- Uma operação sobrescreve o resultado de outra e pode gerar inconsistências;
 - Sobreposição do destino:
 - $a = (\text{destino}, \text{idObjeto}, \text{atributo})$ e $b = (\text{destino}, \text{idObjeto}, \text{atributo})$, ($a < b$).

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200
3	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	2005
...							
16	ed	null	Bruno	Artigo[Título=MultiPrint: Multiuser...]	Ano	2009	2004

Figura 8: Sobreposição de destino.

Fundamentação Teórica

Operações de sobreposição

- Uma operação sobrescreve o resultado de outra e pode gerar inconsistências;
 - Sobreposição da origem:
 - $a = (\text{origem}, \text{idObjeto}, \text{atributo})$ e $b = (\text{destino}, \text{idObjeto}, \text{atributo})$, ($a < b$).

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
9	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046500	01000000
14	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046500	01046000
...							
17	ed	null	Bruno	Artigo[Título= MultiPrint: Multiuser...]	ISSN	01046555	01046500

Figura 8: Sobreposição da origem.

Fundamentação Teórica

Operações de Sobreposição

- Tratamento:
 - Blind – operações de sobreposição não são tratadas;
 - Restrict – operações de sobreposição não são permitidas;
 - Undo –
 - Operações de sobreposição são permitidas;
 - Operação sobreposta é removida do repositório;
 - Transitivas são desfeitas e removidas;
 - Redo –
 - Operações de sobreposição são permitidas;
 - Se sobreposição da origem - operação sobreposta é movida para o final do repositório;
 - Se sobreposição do destino - operação sobreposta é removida do repositório;
 - Transitivas são refeitas e movidas para o final do repositório.

Fundamentação Teórica

Política *Redo*

- Sobreposição do destino:

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200
3	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	2005
...							
16	ed	null	Bruno	Artigo[Título=MultiPrint: Multiuser...]	Ano	2009	2004
17	ed	null	Bruno	Artigo[Título= MultiPrint: Multiuser...]	ISSN	01046555	01046500

Figura 9: Política *redo* para sobreposição do destino.

Fundamentação Teórica

Política *Redo*

- Sobreposição do destino:

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200
3	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Ano	2009	2005
...							
16	ed	null	Bruno	Artigo[Título=MultiPrInt: Multiuser...]	Ano	2009	2004
17	ed	null	Bruno	Artigo[Título= MultiPrInt: Multiuser...]	ISSN	01046555	01046500

Figura 9: Política *redo* para sobreposição do destino.

Fundamentação Teórica

Política *Redo*

- Sobreposição do destino:

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200
3	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Ano	2009	2005
...							
16	ed	null	Bruno	Artigo[Título=MultiPrInt: Multiuser...]	Ano	2009	2004
17	ed	null	Bruno	Artigo[Título= MultiPrInt: Multiuser...]	ISSN	01046555	01046500
3	cp	Bruno	Carlos	Artigo [Título=MultiPrInt: Multiuser...]	Ano	2009	2005

Figura 9: Política *redo* para sobreposição do destino.

Fundamentação Teórica

Política *Redo*

- Sobreposição da origem:

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
9	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046500	01000000
...							
14	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046500	01046000
...							
17	ed	null	Bruno	Artigo[Título= MultiPrint: Multiuser...]	ISSN	01046555	01046500

Figura 10: Política *redo* para sobreposição da origem.

Fundamentação Teórica

Política *Redo*

- Sobreposição da origem:

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
9	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046555	01000000
...							
14	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046555	01046000
...							
17	ed	null	Bruno	Artigo[Título= MultiPrint: Multiuser...]	ISSN	01046555	01046500

Figura 10: Política *redo* para sobreposição da origem.

Fundamentação Teórica

Política *Redo*

- Sobreposição da origem:

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
9	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046555	01000000
...							
14	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046555	01046000
...							
17	ed	null	Bruno	Artigo[Título= MultiPrint: Multiuser...]	ISSN	01046555	01046500
9	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046555	01000000
14	cp	Carlos	Ana	Artigo [Título=MultiPrint: Multiuser...]	ISSN	01046555	01046000

Figura 10: Política *redo* para sobreposição da origem.

Fundamentação Teórica

Validação

- Verificar se as fontes possuem os mesmos dados desde que as operações foram definidas até o momento da reaplicação;
- Tipos:
 - Sem validação;
 - Validação da origem – evita anomalia das fontes consistentes;
 - Validação do destino – evita a anomalia do destino sobrescrito;
 - Validação completa.

