

# TERMISUL E INFORMÁTICA: CAMINHO PERCORRIDO E HORIZONTE A ALCANÇAR

Anna Maria Becker Maciel<sup>1</sup>

**Abstract:** *This paper reports on the quest for the interface between Terminology, Informatics and the Termisul Project. Three landmarks highlight the account, DOS, Windows, and Internet. They characterize the steps of a team of linguists towards the computerized era. As far as the technology advanced, their perception of digital resources changed from being mere operational help in data management to representing sophisticated tools assisting research and eventually to the demand of cooperative work with computer scientists. The difficulties of the early stages encountered in the path tread are remembered. The meaningful results of the efforts taken are described. The challenges of the ongoing research work are outlined.*

**Key words:** *terminology, terminotics, informatics, linguistic research, computerized research.*

## 1. INTRODUÇÃO

Os campos de interesse conjunto da Terminologia e da Informática não são novos e estão cada vez mais próximos. Já Wüster, quando definiu as características dos estudos terminológicos nos anos 60, imaginava uma grande rede virtual unindo terminólogos no espaço cibernético. Hoje, sua visão profética já não é um delírio, pois a Internet torna possível a ampliação de novos horizontes. Costuma-se também afirmar que Wüster (1974), ao enfatizar a multidisciplinaridade da Terminologia, mencionava, entre outras ciências, a Informática. No entanto, conforme adverte Picht (1996, p. 258-259), é preciso esclarecer que ele se referia à interface da documentação, informação e técnica informática e não à disciplina como atualmente a entendemos.

De fato, os terminólogos pioneiros, e o Termisul em seu encaixe, sonhavam com a ferramenta ideal para gerenciar, tratar, processar e armazenar informação especializada. Nesse contexto, surgiu um novo conceito e criou-se um novo termo, Terminótica, denominando a união da Terminologia com a Informática (GOUADEC, 1990, p. 4). Tal conceito situava-se mais no campo da Ciência da Informação do que na área das Ciências Exatas, uma vez que, seguindo os pressupostos normativos wüsterianos, concebia-se a Terminótica como um recurso para operacionalizar os procedimentos da pesquisa, garantir armazenamento e conservação para que as terminologias coletadas não sofressem transformações como as que ocorrem na evolução natural da língua comum.

Os termos das diferentes áreas eram estocados em sistemas informatizados, processados pelo terminólogo e depois recuperados pelo usuário. O que se esperava do informata nos primórdios da Terminologia moderna era a criação de um artefato robusto e confiável que garantisse a guarda e a consistência dos dados e oportunizasse recuperação fácil, ágil e unívoca da informação. Surgiram, a partir dos anos 60, os grandes bancos de dados terminológicos para auxiliar as tarefas do terminólogo e do usuário. Ao primeiro possibilitavam o gerenciamento dos

---

<sup>1</sup> Docente do Programa de Pós-Graduação em Letras, UFRGS e Pesquisadora do Projeto TERMISUL

termos coletados, enquanto ao segundo, tradutor ou redator, os bancos ofereciam a utilização de um gigantesco dicionário eletrônico.

Também o Termisul seguiu a tendência da época e iniciou um longo e interminável caminho em busca da informatização, recuperá-lo implica remontar à introdução do computador no Instituto de Letras. Hoje, ninguém acreditaria que, apenas há vinte anos, o PC na Universidade Federal do Rio Grande do Sul era uma máquina reservada às esferas de alto prestígio acadêmico e amplo conhecimento tecnológico. Mas nem por isso o Projeto desanimou. Como entusiasta da empreitada, coube à autora deste texto relatar como se desenvolveu uma trajetória que ainda continua na busca por um horizonte já não tão longínquo. Não são poucos os episódios pitorescos que ilustram essa história, mas esses não serão aqui relatados; ao invés disso, pretende-se demonstrar como uma equipe, ingênua e ignorante na Ciência da Computação, empreendeu, ao longo dos últimos vinte anos, a união da Terminologia e da Informática, pomposamente denominada Terminótica.

Este relato contemplará três grandes etapas de um árduo percurso que se estende da era do *DOS (Disk Operating System)* à era do *Windows* e daí à atualidade e que, se não revela aquisição do domínio da área, atesta a progressiva mudança de uma perspectiva de informatização. No início, o Projeto via o computador como um arquivo informatizado, uma máquina capaz de armazenar dados alfa-numéricos, referentes aos participantes da equipe, à constituição da biblioteca de referência, aos achados da pesquisa. Não tardou a ficar claro que tais tarefas eram insuficientes, evidenciando a falta de recursos semi-automatizados para gerenciar e tratar, não apenas guardar o que era coletado.

À medida que a tecnologia avançava, a visão da equipe se alargava e exigia mais do computador; além de ser simples máquina de estocar material, ele passou a representar um recurso auxiliar na manutenção, gerenciamento e tratamento dos dados e um meio de colocá-los ao alcance do usuário. Ao chegar o limiar do século XXI, o advento da grande rede mundial, a *web*, desvendou um horizonte cada vez mais amplo e mais próximo. Do interior da sala de reduzidas dimensões dentro do recinto do Instituto de Letras, era possível alcançar a imensidão do espaço cibernético: o conhecimento e a experiência de outros pesquisadores, as fontes documentárias essenciais inteiramente digitalizadas, todo um universo de informação encontrava-se ao alcance de um simples toque no teclado.

