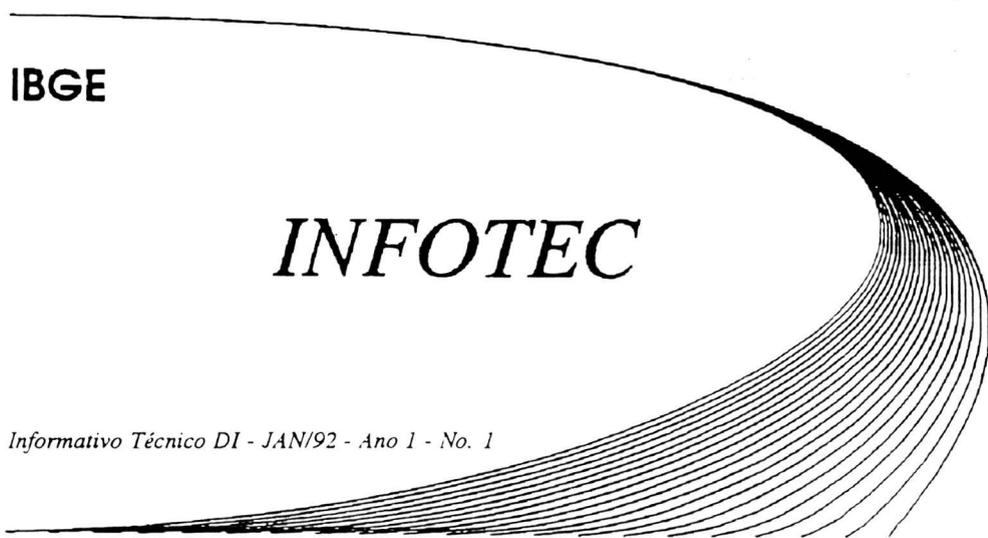


**IBGE**

# *INFOTEC*



*Informativo Técnico DI - JAN/92 - Ano 1 - No. 1*

**CRIPTA**

**Integração à Arquitetura de Sistemas de Informática para as Pesquisas Estatísticas**

**Murilo Cardoso de Castro**

**Reina Marta Hanono**

**DEPARTAMENTO TÉCNICO**

**DI**

## 1 - INTRODUÇÃO

Durante a última década a comunidade de informática objetivou a produção de sistemas de boa qualidade, a um baixo custo e a um bom nível de produtividade no seu desenvolvimento.

Partilhando de tais objetivos, procuramos, no IBGE, um conjunto integrado de ferramentas de uso amigável, visando o aumento de produtividade e qualidade no desenvolvimento e produção de Sistemas de Informática para o Processamento de Dados Estatísticos (Pesquisas/Censos).

Entendemos por Sistema de Informática para o Processamento de Dados Estatísticos (**Figura 1**, em anexo) um conjunto de Aplicações, Dados, Tecnologias e Serviços que permitem realizar as distintas fases do ciclo de vida de uma pesquisa (Apuração: Coleta, Captura, Codificação, Crítica/Correção; Análise: Crítica de Agregados, Exploração e Tabulação; Disseminação; e, Acervo).

Destacando do Sistema de Informática apenas o componente Aplicações, projetamos a Arquitetura deste componente (**Figura 2**, em anexo) de acordo com um Esquema Genérico para Aplicações em Pesquisas Estatísticas, propondo diretrizes metodológicas e tecnológicas, que possam sustentar a comunidade de usuários finais e de especialistas em informática. Dentro deste Esquema, o CRIPTA, ferramenta voltada para a execução das etapas de Crítica e Correção, dentro da fase de Apuração de uma Pesquisa, possibilita:

- 1) - a especificação das regras de críticas de forma simples e padronizada;
- 2) - a geração automática do programa de crítica a partir das especificações dos usuários;
- 3) - a garantia sobre a integridade das informações constantes nas especificações;
- 4) - a integração com Sistemas Gerenciadores de Banco de Dados;
- 5) - a integração com as distintas ferramentas que fazem parte do ambiente de software do IBGE;
- 6) - a confiabilidade e produtividade no desenvolvimento de aplicações de crítica/imputação;
- 7) - a portabilidade do produto gerado (programas de crítica e imputação) para outras plataformas tecnológicas (descentralização);

8) - maior integração entre os membros da equipe de desenvolvimento, em especial entre analista temático e analista de sistemas.

