

# Curriculum vitae Pieter Masereeuw

## Algemene gegevens

naam Masereeuw, Pieter Christiaan  
geboren 5 november 1957 te Amsterdam  
burgerlijke staat gehuwd, 4 kinderen  
adres Anna van Saksenstraat 25; 1901 TH Castricum  
telefoon 0251-670964, 06-20786265  
e-mail [pieter@masereeuw.nl](mailto:pieter@masereeuw.nl)  
zie ook <http://www.masereeuw.nl>  
foto



## Opleiding en ervaring

middelbare school 1970-1976. gymnasium-alfa  
(Frans, Duits, Engels, Grieks,  
Latijn, Nederlands,  
Geschiedenis)

kandidaatsexamen Universiteit van Amsterdam,  
1976-1981, cum laude  
(klassieke taal- en letterkunde)

doctoraalexamen Universiteit van Amsterdam,  
1981-1988, cum laude  
(hoofdvak Latijnse taalkunde,  
bijvakken informatica en  
algemene taalwetenschap)

programmeer- en scriptingtalen Java, Javascript, C++, C, Pascal/  
Delphi, Prolog, Smalltalk, Perl,  
Unix shells, JSP, PHP, ACL  
(Arbortext Command Language),  
8086-Assembler

platforms Unix (Solaris, Ultrix, AIX, Linux),  
VMS, AOS/VS, Macintosh,  
Windows NT

formalismes SGML, HTML, XML, DITA, XSLT,  
SVG, XSLT, XSL-FO, XML  
Schema (XSD), lex/yacc,  
javaCC, SQL, ..

## Overzicht vorige betrekkingen

Vanaf medio januari 2008	Zelfstandig als freelancer via <i>ZZP Oké Pieter Masereeuw</i> . Mijn klanten: Sanoma Learning, Uitgeverij Malmberg, Ten Brink Offset, NEN (Nederlandse normen), Ambrac/Boom Uitgevers Den Haag, Ambrac/Kluwer, D-reizen, Uitgeverij Pegasus, Politieacademie, Nationale Politie, DCT (o.a. voor Vanderlande Industries, DAF, Nedtrain en Damen Shipyards), CAK (via Informaat), Instituut voor Nederlandse Lexicologie, iCtrl (voor Lloyds Register Rail Europe en voor Nedtrain), Uitgeverij Springer/BSL, The Docworkers, SDU, Van Dale Lexicografie, Uitgeverij Thieme, Logica (voor ING/Postbank), TNT Post Electronisch zakendoen,
1/1/2004-31/1/2008	Coördinator ICT bij Alfabase te Alphen aan den Rijn. Tevens 1 dag per week free-lance werk.
1/11/2000-31-/12/2003	Teamleider IT bij PlantijnCasparie Data te Heerhugowaard (4 dagen per week). Tevens 1 dag per week free-lance werk. Einde betrekking wegens bedrijfssluiting.
1/06/1999 - 31/10/2000	Software engineer bij Ambrac B.V. te Utrecht.
1/01/1997 - 30/05/1999	Systeemanalist/programmeur bij het Instituut voor Nederlandse Lexicologie (INL) te Leiden.
1/12/1995 - 31/12/1996	Systeemontwikkelaar bij Informaat te Baarn.
1/11/1986 - 1/12/1995	Systeemanalist/programmeur bij het Computercentrum Letteren van de UvA.
1/03/1984 - 1/11/1986	Technisch programmeur bij twee ZWO-projecten (conversie van het Longman woordenboek en toepassing daarvan bij het automatisch ontleden van het Engels)
1/05/1983 - 1/07/1985	Programmeur/systeembeheerder bij het Computercentrum Letteren (CCL) van de UvA

1/09/1979 - 1/09/1983	Assistent bij de vakgroep Latijn van de Universiteit van Amsterdam (UvA) met als taak het vervaardigen van programmatuur voor het ontsluiten van Latijns corpusmateriaal
1/09/1979 - 1/03/1984	Studentenbaantje: PTT Telefoondistrict Amsterdam (internationale telefoonverbindingen)

### **Zelfstandig als freelancer via *ZZP Oké Pieter Masereeuw*, vanaf medio januari 2008**

Als freelancer probeer ik zoveel mogelijk opdrachten te krijgen die aansluiten bij mijn ervaring en belangstelling. Die liggen vooral op het gebied van de teksttechnologie. U kunt hierbij denken aan:

- Projecten op het gebied van de ontwikkeling van software voor redactionele activiteiten (waaronder invoerhulpmiddelen, conversie en DTD/Schema-ontwerp)
- Ontsluiting van tekstbestanden (zoeken in vrije tekst, woordenboeken)
- Conversie naar publicatieformaat, zowel print/druk als web. Bij conversie naar print- en drukwerk werk ik samen met specialisten op het gebied van InDesign, FrameMaker, XyVision en 3B2. Ook verzorg ik zelf opmaakprogrammatuur middels XSL-FO.

Bij dit soort projecten maak ik bij voorkeur gebruik van open standaarden (waaronder XML, SGML, XSLT, XSL-FO, SVG en XHTML). Daarnaast maak ik voor het ontwikkelen van applicaties bij voorkeur gebruik van ontwikkelomgevingen die toelaten dat mijn software, inclusief gebruikersinterface, op meerdere platforms kan draaien. Daaruit volgt een voorliefde voor de programmeertaal Java en Java Swing.

Tot mijn recente opdrachtgevers behoren:

- **KOOP (Kennis- en Exploitatiecentrum Officiële Overheidspublicaties)**. Beheerder en ontwikkelaar van diverse overheidswebsites en zoeksystemen van overheidspublicaties. Gebruikte technieken: XML, Schema, XSLT, XSLWeb, Solr Cloud, Interactive XSLT, JSON, Apache configuratie, webservices (REST), low level HTTP.
- **Uitgeverij Malmberg** - het programmeren van een conversie (XSLT) voor een wiskundeonderwijsmethode; tevens ontwikkeling van XML-Schema, Schematron en

XQuery-scripts. Het schrijven van conversiesoftware om Word- en andere Officebestanden om te zetten naar het eigen XML-formaat van Malmberg/Sanoma. Ontwikkelen van Javasoftware voor een intelligent ZIP-tool en het beheer van persistente identifiers.

- **Van Gogh Museum** - het schrijven van een conversie van Wordbestanden naar TEI voor de dagboeken van Jo Bonger, de echtgenote van Theo van Gogh. Tevens ontwikkelen van redactionele hulpmiddelen in Oxygen XML Editor.
- **International Baccalaureat (IBO)** - het maken van een conversie van Wordbestanden naar JATS XML en het ontwikkelen van een webapplicatie voor de bediening daarvan.
- **EZ-base** - aanpassen van XSLT-stylesheets voor het renderen van XML-bestanden voor catalogi naar PDF middels XSL-FO.
- **Instituut voor de Nederlandse Taal (INT)** - het maken van een Rich Internet Application voor de historische woordenboeken van het instituut, op basis van Interactief XSLT (Saxon-JS) en Bootstrap. Zie <http://gtb.ivdnt.org>. Daarnaast veel conversiescripts op basis van XSLT en XProc voor het converteren van bestanden naar TEI-formaat.
- **Juridisch Woordenboek Spaans** (privé-financiering) - ontwerp van een woordenboekschema (XSD), plugin voor Oxygen XML Editor (Oxygen Java API, XSLT, CSS) voor een redactie zonder al te grote computervaardigheden, inrichten ExistDB om dagelijks automatisch rapporten te genereren over de stand van zaken van het redactieproces.
- **Fryske Akademy** - inrichten/aanpassen van een redactionele omgeving voor een Nederlands-Fries woordenboek (Java, MySQL, Schema, XSLT) en het ontsluiten van dat woordenboek middels een webinterface.
- **Sanoma Learning** - conversie van XML naar JSON t.b.v. onderwijssoftware.
- **DCT** - Automatische conversie van een grote verzameling officebestanden (Word, Excel) naar DITA om daarna de gegenereerde DITA-bestanden via de DITA Open Toolkit om te zetten naar grote PDF-bestanden.
- **NEN** - het ontwikkelen van een XSLT stylesheet om NEN documenten in HTML weer te geven.
- **Ten Brink Offset** - het helpen bij het ontwikkelen van diverse order-intake- en workflow-systemen, vooral in Java.
- **Ambrac/Boom Uitgevers Den Haag** - het programmeren van een Solr indexeersysteem op basis van Jenkins, rsync en de Linux shell.