Diante disso, o Projeto procurou atualizar a metodologia de seus procedimentos de pesquisa e de divulgação de seu trabalho. As consequências dessa realidade foram duas: em primeiro lugar, a mudança da pesquisa manual em papel para a pesquisa em textos digitalizados; em segundo, a disponibilização *on-line* do fruto de trabalhos teóricos e aplicados da equipe. Nesse direcionamento, foram adotados como auxiliares de pesquisa *softwares* de análise textual, e as portas virtuais do Projeto se abriram para oferecer aos interessados a oportunidade de conhecer o material acumulado ao longo de vinte anos e experimentar ferramentas digitais básicas. Essa é a etapa do percurso que a equipe ora vivencia, o horizonte ainda não foi alcançado e nele vislumbramos a tão almejada parceria com pesquisadores da Ciência da Computação e não somente com técnicos em informatização.

Não se pode dizer que as diferentes etapas desta trajetória tenham sido estanques e nitidamente demarcadas; por vezes elas se sobrepõem, tornando impossível afirmar quando uma começa e outra termina; no entanto, suas características bem descrevem a evolução do uso dos recursos informatizados no Termisul. Essa evolução não termina aqui, o relato se conclui apontando as perspectivas futuras da pesquisa no Projeto a realizar-se nos próximos anos.

## **2. PRIMEIROS PASSOS**

Desde sua origem, o Termisul busca o suporte informatizado ideal. Ainda que carecendo de equipamento próprio e assistência especializada, a equipe ousou penetrar no universo sofisticado e misterioso da Informática. Sem possuir um computador e sem contar com o auxílio

de técnicos, professoras e Bolsistas de Iniciação Científica (BIC) do Instituto de Letras enfrentaram o desafio da informatização. Para tanto, analisaram as potencialidades dos mais variados programas informatizados à luz dos objetivos a alcançar, sempre enfrentando a precariedade dos meios financeiros e dos recursos humanos da situação.

Embora não tendo recebido nenhum treinamento específico, nenhuma espécie de letramento, para não dizer alfabetização, na área, fazendo uso apenas de dedicação, coragem e uma grande dose de ingenuidade, iniciou-se a informatização do Projeto em uma máquina emprestada nas horas vagas pela secretaria de um dos departamentos do Instituto de Letras. A sistemática da organização das atividades foi esboçada visando o registro de dados administrativos do grupo, fontes documentárias e dados terminológicos, tudo de maneira muito simples, em disquetes flexíveis de 5"1/4, com a "fantástica" capacidade de 160 a 320 Kb de armazenamento. Conservados em estojos plásticos, os disquetes deram origem ao primeiro Banco de Dados Termisul, BDT; levados de lá para cá dentro de uma bolsa, literalmente dando a essa sigla o significado de Bolsa de Dados Termisul. Nesse ínterim, três acontecimentos trouxeram novo alento à busca da informatização: a visita do Dr. Felix Mayer (1991), a doação de um computador e o estágio no Canadá de dois membros do Projeto, a coordenadora, em 1992, e a vice-coordenadora e autora deste texto, em 1993.

Dr. Mayer, então professor de Tradução e Interpretação da Universidade de Saarbruck, Alemanha, desvendou para a equipe, em um curso de extensão chamado "Terminografia Informatizada: fundamentação teórica e metodologia", as potencialidades que a utilização da tecnologia oferecia ao estudioso da Terminologia. Mostrou as maravilhas que a máquina podia fazer, quando equipada com um programa adequado, no caso o MultiTerm, Trados, e alimentada por pesquisadores treinados. Desde então, esse *software* comercial produzido na Suíça foi o sonho de consumo do Termisul. No entanto, se não foi possível adquiri-lo nessa época, ao menos o entusiasmo da equipe viu-se recompensado pela vinda do primeiro computador do Termisul, um IBM PC XT com sistema operacional MS-DOS, doado pelo Reitor Prof. Dr. Tuikson Dik.

A visita aos centros de excelência em terminologia do Canadá, graças à bolsa *Faculty Enrichment Program* concedida pela Embaixada do Canadá no Brasil, teve grande influência na história do Termisul não somente no que concerne à reflexão teórica, mas de maneira especial no que diz respeito à utilização da tecnologia. Os grandes bancos de dados terminológicos, o BTQ (*Banque de Terminologie du Québec*) e o *TERMIUM*, banco oficial do governo canadense em Ottawa, impressionaram enormemente as duas pesquisadoras do Termisul, que passaram a sonhar com um BDTermisul de maior alcance, integrado em uma futura estrutura de redes de âmbito local, nacional e internacional (MACIEL, 1993).

Na euforia da posse de um computador, procurou-se encontrar o programa que melhor atendesse às exigências de armazenamento e recuperação das informações, proporcionasse a construção de uma grade de apoio à pesquisa e que pudesse, acima de tudo, ser adquirido com os recursos de um parco orçamento. Finalmente, foi feita a opção pelo software MicroISIS, 2.3, uma adaptação para microcomputadores do sistema CDS/ISIS (*Computerized Documentation System/ Integrated Set of Information*), desenvolvido pela UNESO e distribuído gratuitamente no Brasil pelo IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia). Com as respectivas instruções de uso, o MicroISIS foi conseguido através do CPD (Centro de Computação de Dados) da Universidade.