## **2 - CRIPTA NA ARQUITETURA DE SISTEMAS DE INFORMÁTICA PARA AS PESQUISAS ESTATÍSTICAS**

Considerando a importância de se ter uma linguagem padronizada para a especificação dos Planos de Crítica de qualquer pesquisa, independente dos recursos de informática que esta pesquisa venha a adotar, entendemos que: **a integração do CRIPTA na Arquitetura proposta é um fator fundamental.** Ao adotá-lo, o usuário não está preso a uma "camisa de força", pois o sistema está preparado para se combinar com as distintas ferramentas do Esquema Genérico, segundo as necessidades do usuário.

A integração pode ser vista através do interfaceamento com outros sistemas ou ferramentas existentes no Esquema, sejam para o sistema central IBM ou sistemas distribuídos, assim como através das funções que o próprio CRIPTA realiza, o que pode ser resumido na **Figura 3**, em anexo.

A seguir descrevemos as soluções de interface entre o CRIPTA e cada uma das ferramentas do Esquema Genérico que modela a Arquitetura de Aplicações do Sistema de Informática para as Pesquisas Estatísticas. Antes, todavia, cabe apresentar o aperfeiçoamento aplicado ao CRIPTA, tendo em vista a política de descentralização, especialmente da fase de apuração de pesquisas.

### **2.1 - APERFEIÇOAMENTO DO SISTEMA CRIPTA (PROCESSAMENTO DESCENTRALIZADO)**

#### **OBJETIVO:**

Em função das diretrizes de descentralização, retratadas na própria Arquitetura do Sistema de Informática para as Pesquisas, se decidiu dotar o sistema CRIPTA existente atualmente na instituição da característica de **portabilidade**, adaptando o sistema para que seu produto (programas de crítica e atualização) possa ser processado em diferentes plataformas, além do sistema central IBM.

#### **DESCRIÇÃO:**

Analisando-se a forma de desenvolvimento de aplicações para a fase de Apuração de Pesquisas na instituição, verificamos que a especificação e implementação, da etapa de Crítica e Correção, são feitas de modo centralizado.

Aliando este fato à necessidade de se dispor deste aperfeiçoamento do CRIPTA em um prazo reduzido, concluímos que: a alternativa que provocaria menor número de alterações no sistema é a de reespecificar e implementar somente a fase de geração de código do sistema CRIPTA de forma a gerar os programas de crítica e correção em linguagem "C".

Com esta nova versão, os usuários continuarão fazendo suas especificações no sistema central, sendo que o produto gerado poderá rodar em plataformas descentralizadas, dada a grande portabilidade oferecida pela linguagem de programação "C".

No caso de plataformas UNIX com o sistema de banco de dados, TSGBD (disponível nos supermicros e superminis), o programa de crítica/atualização gerado terá interface com o TSGBD, e permitirá a correção dos dados em um único ciclo 'on-line', após a primeira passagem de crítica em 'batch`.

**ESTADO:** Em fase de teste com a simulação da PNAD.

**FORMA DE EXECUÇÃO:** No sistema central IBM/VM, através da **ECRIPTAX**; ou no supermicro EDISA/EDIX, através da **CRIPTAX**.

## 2.2 - INTEGRAÇÃO CRIPTA-ATLAS

### **OBJETIVO:**

Permitir a recuperação do histórico, isto é, que pesquisas para as quais existam dicionários ATLAS para anos anteriores possam ser apuradas em anos subseqüentes, mediante a utilização do CRIPTA, assim como que, pesquisas descritas em CRIPTA, possam eventualmente utilizar os processos de **Tabulação do ATLAS**, ainda não disponíveis no CRIPTA.

### **DESCRIÇÃO:**

Primeiramente, vale ressaltar que as informações de dicionário contidas na Base de Metadados podem ser utilizadas indistintamente pelos sistemas ATLAS e CRIPTA.

Existe, no entanto, um utilitário (CONVERSOR) em ambiente convencional, que permite criar dicionários CRIPTA a partir de dicionários ATLAS.

**ESTADO:** Já implementado

**FORMA DE EXECUÇÃO:** No sistema central **IBM/VM**, através da **ECRIPTA**.

### 2.3 - **INTERFACE CRIPTA-DIA**

#### **OBJETIVOS:**

I) Permitir:

- a) a geração em ambiente convencional dos Dicionários DIA (FIPOLO e FIVAR) a partir de dicionários residentes na Base de Metadados;
- b) a geração em ambiente convencional dos conjuntos de Regras de Imputação Determinística (RID) e Regras de Edição (EDIT), a partir do plano de crítica (PCAUT) residente na Base de Metadados;
- c) a reformatação do arquivo a ser processado pelo DIA. Este deve ser um arquivo zonado;

objetivando a **utilização da metodologia Fellegi & Holt para imputação de dados qualitativos**, assim como possibilitando a **consistência automática de regras de crítica para variáveis intra-registros**.