- **Ambrac/Kluwer Law International** - het reverse engineeren van de werking van scripts binnen het content management-systeem Sigmalink.
- **Pegasus Uitgeverij** - technische XML-ondersteuning bij verschillende woordenboeken.
- **Politieacademie**- het maken van een DITA- (XML-) specialisatie voor de Politieacademie. Het inrichten van een XML-editor (XMetaL, Oxygen), het converteren van bestaande teksten naar het gespecialiseerde DITA-formaat. Meewerken aan de selectie van een contentmanagementsysteem.
- **DCT** - Proof of concept: publiceren van Dita Topics via Sharepoint en Dita Exchange voor een landelijke energiedistributeur.
- **CAK** (via Informaat en BeInformed): XSLT-werk (inclusief CSS en wat Javascript) aan het BeInformed-webportaal van het CAK en het genereren van standaard-brieven en documenten (XSL-FO).
- **DCT** - Maken van een Javascriptapplicatie voor Damen Shipyards om op zee te navigeren (!) door de bij het schip en de scheepssystemen behorende documentatie.
- **DCT** - Aanpassen van de DITA Open Toolkit (Webhelp en PDF via XSL-FO) voor Damen Shipyards.
- **DCT** - Het maken van een DITA specialisatie voor Nedtrain en het aanpassen van de DITA Open Toolkit.
- **DCT** - Wijzigen van complexe XSL-FO tabellen voor Vanderlande Industries.
- **INL** (Instituut voor Nederlandse Lexicologie)
  - Advies over het gebruik van DITA en DITA editors binnen een groot taalkundig project
  - Conversie van talloze bestandsformaten (waaronder kranten) naar de TEI DTD
  - Het bouwen van een user interface voor het interactief corrigeren van TEI documenten met behulp van een Java Swing interface
- **iCtrl** - het bouwen van een op componenten gebaseerde XML-redactieomgeving voor de treinonderhoudmanuals van Nedtrain. Het systeem gebruikt XMLMind (XML editor), Ovidius TCToolbox (CMS), XSLT 2.0, CSS, Java en Prince XML (voor PDF-rendering). De online-variant is voorzien van een HTML interface dat met behulp van Javascript/JQuery vergelijking van nieuwe en oude versies mogelijk maakt.
- **iCtrl** - ontwikkelen van een XML-Schema en inrichten van een XML-editor (XMLMind) met Java, ten behoeve van het opschonen van een database bij een toeleverancier van de Nederlandse Spoorwegen.
- **Springer/BSL (Bohn Stafleu van Loghum)** - XML conversie naar het ePub-formaat voor e-readers - het maken van een tool dat die conversie doet, voorzien van een Java Swing-gebruikersinterface.
- **Van Dale Lexicografie**
  - advies op technisch-redactioneel gebied;

- schema-ontwikkeling; conversie van diverse interne formaten (met XSLT 2.0) naar één XML-schema (W3C XSD);
- Interactieve tools voor de exploratie van grote woordenboekbestand, op basis van SAX en Java Swing
- **BSL/Springer** - Software voor het maken van e-books in het ePub-formaat
- **The Docworkers**
  - conversie van 3B2-bestanden naar XML;
  - ontwikkelen van een XML-roundtrip (Word-DocBook-Word);
  - Software voor het maken van e-books in het ePub-formaat met een Java Swing-userinterface.
- **SDU** - DTD's en upconversie voor de Nederlandse elektronische regelgeving;
- **Thieme** - oplossen van een fout in een Apache Cocoon pipeline (reparatie van een XSLT component);
- **Logica (for ING/Postbank)** - bugfixes in de software voor de integratie van ING en Postbank (software geschreven in C++ en Easel);
- **TNT Post Electronisch Zakendoen** - automatisch van een demo-website voor elektronisch factureren (Java, JBoss, Seam), ondersteund door een Java Webstart Swing-applicatie;
- **Sheridan Europe** - Java programmeren voor het distribueren van POD-opdrachten over printbedrijven in de hele wereld. Hierbij werd gebruik gemaakt van low level Java API's waaronder mail, ftp, pdf-manipulatie/ itext en Java Swing voor Webstart GUI's.

