Healthy CDDs

Strategies for building a Custom Debian Distribution

Andreas Tille

Debian Conference 8

Mar del Plata, August 12, 2008

Structure

1. History
2. Development
   - Measures for development
   - Growth of Metapackages
   - Activity of members
3. Comparison of CDDs
   - Graphing activity
   - Conclusion

Structure

1. History
2. Development
   - Measures for development
   - Growth of Metapackages
   - Activity of members
3. Comparison of CDDs
   - Graphing activity
   - Conclusion

Structure

1. History
2. Development
   - Measures for development
   - Growth of Metapackages
   - Activity of members
3. Comparison of CDDs
   - Graphing activity
   - Conclusion

Libre Software Meeting (DebConf 1) - Bordeaux 2001

There is a laptop open

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- → DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)

→ DebConf 3 Oslo “Custom Debian Distributions”

- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)

→ DebConf 3 Oslo “Custom Debian Distributions”

- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)
- ⟶ DebConf 3 Oslo “Custom Debian Distributions”
- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

Who are the others?

- *Debian Junior* (2000)
- *Debian Med* (2002)
- *Debian Edu (SkoleLinux)* (2002)

→ DebConf 3 Oslo “Custom Debian Distributions”

- *Debian* Accessibility (2003)
- *Debian* Desktop (2003)
- *Debian* Lex (2003)
- *Debian* Nonprofit (2003)
- *Debian GIS* (2004)
- *DebiChem* (2004)
- *Debian* Enterprise (2007)

1 History

2 Development
- Measures for development
- Growth of Metapackages
- Activity of members

3 Comparison of CDDs
- Graphing activity
- Conclusion

Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

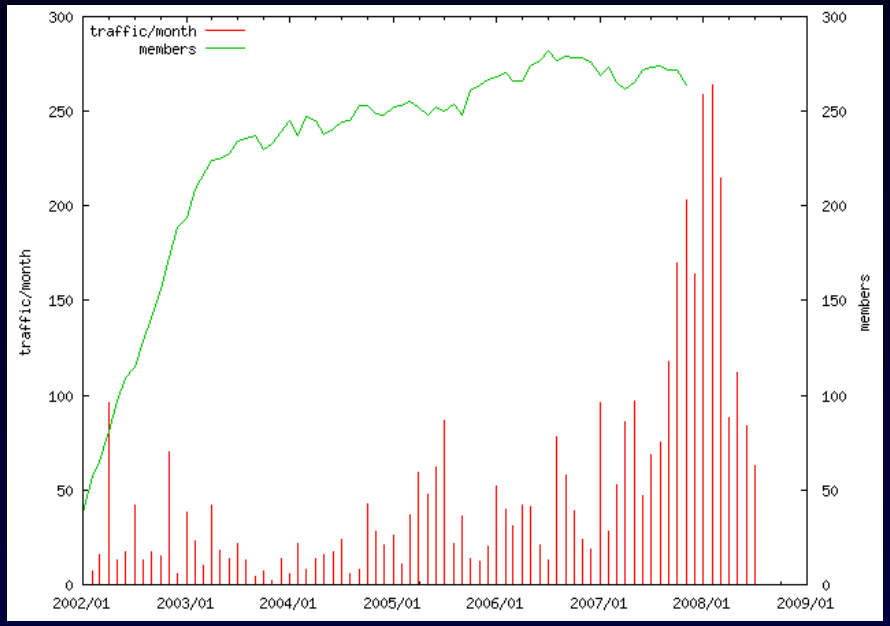
Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

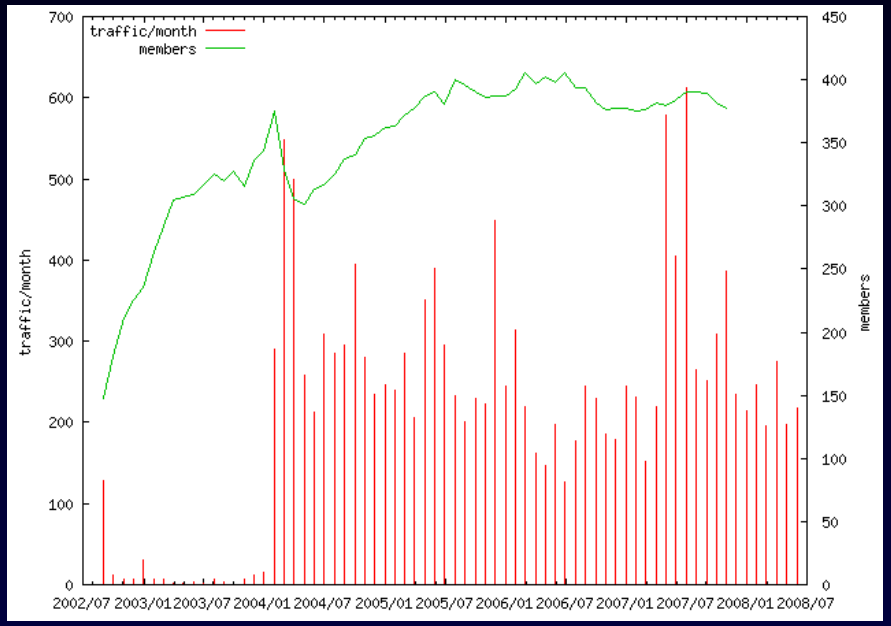
Considering some numbers ...