Teve início, então, a capacitação da bolsista responsável e de sua orientadora, ambas amadoras, sem conhecimentos técnicos de computação e áreas afins, que se dispuseram a seguir cursos ministrados para bibliotecários em Porto Alegre e em Florianópolis. Como seu campo de trabalho e seus propósitos eram outros, elas tiveram de adaptar as funcionalidades destinadas ao gerenciamento de bibliotecas às necessidades de um incipiente projeto de pesquisa terminológica. Sua inexperiência, conjugada às dificuldades derivadas da interface pouco amigável e da complexidade do funcionamento do aplicativo, não impediu a implementação em caráter experimental do BDTermisul como um sistema de bases de dados relacionais constituído por sete minibases desenhadas "sob medida" para atender os objetivos do Termisul. As bases

eram as seguintes: TERMIN, LEGIS, BIB, BOLSA, RCHUM e EVENT; as duas primeiras eram destinadas às fichas terminológicas e às fontes documentárias respectivas, as outras, aos dados burocráticos e administrativos das atividades em andamento (GAVENSKI, 1993).

A primeira experiência de informatização do Projeto feita com o MicroISIS, em que pese a boa vontade da equipe, não se mostrou plenamente satisfatória, pois cedo apresentou dificuldades para o perfeito funcionamento e a adequação às exigências crescentes da pesquisa. Ao mesmo tempo, o estágio no Canadá confirmou o quanto a equipe carecia de conhecimentos na área da informática e enfatizou a urgência de recorrer aos especialistas. Compreendeu-se que não era mais possível continuar sem a assistência dos informatas e buscou-se socorro. Desde então, alunos da Ciência da Computação, da Matemática Computacional e da Engenharia Elétrica começaram a se engajar no Projeto como bolsistas, orientados pelos próprios terminólogos. Esses alunos constituíram o único suporte da área proporcionado ao Projeto em toda sua história; a seu comprometimento e paciência com linguistas “analfabetos” deve-se o desenvolvimento da informatização do Termisul.

Imediatamente os novos bolsistas procuraram contornar os problemas criados pelo amadorismo da informatização em curso. Tentaram utilizar os recursos do próprio sistema, utilizando o ISIS-Pascal, mas, diante da navegabilidade complexa e da interface não amigável do programa, propuseram a utilização da linguagem Clipper em uma modelagem de base de dados padrão xBase. Com a concordância da equipe, foi aceita a proposta. Definidos os campos de acordo com as fichas terminológicas elaboradas pelas pesquisadoras, iniciou-se a construção de um aplicativo “feito em casa”, verdadeiramente artesanal. Assim, nasceu o sistema de armazenamento, manipulação e recuperação de dados, TermIN-TermOUT (MACIEL, PARMEGGIANI, MALLMANN, PEREIRA, 1994).

TermIN era a base de trabalho, que oferecia uma estrutura fixa de ficha terminológica na qual os dados eram facilmente inseridos, modificados ou eliminados de acordo com o andamento da pesquisa. Os campos tinham extensão predeterminada com exceção do campo destinado à definição, que era capaz de comportar a coleção de informações necessárias para fundamentar a posterior elaboração do texto definitivo. As fichas eram interligadas por uma rede de remissivas e podiam ser acessadas de várias maneiras a partir de diferentes campos.

TermOUT era a base de leitura destinada à consulta do usuário, que podia receber o material da base de trabalho automaticamente, sendo os dados validados sempre que solicitados pelo administrador e nunca pelo usuário. A base dispunha de um sistema ágil e prático de acesso às informações, que buscava a ficha de um termo através de uma barra de rolagem e permitia a visualização de todos os campos relativos ao termo selecionado e suas referências cruzadas. Oferecia ainda um módulo de geração de relatórios para impressão de listagens totalmente configuráveis assim como permitia exibi-los na tela e/ou salvá-los em arquivo texto.

O resultado da construção do TermIN-TermOUT gerou muitas expectativas. Esperava-se que o sistema fosse utilizado para a divulgação do dicionário da terminologia jurídico-ambiental que estava em elaboração. Imaginava-se que uma versão vazia do sistema pudesse ser divulgada em disquetes rígidos de 3” ½, ou através de um sistema de redes. Conjeturava-se também distribuí-la a terminólogos e tradutores que desejassem criar suas próprias bases de dados, e chegou-se a elaborar um protótipo de aplicativo, o TermTrad. Em suma, o sistema alimentou muitos sonhos do Termisul.

A essa altura dos acontecimentos, às escondidas, contrariando a decisão unânime da equipe, esta autora (único voto vencido), com a cumplicidade dos bolsistas, coautores do programa, teve a ousadia de colocar uma versão do TermOUT no servidor Unix do Instituto de Informática. Assim, no endereço de um deles <[www.inf.ufrgs/~aap/termisul](http://www.inf.ufrgs/~aap/termisul)>, o primeiro dicionário eletrônico do Termisul esteve à disposição do usuário durante três anos, de 1993 a 1996, quando perdemos o privilégio da hospedagem. Diante do êxito da façanha, reconhecida desde Viena pela seção internacional de terminologia da INFOTERM bem como por outros terminólogos estrangeiros, a equipe rendeu-se à evidência do sucesso e festejou a oportunidade do convívio na esfera cibernética. Somente mais tarde, quando o Instituto de Letras pode compartilhar da Internet, o Termisul criou sua primeira página oficial no servidor Orion da universidade. Mas

isso já pertence à etapa seguinte desta história: a era do Windows.

### 3. A ERA DO WINDOWS

O segundo momento da trajetória aqui relatada se caracteriza pela crescente aplicação da pesquisa terminológica em produtos terminográficos e a mudança dos procedimentos desenvolvidos na coleta de dados, correspondendo a um período efervescente de demanda de recursos computadorizados. A essa época o Termisul havia adquirido dois microcomputadores PC-386DX, operando com o sistema operacional *Windows*. Os aplicativos em uso já não mais satisfaziam ao gerenciamento da pesquisa, que se tornava muito mais complexa, à editoração de seus resultados em publicações tradicionais e eletrônicas e, principalmente, à coleta das informações. Os grandes marcos dessa etapa do Termisul dominada pela tecnologia *Windows* foram a elaboração e o lançamento de dicionários e glossários e o desenvolvimento da pesquisa terminológica com *corpora* digitalizados.