## **Alfabase Cross Media Solutions, 2004-januari 2008**

Alfabase Cross Media Solutions is onderdeel van de Amerikaans-Europese Sheridan-groep. Het is een grafisch bedrijf dat zich van drukkerijen en printing-on-demand-bedrijven onderscheidt door geavanceerde prepublishingactiviteiten en een hoge mate van automatisering. Een andere vorm van automatisering heeft betrekking op automatische order-intake en -verwerking.

Prepublishing bij Alfabase heeft vooral betrekking op het volledig automatisch (dus zonder menselijke tussenkomst) omvormen van gestructureerde bestanden (vooral XML en SGML) naar aantrekkelijk vormgegeven PDF-bestanden. Een andere activiteit is het slimmer en efficiënter maken van het (handmatige) DTP-proces, vooral door middel van scripting en hulpmiddelen voor redactionele samenwerking via internet.

Automatische order-intake en -verwerking gebeurt vooral ten bate van printing-on-demand-activiteiten van andere bedrijven binnen de Sheridan-groep. Het is een internationale activiteit, waarbij klanten (wereldwijd opererende uitgevers) hun bestanden op een centraal punt aanbieden, waarna de productie over de wereld (Azië, Europa, Amerika) verdeeld wordt om te besparen op distributietijd en -kosten.

Mijn taak binnen Alfabase en Sheridan was het coördineren van ICT-ontwikkelactiviteiten. Dat werk bestond uit een aantal onderdelen:

- Het verrichten van onderzoek op het gebied van prepublishing, automatische PDF-manipulatie en opdrachtverwerking.
- Het geven van technische leiding aan een groep van software-ontwikkelaars.
- Als lid van het managementteam bijdragen aan het beleid van Alfabase en de Sheridangroep met betrekking tot prepublishing en workflow-automatisering; op beide terreinen werk ik op technisch gebied nauw samen met ons Duitse zusterbedrijf [CrossMediaSolutions](#).
- Het zelf uitvoeren van ontwerp- en ontwikkelwerk; een belangrijk product was xzPages, dat bestond uit een combinatie van een op Swing gebaseerde Java Webstart client en een verzameling Unix server scripts voor automatisch typesetten.
- Ondersteunen van commerciële activiteiten. Dit betreft met name het geven van demonstraties en het overleggen met de klant over het inzetten van automatisering bij zijn redactionele activiteiten.

### ***Free-lance activiteiten, één dag per week***

Eén dag per week hield ik mij bezig met zelfstandige activiteiten. Hieronder vallen:

- Cursussen geven over XML en daarop gebaseerde standaarden.
- Webapplicaties (in JSP) voor de beschikbaarheidsregistratie van invalkrachten op basisscholen; verkocht aan verschillende scholenstichtingen.
- Het porteren (van C, Pascal, lex en yacc naar JavaCC en Java) van een applicatie voor de fonologische analyse van taaluitingen van kinderen met spraakproblemen. Deze applicatie zal via Java Webstart aan klinische logopedisten ter beschikking worden gesteld; het gaat hier om de voortzetting van de zgn. FAN-programmatuur die ik bij de Universiteit van Amsterdam heb ontwikkeld en waarvan ik later de rechten heb verkregen. Zie ook [hieronder](#).

- Programmeerwerkzaamheden in Delphi, C, C++ en Java (onder andere heb ik de bij de UvA ontwikkelde parser [Atlas](#) van Unix naar Windows geporteerd en voorzien van een Java Swing-gebruikersinterface).
- Diverse kleine webapplicaties, zoals scoreverwerking voor de Noord-Hollandse dartbond (in PHP) en het maken van een GUI-tool (in Swing) dat het loten van tegenspelers regelt..

### ***Werk-gerelateerde nevenactiviteiten***

- Redacteur van <!ELEMENT, het orgaan van de SGML/XML Users Group Holland).
- Lid van de SIG (Special Interest Group) over XSL-FO van de bovengenoemde Users Group.
- Mede-organiseren van een workshop over up-conversie naar XML met behulp van formele grammatica's, tijdens het SGML/XML-congres 2006.
- Organisatie van een workshop over XSL-FO op de technische dag van het SGML/XML-congres 2004.
- Organisatie van een workshop over [Apache Cocoon](#) op de >technische dag van het SGML/XML-congres 2003.