II) Permitir o acesso do **DIA** a dados residentes no DB2.

#### **DESCRIÇÃO:**

A interface CRIPTA/DIA cria, em ambiente convencional, arquivos no formato do software "**DIA**", a partir de descrições de Dicionários e Planos de Crítica residentes na Base de Metadados, possibilitando a utilização do "**DIA**" nas diversas pesquisas do IBGE.

É composta por dois módulos:

**DICIONÁRIO** - Cria os arquivos "**FIVAR**" (Nome da variável e suas categorias), e "**FIPOLO**" (Nome, posição inicial e tamanho de cada variável);

**PLANO DE CRÍTICA** - Cria os arquivos "**RID**" (Regras de Imputação Determinística) e "**EDIT**" (Regras de Incompatibilidade), a partir de Planos de Crítica no formato **PCAUT**;

que interagem com as informações residentes na Base de Metadados.

Considerando-se a restrição do **DIA** de somente processar arquivos zonados com um único tipo de registro, será necessário tornar disponível aos usuários, um módulo conversor similar ao **TRANSLADO**, que poderia ser um processo do "**CRIPTA**".

Vale ressaltar que, hoje, pode ser utilizado o processo **TRANSLADO** para a geração do arquivo de dados para o **DIA**. Porém, o usuário seria obrigado a fazer as atribuições das variáveis não-zonadas para variáveis auxiliares zonadas, a serem utilizadas no **PROCESSO TRANSLADO**.

- ESTADO:**
- a) a interface para criação dos dicionários **DIA** (**FIVAR**, **FIPOLO**), já está disponível;
  - b) a interface para criação do conjunto de Regras de **EDIÇÃO** (**EDITS**) já está disponível;
  - c) a interface para a geração do conjunto de Regras de Imputação Determinística (**RIDS**) está sendo implementada;
  - d) o processo **TRANSLADO** que permite a geração de arquivos faz parte do **CRIPTA**, e translada as variáveis no formato de armazenamento que elas possuem, que não necessariamente será zonado. Portanto, poderá ser implementado, sem muito esforço, um processo similar que permita transladar, de forma automática, qualquer variável em formato zonado.
  - e) o acesso do **DIA** a dados residentes no **DB2** está em fase de estudo.

**FORMA DE EXECUÇÃO:** No sistema central **IBM/MVS**, através da **CRIPDIA**.

## 2.4 - INTERFACE CRIPTA-SAS

### **OBJETIVO:**

Permitir que as pesquisas descritas mediante dicionário **CRIPTA** e/ou os arquivos derivados das mesmas, gerados pelo **CRIPTA**, possam vir a ser processados para **análise de dados e/ou tabulação**, utilizando as facilidades do **SAS**.

### **DESCRIÇÃO:**

Atualmente existe a interface **METADADOS-SAS** (implementada pelo **DEBAD**) em fase de teste, que cria um arquivo com os comandos de descrição dos dados para ser utilizado como comandos de Input em um programa **SAS**.

Da mesma forma que o Sistema CRIPTA poderia criar de forma automática a descrição dos dados a ser utilizada como comandos de Input para o SAS, diretamente a partir do processo TRASLADO, que permite, hoje, a criação de arquivos derivados.

**FORMA DE EXECUÇÃO:** Através da aplicação de consulta ao Banco de Metadados, **META.**

## **2.5 - INTERFACE CRIPTA-DB2 E CRIPTA-TSGBD**

**OBJETIVOS:**

a) Viabilizar que sejam processados no CRIPTA, arquivos de dados residentes no sistema de banco de dados, DB2 (quando processo centralizado); ou no sistema de banco de dados, TSGBD (para processo descentralizado, em supermicros e/ou superminis);

**DESCRIÇÃO:**

a) As interfaces de acesso ao DB2 ou ao TSGBD não são interfaces genéricas, nem poderiam ser, devido aos distintos modelos lógicos dos Bancos a serem acessados, que poderão estar mapeados em uma ou mais tabelas no DB2/TSGBD;

De toda forma, foi definido para esta interface, uma forma genérica de implementação;

b) A integração da mesma ao CRIPTA é automática, já que no caso de existir esta interface, esta informação deverá existir no Dicionário.