Fundamentação Teórica

Validação Completa

- Anomalia das fontes consistentes:

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "200"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

Fundamentação Teórica

Validação Completa

- Anomalia das fontes consistentes:

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "200"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre-RS"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

Fundamentação Teórica

Validação Completa

- Anomalia das fontes consistentes:
 - Reaplicar a op 10 faria as fontes *Bruno* e *Carlos* ficarem inconsistentes.

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
10	cp	Bruno	Carlos	Artigo [Título=MultiPrInt: Multiuser...]	Local publicação	Porto Alegre, RS, Brasil	Porto Alegre- RS
...							

Figura 11: Anomalia das fontes consistentes.

Fundamentação Teórica

Validação Completa

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
10	cp	Bruno	Carlos	Artigo [Título=MultiPrInt: Multiuser...]	Local publicação	Porto Alegre, RS, Brasil	Porto Alegre- RS
...							

Figura 11: Anomalia das fontes consistentes.

- Pode ser armazenada em *backup*;
- Não volta para o repositório, mesmo que nova alteração na fonte torne a operação válida;
 - Repositório – conjunto finito e gerenciável de op;
 - Reaplicação acarrete o menor custo adicional no processo de integração.

Fundamentação Teórica

Validação Completa

- Anomalia do destino sobrescrito:

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "200"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

Fundamentação Teórica

Validação Completa

- Anomalia do destino sobrescrito:

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "2007"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

Fundamentação Teórica

Validação Completa

- Anomalia do destino sobrescrito:
 - Reaplicar a op 2 faria as fontes *Ana* e *Bruno* consistentes, mas o usuário não toma conhecimento da alteração.

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
...							
2	cp	Ana	Bruno	Artigo [Título=MultiPrnt: Multiuser...]	Ano	2004	200
...							

Figura 12: Anomalia do destino sobrescrito.

Fundamentação Teórica

Reaplicação das Operações

- VRS (*Validate and Reapply in Separate*):
 - Valida as operações antes de reaplicá-las;

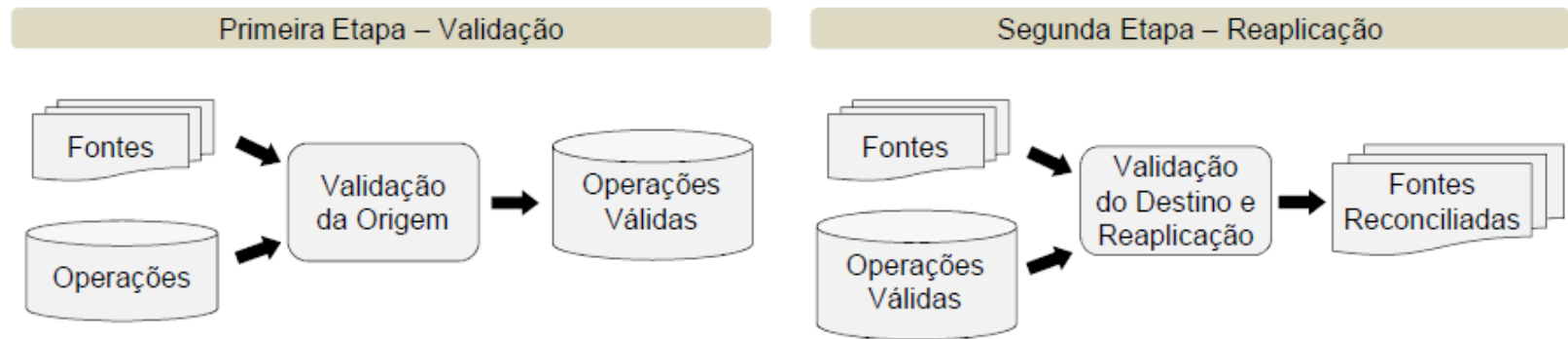


Figura 13: Estágios do método VRS [Tomazela, 2010].

Fundamentação Teórica

Reaplicação das Operações

- VRT (*Validate and Reapply in Tandem*):
 - Valida e reaplica as operações simultaneamente.