A equipe não desistia da busca por um *software* que facilitasse os procedimentos de manutenção e a coleta de dados, embora estivesse consciente da utopia de encontrar a ferramenta ideal para o trabalho terminológico. Diversas tentativas de utilização de programas comerciais de gerenciamento e registro de informações, como FOXPRO e outros, frustraram as expectativas da equipe. Finalmente, a decisão pelo uso da linguagem *Delphi*, *Borland*, bastante difundida pela praticidade de seus recursos no desenvolvimento de programas visuais para *Windows*, trouxe novas esperanças.

Nesse momento, o trabalho a ser encarado era traduzir os dados do sistema TermIN-TermOUT para uma outra estrutura mais adequada à prioridade do Projeto naquele momento, a divulgação de seus produtos terminográficos. A organização das fichas terminológicas originalmente aceitas foi revisada, o número e a extensão dos campos foram aumentados para atender às características da informação cada vez mais complexa que era coletada. Ao mesmo tempo, a editoração de um dicionário papel e sua disponibilização em versão eletrônica juntamente com o livro tornou-se o principal alvo da informatização.

Tudo parecia ser automático e simples de fazer; no entanto, a realidade foi bem diferente, como conta o bolsista que venceu a batalha. Ele relata como o trabalho exigiu “muito esforço, reflexão, criatividade e paciência por parte de todos, da coordenação, dos bolsistas e principalmente daquele a quem competia a maior responsabilidade na tarefa de desenvolver satisfatoriamente o programa” (MORALES, 2001, p.365-6). Mas, superadas as dificuldades, o copião do *Dicionário de Direito Ambiental: Terminologia das Leis do Meio Ambiente*, de ora em diante mencionado como DIC, foi entregue à Editora da Universidade necessitando apenas a revisão gráfica final. Para alívio e satisfação da equipe, o livro, acompanhado de sua versão eletrônica, foi lançado triunfalmente em 1998. A versão eletrônica, denominada TermDIC, apresentava todo o conteúdo do dicionário, termos, ementas da legislação utilizada na coleta dos termos e referências bibliográficas consultadas e permitia as mais variadas formas de consulta, como pesquisa por termos completos ou segmentos, temas, ocorrências da legislação. Além disso, podia ser acessada em um computador IBM PC 486 ou compatível, com memória RAM 16Mb operando em ambiente *Windows 95* ou superior.

Gravar um CD-ROM nessa época era quase uma aventura; até parece mentira, mas não se encontrou onde fazê-lo na nossa Universidade, tampouco em Porto Alegre, e o original do TermDIC foi enviado em disquete pelo correio tradicional para uma firma em São Paulo. Mas a alegria não durou muito. A falta de experiência da equipe resultou em um defeito na programação, descoberto somente quando as cópias já estavam na Gráfica da Universidade prontas para serem inseridas na contracapa do livro. O programa não podia ser executado se instalado em um diretório diferente do diretório original e em sistemas que usassem línguas diferentes do português e inglês. Passado o susto e a decepção, procurou-se minorar o problema divulgando uma atualização corretiva na *homepage* do Termisul

<<http://orion.ufrgs.br/termisul/termdic.html>> no servidor da UFRGS.

Passada mais de uma década da publicação, o TermDic até hoje é motivo de orgulho para a equipe. Foi registrado no INPI (Instituto Nacional de Propriedade Industrial), sob o número 98003170 em 09/11/1998. Integra a produção científica da UFRGS, tendo a autora deste texto como coordenadora do sistema e o Bolsista de Apoio Técnico do CNPq, Diego Francisco de Gastal Morales, como autor do aplicativo e da interface. Devido à cessão dos direitos de divulgação em qualquer mídia à editora Lexikon, que publicou a 2ª edição do Dicionário, o TermDIC não pode ser disponibilizado na página do Termisul.

Ao tempo da revisão final do DIC, já quase a caminho da gráfica, foi adquirido o tão sonhado *MultiTerm Professional* que o Prof. Mayer nos apresentara em 1991. O aplicativo integra um conjunto de ferramentas para a construção de bases de dados terminológicas destinadas a auxiliar o tradutor bem como a facilitar a edição de glossários e dicionários. Comporta dados multilíngues, permitindo a criação de fichas terminológicas diferenciadas em até vinte idiomas de acordo com os propósitos do trabalho terminológico, terminográfico e tradutório a ser realizado. A recuperação da informação através de um sistema de buscas e filtros é ágil e amigável, capaz de ser manipulada pelo usuário sem maior prática em ferramentas digitais. No entanto, a operacionalização da construção das bases não é tão fácil assim e exige esforço, prática do uso de *softwares* de bases de dados e muita dedicação do responsável, ainda mais quando ele trabalha com uma equipe incansável em sugerir modificações na estrutura inicial.

O Multiterm suporta o trabalho em rede de vários operadores simultaneamente na mesma base de dados, aspecto altamente vantajoso para o Termisul que, na época da aquisição, planejava o *Glossário Multilíngüe de Direito Ambiental Internacional*, com a colaboração de uma professora e bolsistas da área jurídica, bem como de bolsistas de Letras, alunos de diferentes línguas (português, espanhol, francês e inglês). Inicialmente, a base foi planejada como uma ferramenta interna do Projeto para o tratamento de seus dados. Em seguida, pensou-se em também elaborar uma versão eletrônica do trabalho que ficaria disponibilizada para o usuário *on-line*. Tal decisão sobrecarregou o trabalho do bolsista de Informática que teve de contornar vários problemas de ordem técnica para garantir a integridade e segurança dos dados e aprimorar a interface (RECH, 2001, p. 367-371).