**ESTADO:** A interface **CRIPTA-DB2** foi testada de forma satisfatória com os dados da PATR89.

A interface **CRIPTA-TSGBD** faz parte do aperfeiçoamento do CRIPTA (item 2.1) e está sendo testada junto com o sistema.

## **2.6 - INTEGRAÇÃO COM PACOTES DE PC-DOS**

### **2.6.1 - REDATAM**

**OBJETIVO:**

Obter as informações necessárias para explorar (**fase de análise**) um arquivo de dados com **REDATAM**, produzindo uma listagem das mesmas, facilitando a entrada destes no REDATAM, e a obtenção do arquivo a partir do qual deverá ser gerada a **BASE REDATAM**.

## DESCRIÇÃO:

O **REDATAM** trabalha com dicionário próprio, compactado na hora da entrada 'on-line'. Devido a este ser um sistema fechado, propomos a implementação de uma interface **CRIPTA-REDATAM**, que facilite aos usuários a geração de um Dicionário REDATAM, assim como a geração do arquivo zonado contendo as variáveis referenciadas.

Para tanto, esta interface deverá recuperar as informações do Dicionário CRIPTA na Base de Metadados, dando para cada variável especificada para a recuperação: a identificação da variável, o nome da variável, os limites inferior e superior, a posição inicial, o tamanho, o número de decimais, o tipo de registro a que pertence, assim como o nome e o código de cada uma das categorias, caso a variável seja categorizada. Além disso, ela deverá gerar um arquivo zonado contendo registros com as variáveis recuperadas, ficando este disponível para ser migrado para o micro.

**ESTADO:** Proposta em estudo.

### 2.6.2 - **IMPS**

#### **OBJETIVO:**

Gerar dicionário IMPS, em ambiente convencional, a partir das informações de dicionários residentes na Base de Metadados, possibilitando, assim, a utilização dos recursos do IMPS para processamento descentralizado.

#### **DESCRIÇÃO:**

Esta interface deverá facilitar a recuperação das informações dos Dicionários CRIPTA residentes na Base de Metadados, gerando o dicionário **IMPS-Batch** no sistema central IBM/MVS, ficando este arquivo pronto para ser migrado para o micro.

Para cada variável deverão ser recuperados: a **Identificação**, a **Posição inicial**, o **Tamanho**, o **Tipo** e a **Característica (Item ou Sub-item)**, assim como o código e o nome das categorias, caso elas sejam categorizadas.

Caso o usuário deseje, deverá ser gerado o arquivo zonado para posterior utilização pelo IMPS (**fase de análise**).

**ESTADO:** Proposta em estudo.

### **3 - CONCLUSÃO**

Muito esforço vem sendo dispendido no sentido de se implementar no IBGE uma Arquitetura de Sistema de Informática para o Processamento de Dados Estatísticos, que proporcione características modernas dentro de uma filosofia de 'sistemas abertos` - a direção da informática nos anos 90.

Dentro deste esforço, o CRIPTA tem um papel de grande importância pela garantia que oferece de geração direta de aplicações de crítica e correção de dados, a partir das especificações dos planos de crítica por parte dos analistas temáticos.

### **BIBLIOGRAFIA**

1. CRIPTA - Proposta de desenvolvimento DI/SUSIS/DISME, Novembro 1986. (disponível na DI/DETEC/DISIA).

2. CRIPTA - Sistema automatizado de Plano de Crítica, Imputação Automática e Especificação de Procedimentos dedicados - DI/GESID, Abril 1987.

3. BARBOSA, D. M. R. & HANONO, R. M. - Estudo das Ferramentas para Apuração de Dados. Revista Brasileira de Estatística, Rio de Janeiro, 49(191), JAN/JUN 1988, 85-100.