### **PlantijnCasparie Heerhugowaard, 2000-2003**

PlantijnCasparie Heerhugowaard was een bedrijf dat in grote mate vergelijkbaar is met mijn huidige werkgever Alfabase. Het beschikte, naast een drukkerij, over een grote pre-publishing- en automatiseringsafdeling. Door de malaise in de grafische industrie werd PlantijnCasparie Heerhugowaard helaas door de moedermaatschappij opgeheven. Ik ben voor de definitieve bedrijfssluiting weggegaan.

Bij PlantijnCasparie was ik teamleider van een team van ontwikkelaars en een web-ontwerper. Verder was ik als programmeur ook inhoudelijk betrokken bij projecten.

Via detachering was ik ook betrokken bij conversieprojecten van Kluwer Alphen/Deventer en bij het ontwikkelen van een internetrouteplanner voor Rijkswaterstaat (die rekening hield met file-verwachtingen).

### ***Werkzaamheden***

Bij PlantijnCasparie hielden mijn team en ik ons bezig de volgende soorten werkzaamheden:

- *E-commerce*  
De webapplicaties werden aanvankelijk ontwikkeld in



JSP met XML, maar later met behulp van Apache Cocoon.

- *Woordenboeken en encyclopedieën op CD-ROM en internet*

Een voorbeeld is de CD-ROM die behoort bij het woordenboek Latijn-Nederlands (hierop is de website <http://www.latijnnederlands.nl> gebaseerd. Bij het maken van webapplicaties maakten wij gebruik van Java Server Pages (JSP) en veel projecten van Apache: Apache Tomcat, Apache Cocoon, Apache FOP en Apache Lucene.

- *SGML/XML-applicaties*

Het gaat hier enerzijds om programmatuur die als hulpmiddel diende bij redactionele activiteiten en anderzijds om programmatuur die automatische opmaak mogelijk maakte. Automatische opmaak werd verricht met een traditioneel zetsysteem (3B2), maar ook (experimenteel) met XSL-FO.

- *Order-intake*

Aanpassen van een interactief (Java Swing) programma waarmee order managers de procesgang van de SGML- en PDF-verwerking van wetenschappelijke tijdschriften konden regelen

## **Ambrac, 1999-2000**

Ambrac is een bedrijf dat diensten verleent op het gebied van XML en SGML. Bij Ambrac hield ik me bezig met de volgende activiteiten:

- *Content-management*

Het inrichten van het content-managementsysteem *Sigmalink* voor Kluwer Law International te Den Haag.

- *Conversies*

Enerzijds het ontwikkelen van tools om het doel van conversies te vergemakkelijken, anderzijds het uitvoeren van conversies met behulp standaard-tools als Perl en (toen nog in ontwikkeling) XSLT.

- *Redactionele tools*

Maken van scripts in *Adept\*Editor*, voornamelijk bij het opzetten van een redactie-omgeving bij Samsom uitgeverij (nu Kluwer Alphen).

## **Instituut voor Nederlandse Lexicologie (INL), 1997-1999**

Bij het INL heb ik me als enige "zware" ontwikkelaar beziggehouden met de volgende soorten programmatuur:

- *Corpus-retrievalprogrammatuur*  
Bij het INL heb ik me beziggehouden met het ontwikkelen van een EU-gefinancierd nieuw corpus-retrievalprogramma. Bijzondere eigenschappen van dit programma waren:
  - Client/server: de client in Java (applets), het serverdeel in Java en C++;
  - Object-oriëntatie; hierdoor is een grote mate van parametrizeerbaarheid verkregen. Zelfs de vraagtaal is parametrizeerbaar;
  - Java-applets (voor de client);
  - Servlets (toen zeer nieuw; JSP was er nog niet);
  - Gedistribueerde applicaties: de servlet geeft via sockets opdrachten door aan processen op andere machines;
  - Multithreading, zowel in C++ (Posix-threads) als in Java;
  - SGML/XML-aware zoeken;
  - > 40 miljoen woorden;
  - Combinatie van zoeken met behulp van bitvectoren en reguliere expressies.
- *Conversie- en validatiesystemen*  
Speciaal voor het doen van conversies ontwikkelde ik de scripttaal *Taggle* - dit omdat het voor de hand liggende bestaande tool *Perl* minder geschikt is voor operaties die zich moeten afspelen binnen specifieke domeinen in een tekst. De Taggle-taal bleek ook handig bij HTML-conversie en CGI-scripts. Taggle is geschreven in Gnu Bison (yacc), flex en Gnu C. Er zijn versies gemaakt voor Sun/Solaris, Linux en (experimenteel) Windows NT. Een Taggle-script werd gecompileerd; daarbij werd, via C, native code geproduceerd.
- *Lexicografische programmatuur*  
Nieuwe versies van het Groene boekje (Woordenlijst Nederlandse taal) werden grotendeels bewerkt door middel van Taggle-scripts.