A publicação do *Glossário Multilíngüe* em formato papel tradicional foi realizada em 2004, mas sua disponibilização *on-line* não foi efetuada. Problemas relativos à cessão de direitos autorais do aplicativo comercialmente adquirido, embora usado para fins não lucrativos, impediram que o plano de colocar a obra na Internet chegasse a termo. Talvez agora que a versão do Multiterm então usada já foi ultrapassada e substituída por um sistema muito mais complexo e avançado em funcionalidade, seja possível uma nova tentativa, pois é uma lástima que um projeto de tal monta fique relegado ao esquecimento. A utilização do Multiterm no trabalho multilíngüe resultou positiva e levou a equipe a planejar outro projeto, dessa vez envolvendo mais uma língua, o alemão. Com apoio na experiência anterior, procurou-se planejar cuidadosamente a macro e microestruturas do *Glossário de Gestão Ambiental* em cinco idiomas; sua publicação data de 2007.

Completado o ciclo de três obras de referência da terminologia ambiental, o Termisul empreendeu um novo projeto de pesquisa com maiores e mais ambiciosas exigências de suporte informatizado. Mas, antes de tratar desse tópico, é preciso relatar o avanço nos procedimentos de coleta de documentos e de pesquisa textual que desde os meados dos anos 90 tivera lugar no Termisul e que também é um marco considerável na trajetória de informatização.

Quando a Internet alcançou o Instituto de Letras, a penosa busca manual em folhas impressas foi relegada ao passado e substituída pela consulta das fontes documentárias na tela do computador. Até 1996-97, os termisulenses folheavam as páginas do *Diário Oficial da União* (DOU) na seção Meio Ambiente, em busca de candidatos a termo jurídico-ambiental e faziam fotocópias do material relevante para submetê-lo à análise e discussão na reunião semanal da equipe. Pastas arquivo de cartolina tipo A-Z se multiplicavam nas prateleiras da sala 230B, constituindo o *corpus* textual concreto, não virtual da pesquisa. Além de usar o DOU como

fonte de termos, recorria-se também à alternativa de requisitar cópia de diplomas legais sobre determinado tema e esperar até que o serviço do PRODASEN (Processamento de Dados do Senado) atendesse ao pedido. Ambos os procedimentos exigiam que bolsistas e/ou professoras se deslocassem do Campus do Vale para a Biblioteca Central da UFRGS ou para a Assembleia Legislativa do Estado, um processo lento que consumia tempo e esforço.

A integração do Instituto de Letras na Internet permitiu acessar o próprio Senado Federal e baixar os documentos a qualquer momento, ou melhor, sempre que o sistema estivesse no ar. Várias vezes por dia o Senado era visitado pelos membros do Termisul, que também recorriam a outros organismos públicos no país relacionados com o meio ambiente, bem como à ONU e a *sites* oficiais dos países das línguas contempladas pelo Projeto. Assim, pouco a pouco, de grão em grão, utilizando simplesmente os comandos Copiar e Colar, foram coletados da Internet textos, um depois do outro, e formou-se o primeiro *corpus* de linguagem legislativa do Termisul, origem do que seria mais tarde organizado como a atual Base Legis (MACIEL, SILVEIRA, MARRA, 2008).

Diante do acúmulo de material coletado, era necessário que a máquina auxiliasse o pesquisador nos procedimentos de pesquisa. A leitura manual era impraticável, pois utilizar o comando Localizar do processador de texto *Microsoft Word* em um conjunto de algumas dezenas de textos na procura de candidatos a termo era moroso, ineficiente e injustificável quando se dispunha de computadores. Reconhecer e assinalar, a olho nu, com lápis colorido na mão, elementos definitórios para discussões posteriores ainda era pior. A assistência de um programa informatizado que agilizasse a operacionalização dessas e de outras tarefas era urgente. Diante disso, recorreu-se ao conjunto de ferramentas de análise textual *WordSmith Tools* 3.1 (SCOTT, 1999), que foi adquirido pela Internet e que passou a ser usado imediatamente.

A essa época, muito contribuiu para o avanço da informatização dos procedimentos de pesquisa do Termisul o Curso Livre ministrado pelo Prof. Dr. Tony Berber Sardinha, da Pontifícia Universidade Católica de São Paulo, sobre a abordagem de pesquisa linguística com *corpora* textuais na perspectiva da Linguística de *Corpus* e com a utilização desse *software*. Procurando cada vez mais conhecer o potencial de tal abordagem e tirar proveito de múltiplas funcionalidades do *WordSmith*, o Projeto passou a utilizá-lo como rotina preferencial de trabalho. Assistindo o pesquisador no reconhecimento de candidatos a termo e na identificação de padrões linguísticos, alinhando textos, auxiliando em inúmeras tarefas, o aplicativo continua contribuindo para o progresso da pesquisa terminológica do Projeto. Hoje, o Termisul utiliza, além dessa, outras ferramentas de análise textual, como o *Ant.Conc* 2.3.1 (ANTHONY, 2007), e cria ferramentas sob medida para atender às necessidades da pesquisa em andamento. Esse assunto será tratado no próximo tópico.