4. CRIPTA - Especificação Lógica - DI/GESIE (disponível na DI/DETEC/DISIA).

5. BARBOSA, D. M. R. - CRIPTA - Arquitetura do Sistema - (disponível na DI/DETEC/DISIA).

6. BARBOSA, D. M. R. - CRIPTA - Estrutura da Gramática - (disponível na DI/DETEC/DISIA).

7. CRIPTA - Avaliação do Sistema durante o Beta Teste - (disponível na DI/DETEC/DISIA).

8. CABRAL, M. S. S. - Um Esquema para a Apuração de Pesquisas Estatísticas do IBGE. Artigo Técnico A011/90.

9. CABRAL, M. S. S. - O Esquema Genérico para a Apuração das Pesquisas do Statistics Canada. Artigo Técnico A005/90.

10. CRIPTA - Manual da Linguagem (disponível na Biblioteca da DI ou através da ECRIPTA, no VM).

11. CRIPTA - Manual do Dicionário da Aplicação (disponível na Biblioteca da DI ou através da ECRIPTA, no VM).

12. CRIPTA - Manual do Plano de Crítica Automatizado (disponível na Biblioteca da DI ou através da ECRIPTA, no VM).

13. CRIPTA - Manual de Estudo de Casos (disponível na Biblioteca da DI).

14. GUEDES, A. P. - Banco de Metadados. Novembro, 1989, DEBAD/GEBAD (disponível na biblioteca do IBGE).

**INSTITUIÇÃO**

FABRICA DE DADOS

**SOCIEDADE**



FIGURA 1

# ARQUITETURA DE APLICAÇÕES

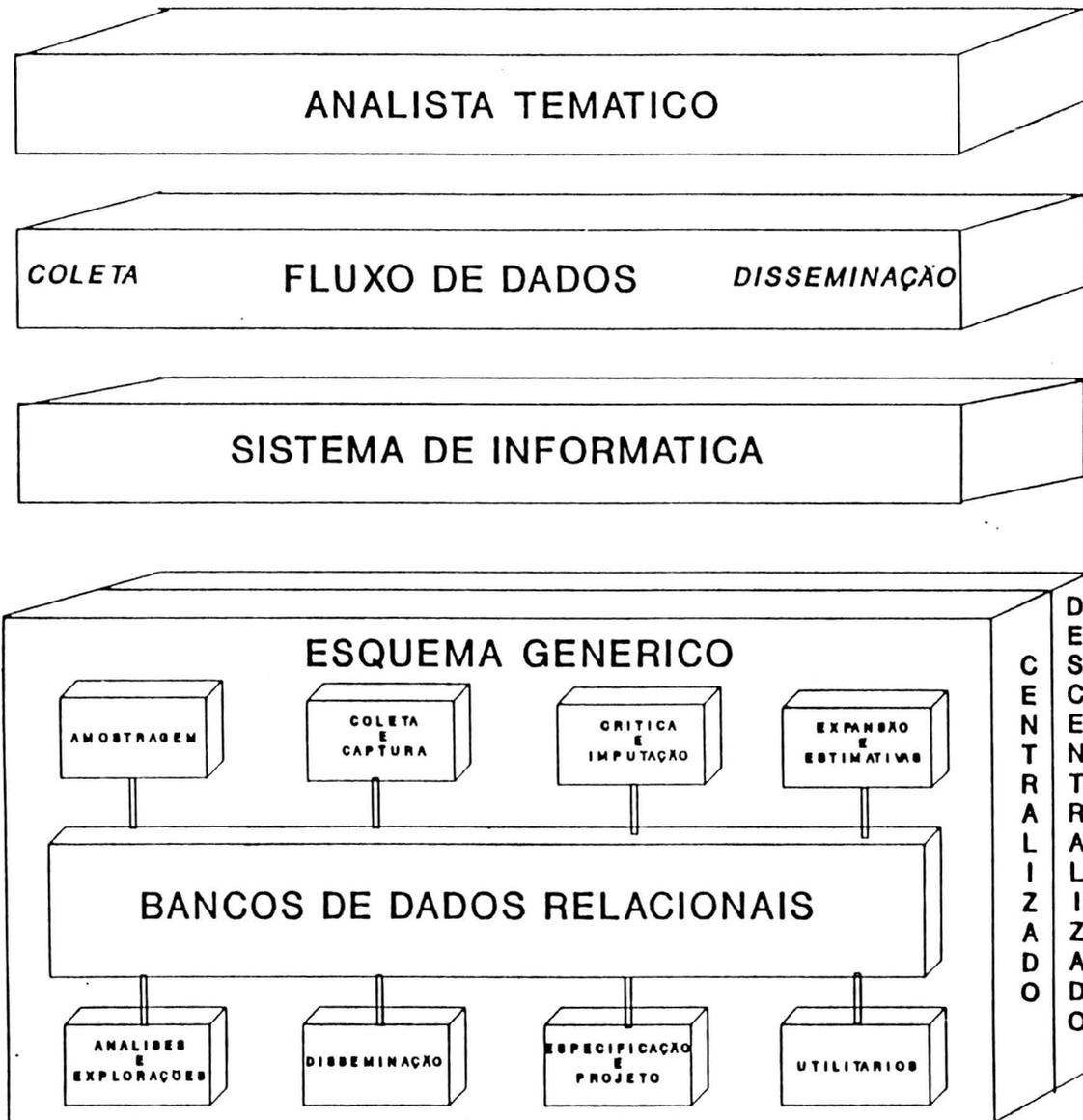


FIGURA 2

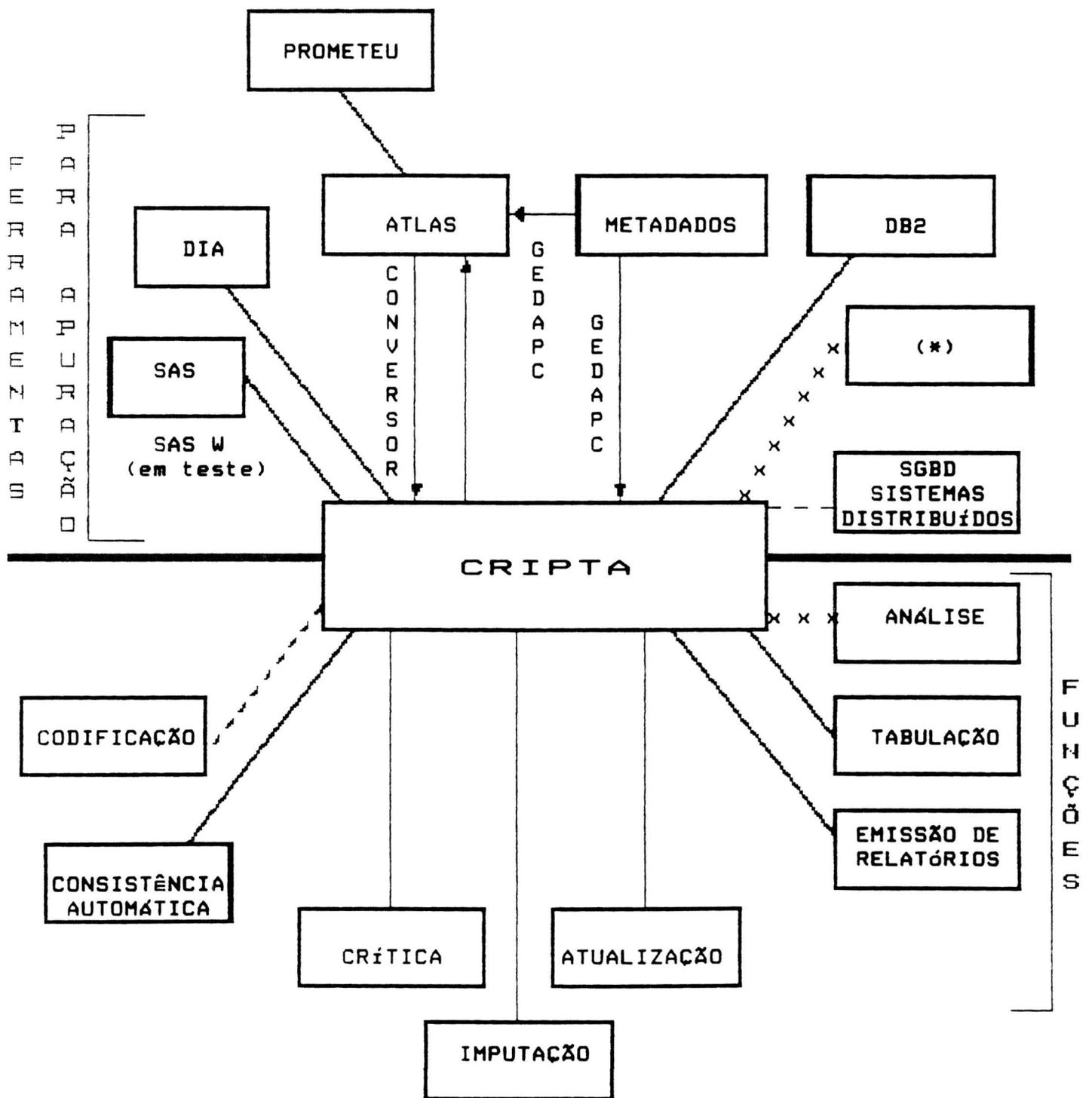


Figura 3

(\*) Pacotes para Sistemas Distribuídos (PC-DOS e UNIX), que possibilitam o processamento de dados estatísticos.

- Já existe
- - - Em desenvolvimento
- x x x Proposta de desenvolvimento