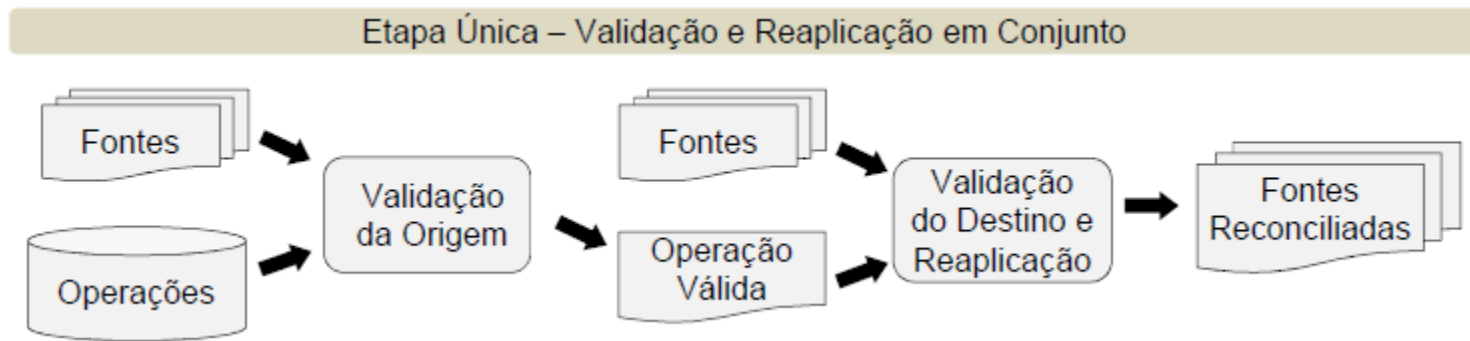


Figura 14: Estágios do método VRT [Tomazela, 2010].

Fundamentação Teórica

VRT

Fonte: Ana

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "2004"

Revista: "Journal of the Brazilian Computer Society"

Página inicial: "5"

Página final: "20"

ISSN: "01046000"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "2007"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200

Fundamentação Teórica

VRT

Fonte: Ana

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "2004"

Revista: "Journal of the Brazilian Computer Society"

Página inicial: "5"

Página final: "20"

ISSN: "01046000"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "2007"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre, RS, Brasil"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
2	cp	Ana	Bruno	Artigo [Título=MultiPrint: Multiuser...]	Ano	2004	200

Fundamentação Teórica

VRT

Fonte: Bruno

Artigo

Título: "MultiPrint: Multiuser Provenance Model for Data Integration"

Ano: "200"

Revista: "JBCS"

Página inicial: "5"

Página final: "24"

ISSN: "01046500"

Local publicação: "Porto Alegre-RS"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Bruno"

Ordem citação: "2"

Autor

Nome: "Carlos"

Ordem citação: "3"

Fonte: Carlos

Artigo

Título: "Multiprint: multiuser provenance model for data integration"

Ano: "2005"

Revista: "Journal of the Brazilian Comp. Society"

Página inicial: "1"

Página final: "20"

ISSN: "01000000"

Local publicação: "Porto Alegre-RS"

Autores

Autor

Nome: "Ana"

Ordem citação: "1"

Autor

Nome: "Carlos"

Ordem citação: "2"

Autor

Nome: "Bruno"

Ordem citação: "3"

Autor

Nome: "Daniel"

Ordem citação: "4"

id	op	origem	destino	idObjeto	atributo	valorOrigem	valorDestino
10	cp	Bruno	Carlos	Artigo [Título=MultiPrint: Multiuser...]	Local publicação	Porto Alegre, RS, Brasil	Porto Alegre-RS

Fundamentação Teórica

Reordenação Segura

- Visa minimizar a quantidade de carregamento das fontes durante a replicação;
- Não altera a ordem das operações transitivas, de operações que alteram a chave e operações de sobreposição;
- Melhora o desempenho do VRT, sempre que o repositório for passível de reordenação.

Fundamentação Teórica

Modelo de Procedência do PrInt

- Modelo de procedência:
 - *Why e where-provenance* – armazena a origem dos dados de cada operação e as alterações realizadas;
 - Procedência de cada operação de cópia, inserção e remoção, usadas no processo de integração.
 - Coleta automática, tratando as operações que se sobrepõem como forma de manutenção;
 - Procedência armazenada separada dos dados, em um repositório de operações;
 - Processo de integração não tem permissão para escrever nas fontes.
 - Consulta por filtro e por rastreamento, no desenvolvimento de métodos de reaplicação de operações.
 - VRS – rastreamento;
 - VRT – filtro.

Referências

- [Tomazela 2010] Tomazela, B. (2010). MPPI: um modelo de procedência para subsidiar processos de integração. Dissertação de mestrado, Universidade de São Paulo, São Carlos, SP.
- [Tomazela *et al.* 2010] Tomazela, B., Hara, C. S., Ciferri, R. R., and Ciferri, C. D. A. (2010). Print: a provenance model to support integration processes. In Proceedings of the 19th International Conference on Information and Knowledge Management (CIKM), pages 1349-1352.
- [Tomazela *et al.* 2013] Tomazela, B., Hara, C. S., Ciferri, R. R., Ciferri, C. D. A. (2013). Empowering Integration Processes with Data Provenance. *Data & Knowledge Engineering*. Aceito para publicação.

PrInt: a Provenance model to support Integration processes

Disciplina: Procedência de Dados e Data Warehousing

Aluna: Dayse Silveira de Almeida

Professora: Cristina D. A. Ciferri

Obrigada.