#### **4. TERMISUL NA WEB**

Como foi anteriormente relatado, a estreia do Termisul na *web* aconteceu em 1993, no servidor do Instituto de Informática, no endereço de um dos alunos bolsistas. A página, mantida durante três anos, apresentava breves informações sobre o Projeto, seus objetivos e sua equipe. Uma senha de *log-in*, fornecida gratuitamente, dava acesso a uma versão-teste do dicionário de Direito Ambiental, cadastrava os usuários e mantinha a equipe a par do número de visitantes da página, nome e instituição de origem. Na sua simplicidade, essa primeira página do Termisul pode ser considerada um marco na história da Terminologia não só na UFRGS como no Estado e muito envaideceu a equipe, embora tivesse sido inaugurada às escondidas. Quando o Instituto de Letras passou a integrar a Internet, em 1997, conectando seus computadores ao sistema geral da Universidade, o Projeto, no endereço <[www.orion.ufrgs.br/termisul](http://www.orion.ufrgs.br/termisul)>, teve a oportunidade de construir uma nova página. Era uma *homepage* de caráter informativo, divulgava a presença do Projeto no mundo acadêmico e em especial no novel universo daqueles que se dedicavam à

Terminologia no Brasil. Fornecia dados gerais sobre os projetos em andamento e sobre as atividades de seus membros participantes.

Em julho de 2006, o Termisul, tendo concluído os projetos de pesquisa e divulgação da terminologia da temática ambiental conforme se propusera, empreendeu um novo trabalho, construído sobre os projetos anteriores, com o objetivo de socializar mais ainda o conhecimento desenvolvido e a experiência vivenciada ao longo de 15 anos de reflexão e pesquisa. Nesse propósito, considerando o cabedal acumulado, fruto do esforço e cooperação de professoras e bolsistas, a equipe decidiu divulgar na Internet seu fazer e sua pesquisa teórica e aplicada. Para tanto, era preciso em primeiro lugar reformular sua presença na Internet. Esse foi o novo desafio planejado para o período de 2006-2010 e se intitulou Acervo Termisul: padrões da linguagem legal, normativa e tecnocientífica.

Mais uma vez, o Termisul enfrentou o desafio de substituir a antiga *homepage* por um *website* no endereço <[www.ufrgs.br/termisul](http://www.ufrgs.br/termisul)>, tarefa por demais ousada para professores e bolsistas do Curso de Letras, sem a devida formação técnica e em uma época de avanços vertiginosos da tecnologia da informatização. Não sendo possível contar com alunos da Ciência da Computação nesse momento, a tarefa foi feita com a prata da casa. Com grande interesse e comprometimento, dois bolsistas bastante familiarizados com o computador e seus mistérios se dispuseram para a difícil empreitada. Procurando subsídios aqui e acolá na Internet e junto a informatas da universidade, conseguiram o extraordinário feito de transformar a singela página de divulgação da existência do Termisul em um *website* dinâmico original e atraente, de ora em diante referido como o Acervo.

Ainda que dois bolsistas e suas respectivas orientadoras respondessem pela organização e informatização da página, o Acervo é uma realização de toda a equipe, professoras e estudantes do Termisul e sua implementação no ambiente virtual não teria sido realizada sem um grande esforço conjunto. O primeiro passo da tarefa foi definir, dentro da meta geral de socialização, o objetivo específico de estimular a pesquisa da linguagem usada pelos especialistas das mais diferentes áreas do conhecimento. O público-alvo foi caracterizado como egressos e estudantes dos cursos de Letras, linguistas, terminólogos, tradutores e redatores técnico-científicos. Nesse contexto, a equipe apoiou-se no quadro teórico fundamental da Terminologia Comunicativa e da Terminologia Textual, norte que orienta o Termisul desde sua origem, e buscou o suporte da Linguística de *Corpus*. Para alcançar o objetivo proposto de incentivar a pesquisa linguística, recorreu também aos pressupostos didático-pedagógicos da aprendizagem significativa sob o prisma da Educação a Distância.

Definidos o objetivo e o usuário, iniciou-se a divisão das responsabilidades entre dois grupos: o primeiro, encarregado da análise, seleção e tratamento do material disponível e o segundo, do planejamento do conteúdo a ser incorporado, desenho, arquitetura e funcionalidade do *site*. No entanto, no decorrer do tempo, toda a equipe passou a interagir; sugestões, palpites, detalhes e reformulações na estrutura e na funcionalidade eram continuamente discutidos, transformando a construção do Acervo em um verdadeiro trabalho colaborativo.

A proposta gráfica procurou favorecer a navegação do usuário, permitindo-lhe escolher o tema de seu interesse. O esquema geral da interface e a distribuição do conteúdo obedeceram ao formato de diagrama nos moldes de mapas conceituais. O conteúdo foi estruturado em blocos, e esses, divididos em seções contendo textos informativos interligados pelo sistema de hipertexto e dotados de janelas *pop-up*, favorecendo a busca dinâmica da informação. O primeiro bloco é destinado aos dados gerais sobre o Grupo Termisul; os outros apresentam os Objetos de Aprendizagem, a Biblioteca Virtual, os *corpora* textuais – denominados Bases – e as ferramentas; o último, uma amostragem do Projeto de pesquisa em andamento, Combinatórias Léxicas Especializadas – CLEs (BEVILACQUA; REUILLARD, 2010, p. 433-448).

Para oportunizar ao usuário a prática da pesquisa linguística informatizada, o Acervo oferece ferramentas para serem experimentadas nos *corpora* das Bases Textuais, que contemplam diferentes áreas temáticas e diferentes línguas; além do português, o Acervo conta com o alemão, o espanhol, o francês, o inglês e o italiano. São ferramentas básicas para procedimentos de pesquisa linguística: listadores de palavra, que podem ser levantadas individualmente, em



seqüências de n-itens (*n-gramas*) ou agrupadas ao redor de uma palavra chave (*clusters*); concordanciadores para textos analisados individualmente ou em pares de original e tradução. Seu manuseio é orientado por informações em janelas *pop-up*.