## **Informaat, 1995-1996**

Bij Informaat werkte ik mee aan een systeem (*Dox*) waarbij bedrijfsinformatie op een gecontroleerde manier werd ingevoerd

om vervolgens naar verschillende media (onder andere drukwerk, RTF-bestanden en hypertext-bestanden) te worden geëxporteerd.

Het ontwikkelwerk gebeurde in de oertaal voor objectgeoriënteerd programmeren: *Smalltalk*.

## **UvA: Vakgroep Latijn en Computercentrum Letteren (CCL), 1979-1995**

Het Computercentrum Letteren was een dienstverlenende afdeling voor de Faculteit der Letteren van de Universiteit van Amsterdam. Het stelde apparatuur en standaardsoftware beschikbaar en ontwikkelde waar nodig specifieke applicaties ten bate van letterenonderzoek. Ik hield mij vooral bezig met de volgende werkzaamheden:

- *Ontsluiting grote tekstcorpora*  
Met deze programmatuur, ontwikkeld door Ir. G.J. van der Steen, konden taal- en letterkundigen zoeken naar verschijnselen in corpora: grote teksten - taal en slecht gestructureerde data. Tevens ontwikkelden wij programmatuur om meer met die corpora te kunnen doen, zoals het aanbrengen van taalkundige codes.
- *Conversie- en validatiesystemen*  
De door mij en mijn CCL-collega's ontwikkelde parser-generator *Atlas* (voorheen: *Parspat*, tegenwoordig *AddXML*) was niet alleen geschikt voor controle en analyse, maar ook voor op regels gebaseerde conversie (transductie). Mijn eigen toepassing, in het kader van mijn afstuderen, is *Latinflexion*, een computergrammatica voor de Latijnse vormleer en een aansluitend semi-automatisch desambiguëringssysteem. Ontwikkeld in het kader van mijn doctoraalscriptie en later verkocht aan de Katholieke Universit t Eichst tt (Duitsland) en de Vakgroep Latijn van de Universiteit van Amsterdam.
- *Gestructureerd invoeren*  
Het ontwikkelen van *DictEdit*, een Macintosh-programma voor het consistent invoeren van woordenboeken, is van 1992 tot 1995 een van mijn belangrijkste bezigheden geweest bij het CCL. Op basis van een definitiebestand werden formulieren gepresenteerd in een Macintosh-venster. De door de gebruiker ingevoerde gegevens werden intern opgeslagen in SGML. Met behulp van een transformatiemechanisme (DSSSL was er toen nog niet, laat staan XSLT) kon het SGML worden omgezet naar woordenboekachtige uitvoer. DictEdit werd

gepresenteerd tijdens het Euralex-congres over Lexicologie te Tampere (Finland).

- *Fan*  
Fonologische/logopedische analyse van het Nederlands van taallerende jonge kinderen. Bij mijn vertrek heb ik de rechten van dit programma gekocht; het programma wordt nu, op zeer beperkte schaal, door mij en een ex-collega geëxploiteerd. Inmiddels werk ik aan een nieuwe versie, gebaseerd op Java Webstart.
- *CM/LTree*  
Dit is een portable cache-memory en indexeringslibrary (in C) voor de opslag en retrieval van grote lexicale bestanden. Ingezet bij diverse CCL-projecten, onder andere bij de door de Europese Unie gefinancierde woordenboekprojecten *Acquilex* en *Sift*. Versies zijn beschikbaar voor Dos/Windows, Macintosh en Unix.

### ***Werk-gerelateerde nevenactiviteiten***

- Lid van een inter-universitaire werkgroep voor software-ontwikkeling binnen de humaniora.
- Bestuurslid en penningmeester van de Stichting TDH (Tekstcorpora en Databestanden in de Humaniora).