A implementação de tais ferramentas iniciou-se preliminarmente em caráter experimental em uma parceria com o Projeto TextQuim do Instituto de Letras. No entanto, como será relatado a seguir, passada a primeira fase do Projeto Acervo, em 2008, o Termisul conseguiu novamente a colaboração dos alunos do curso de Ciência da Computação, os quais assumiram integralmente a tarefa. Eles revisaram as ferramentas e reformularam suas construções, abandonando a programação em PERL pela programação em PHP, que está em plena operacionalização. Novas funcionalidades foram criadas; entre essas, devem ser citadas as que foram introduzidas no extrator de *n-gramas*, incluindo um gerador de *clusters* e oferecendo opções de filtragem dos dados (escolha de diferentes extensões de *n-gramas* e *clusters*, distintas frequências, possibilidade de inclusão de *stoplists*). A equipe considera essas melhorias como parte dos resultados já obtidos pelo novo projeto, uma vez que elas decorrem das diferentes situações encontradas na extração das CLEs.

Os novos responsáveis pelas ferramentas encontraram um trabalho artesanal, reformulado inúmeras vezes, com sérios defeitos estruturais, lapsos e equívocos, quase uma colcha de retalhos. Com empenho, trataram de aprimorar e criar novas funcionalidades de acordo com as exigências do desenvolvimento da pesquisa. No entanto, frequentemente se defrontaram com problemas não resolvidos que prejudicaram a obtenção do resultado esperado. Graças ao cuidado desses futuros profissionais da Computação, o *website* Acervo Termisul se mantém em plena operação e consegue se aproximar gradativamente da meta de incentivo à pesquisa terminológica e ao mesmo tempo executar os procedimentos indispensáveis ao avanço das atividades empreendidas pela equipe.

## 5. HORIZONTE A ALCANÇAR

Conforme foi aqui relatado, o Termisul, desde seus primeiros passos em 1991, procurou acompanhar os pioneiros da Terminologia, encarando o papel da Informática como um processo de informatização, isto é, um processo priorizando o uso de tecnologia informatizada. Tendo, desde o início, optado pela busca da Terminologia no real contexto de uso, o Projeto empreendeu a coleta dos termos em folhas impressas e formou um *corpus* de pesquisa de documentos copiados em papel. Sua informatização visava recursos computadorizados que pudessem agilizar os procedimentos administrativos, tais como estocar achados, organizá-los e colocá-los à disposição do usuário em produtos em formato tradicional ou eletrônico.

Atualmente, para o Termisul como para os terminólogos contemporâneos, o conceito de Informática vai muito além do conceito da utilização de recursos tecnológicos operacionais. De provedora de instrumentos úteis para o trabalho terminológico e para os fins terminográficos, a Informática agora é vista como parceira de pesquisas em pé de igualdade. Hoje, não se busca somente o fornecimento de serviços tecnológicos, tampouco se quer apenas suplementar os recursos linguísticos de que a Informática carece. Nessa perspectiva, a Terminologia procura uma aproximação em via de mão dupla que favoreça a contribuição teórica e aplicada de ambos os campos.

Entretanto, terminólogos e informatas ainda se olham com certa desconfiança, sem bem compreender o raciocínio, a metodologia e a relevância dos respectivos propósitos. A velha rixa racionalismo versus empirismo, problemática de diferentes escolas de pensamento, não foi completamente superada. Ideias preconcebidas assustam e, como outrora os troianos, todos têm medo dos gregos, mesmo que tragam presentes. Enquanto isso, a interface da Terminologia e da Informática que, de outro modo, poderia gerar troca de experiências, intercâmbio de pesquisas, aprofundamento de teorias e otimização aplicações, não é devidamente explorada.

Se em centros de excelência acadêmica, dentro e fora do Brasil, tal não acontece, em muitas

universidades, ao contrário, essa mentalidade é a regra e prejudica ambas as áreas. Ainda que tentem utilizar as ferramentas disponibilizadas na *web*, os linguistas sem formação específica de informática não conseguem operacionalizá-las de acordo com os propósitos visados, tampouco compreendem as estatísticas, as heurísticas e os algoritmos necessários para os procedimentos metodológicos a serem adotados em uma análise baseada em princípios das ciências exatas. Assim, sentem-se frustrados diante das longas e penosas tarefas manuais que são obrigados a enfrentar, despendendo esforços, tempo e comprometendo a objetividade de decisões diante dos dados encontrados.

O Termisul, em seu percurso de vinte anos, reuniu um cabedal de dados teóricos e empíricos que, embora já tenham gerado uma produção respeitável, ainda podem ser muito mais explorados. Seu projeto, ora em andamento, enfoca as combinatórias léxicas utilizadas nas linguagens especializadas, as CLEs, (BEVILACQUA; MACIEL; REUILLARD; SCHEEREN, 2010b), tema que tem merecido a atenção da equipe como atestam teses, dissertações e artigos produzidos incorporados à Biblioteca Virtual do Acervo. A pesquisa avança, muitas tarefas podem ser assistidas pelo computador, outras exigem o trabalho manual de professoras e bolsistas. Profundas e detalhadas análises são procedidas nos *corpora* na busca de itens lexicais candidatos a CLE. A pesquisa essencialmente linguística acumula dados preciosos para o Processamento da Linguagem Natural (PLN) e poderia ser compartilhada por outros projetos que trabalham o mesmo objeto sob a perspectiva da Ciência da Computação.

Com efeito, o tema das colocações ou fraseologismos nas áreas especializadas atrai, cada vez mais, a atenção dos pesquisadores, a tal ponto que eventos inteiramente dedicados à sua problemática se multiplicam ano a ano. A eles acorrem os informatas em grande número e alguns linguistas ousados. Nessas ocasiões, no entanto, cumpre observar, *workshops*, ciclos de palestras, cursos de curta duração, encontros, manuais, *papers*, demonstrações descortinam horizontes, mostram caminhos, mas não oferecem o apoio estável e continuado de que os linguistas sem formação nas Ciências Exatas carecem. Por outro lado, parece que, não raras vezes, nas pesquisas empreendidas sob a ótica da Linguística Computacional, a participação do linguista é periférica, limitando-se a fornecer os dados ao informata que, como em um passe de mágica, os transforma em fórmulas e submete a estatísticas. Na lista das referências bibliográficas de publicações sobre PLN e temas afins, as obras de cunho linguístico são escassas, aparecem em ínfimas proporções ou em fontes secundárias. A autoria principal do texto normalmente pertence ao informata, não raras vezes no conjunto dos autores, encontra-se apenas um pesquisador da área de Letras.

Nesse contexto, a Terminologia propõe à Informática uma cooperação no sentido pleno de operação conjunta paritária formalizada entre terminólogos e informatas. Nessa cooperação, os dois grupos têm o mesmo nível de importância e responsabilidade, compartilham tarefas, repartem achados, trocam recursos, dividem experiências e buscam metas paralelas e, ao mesmo tempo, mantêm sua própria individualidade, seus pressupostos teóricos e propósitos aplicados. Tal relacionamento com a Informática, o Termisul vislumbra no horizonte ao completar vinte anos da busca da informatização. As pioneiras da “velha guarda” iniciaram um percurso que a nova geração deverá levar até o horizonte.

## BIBLIOGRAFIA

- ANTHONY, L. AntConc 3.2.1 w Disponível em  
<[http://www.antlab.sci.waseda.ac.jp/antconc\\_index.html](http://www.antlab.sci.waseda.ac.jp/antconc_index.html)> Arquivo acessado em 10/03/2009.  
BEVILACQUA, C. R.; REUILLARD, P. C. R. Acervo Termisul: padrões da Linguagem legal, normativa e científica. In: ISQUERDO, A.N.; FINATTO, M.J.B. *As ciências do léxico: lexicologia, lexicografia, terminologia*. Campo Grande: Editora UFMS; Porto Alegre: Editora UFRGS, 2010a. p. 433-448.  
BEVILACQUA, C. R.; MACIEL, A. M. B.; REUILLARD, P. C. R.; SCHEEREN, C.M.

- Combinatórias Léxicas Especializadas: etapas prévias para identificação e tratamento. In: *XII Simposio de Riterm, 2010*. Buenos Aires: (no prelo), 2010b.
- GAVENSKI, M. M. MicroSIS: uma experiência no gerenciamento de dados terminológicos. *Cadernos do IL*, Porto Alegre: UFRGS, p. 141-151, jul. 1993.
- GOUADEC, D. *Terminologie: constitution des données*. Paris: AFNOR, 1990.
- MACIEL, A. M. B. Termisul e Terminótica. *Cadernos do IL*, Porto Alegre: UFRGS, p. 133-139, jul. 1993.
- MACIEL, A. M. B.; PARMEGGIANI, A. A.; MALLMANN, F.; PEREIRA, G. A. Banco de dados de termos jurídico-ambientais. In: *Simpósio Iberoamericano de Terminología Riterm 4*, "Terminología y Desarrollo", Buenos Aires: Secretaria de Ciencia y Tecnología de La Nación, Subsecretaria de Informática y Desarrollo, Unión Latina, s/d. v. 2, p. 133-136, 1994.
- MACIEL, A. M. B.; MARRA, D. N.; SILVEIRA, D. A. Base Legis: site multilíngüe da linguagem legal. In: TAGNIN, S. E. O.; VALE, O. A. *Avanços da Lingüística de Corpus no Brasil*. São Paulo: Humanistas, 2008. p. 193-245.
- MORALES, D. F. G. TermDIC: um dicionário eletrônico. In: KRIEGER, M. G.; MACIEL, A. M. B. (orgs). *Temas de Terminologia*. Porto Alegre: Editora da Universidade/Humanitas, 2001. p. 364-366.
- PICHT, H. Em Record d'E. Wüster: La multidisciplinarietat de la terminologia. In: CABRÉ, M.T. (dir). *Terminología: selecció de textos d'E.Wüster*. Barcelona: Universitat de Barcelona, Servei de Llengua Catalana, 1996. p. 225-280.
- RECH, R. O. A informatização do Glossário Multilíngüe de Direito Ambiental Internacional: uma experiência com o MultiTerm. In: KRIEGER, M. G.; MACIEL, A. M. B. (orgs). *Temas de Terminologia*. Porto Alegre: Editora da Universidade/Humanitas, 2001. p. 367-371.
- SCOTT, M. WordSmith Tools 3. Oxford: Oxford University Press, 1999
- Disponível em: <http://www.lexically.net/downloads/version3/download.htm> Arquivo acessado em 07/07/2003.
- WÜSTER, E. Die allgemeine Terminologielehre - ein Grenzgebiet zwischen Sprachwissenschaft, Logik, Ontologie, Informatik un den Sachwissenschaften. *Linguistics*, n. 19, p.61-106, 1974.