- Number of users?
  - Popularity contest?
  - Specific software has usually low popcon numbers
  - Usage of metapackages might give some hints but popcon works bad for these - the “content” is not actually used
- Number of developers
  - Number of Vcs commits $\longrightarrow$ not every CDD uses common Vcs
  - Number of mailing list subscribers $\longrightarrow$ people tend to stay subscribed

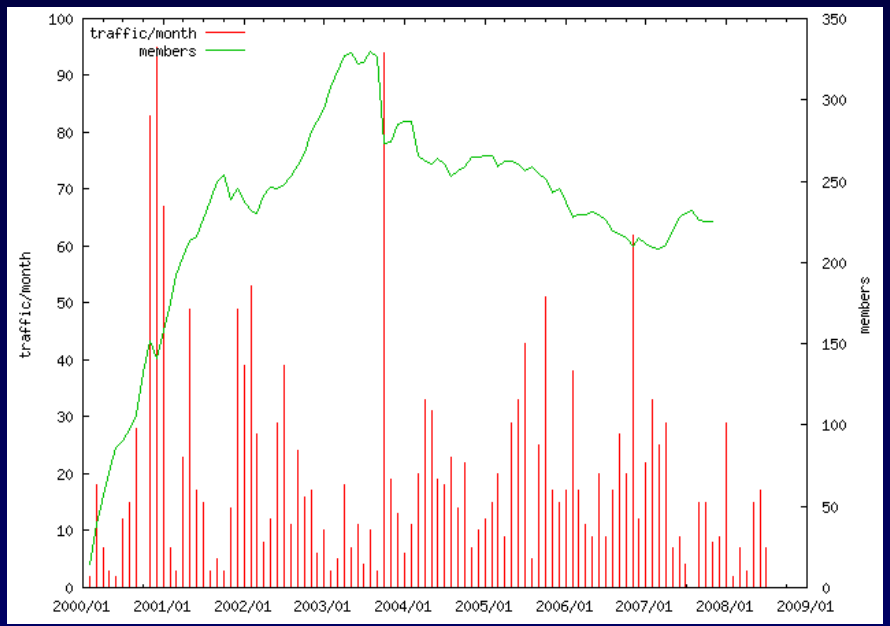
Debian Med



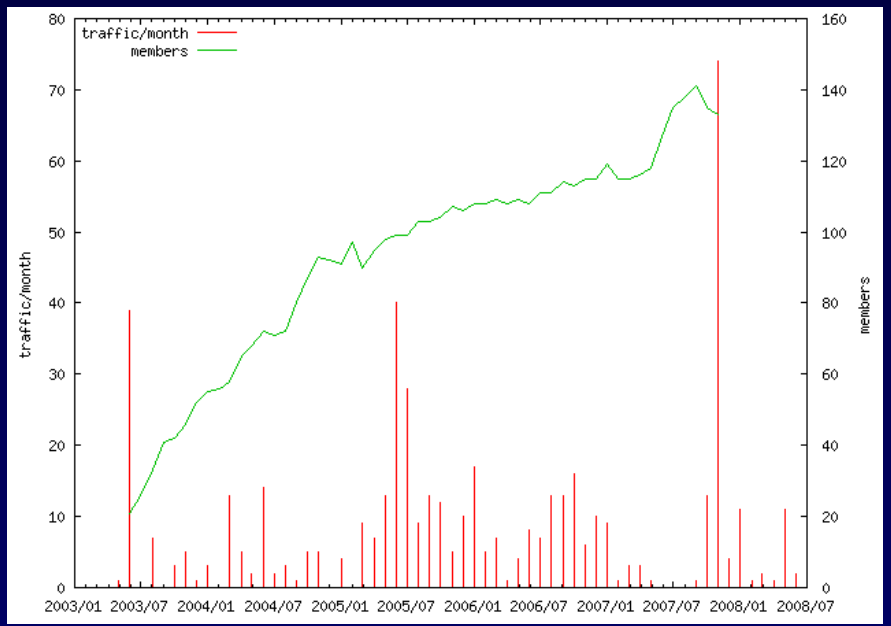
Debian Edu



Debian Junior



Debian Lex



Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

Signal, noise, artefacts

- Not only SPAM is noise
- (Un)subscribe
- Debian Junior list peak in October 2003:
  - Replacement for Abiword: LyX? Openoffice? → 43 postings
  - Philosophy (was Re: Replacement for Abiword: LyX? Openoffice?) → 33 postings
  - Future replacement for Abiword: Tux Writer? → 3 postings
  - 79 postings about specific topic out of 89 non-SPAM
- Intended and unintended robots
  - December 2005: 92 messages by `gentoo-\w+\+help@gentoo.org` about subscription status to *debian-desktop@lists.debian.org*
  - Messages by `Debian Installer`, `bugzilla-skolelinux`, `Archive Administrator`, ...

1 History

2 Development
- Measures for development
- Growth of Metapackages
- Activity of members

3 Comparison of CDDs
- Graphing activity
- Conclusion

Dependency of metapackages

- Technical output - number of metapackages and their dependencies
- Probably good measure
- Hard to obtain because not everything is in Vcs
- Querying *snapshot.debian.net* possible but really time consuming

Dependency of metapackages

- Technical output - number of metapackages and their dependencies
- Probably good measure
- Hard to obtain because not everything is in Vcs
- Querying *snapshot.debian.net* possible but really time consuming

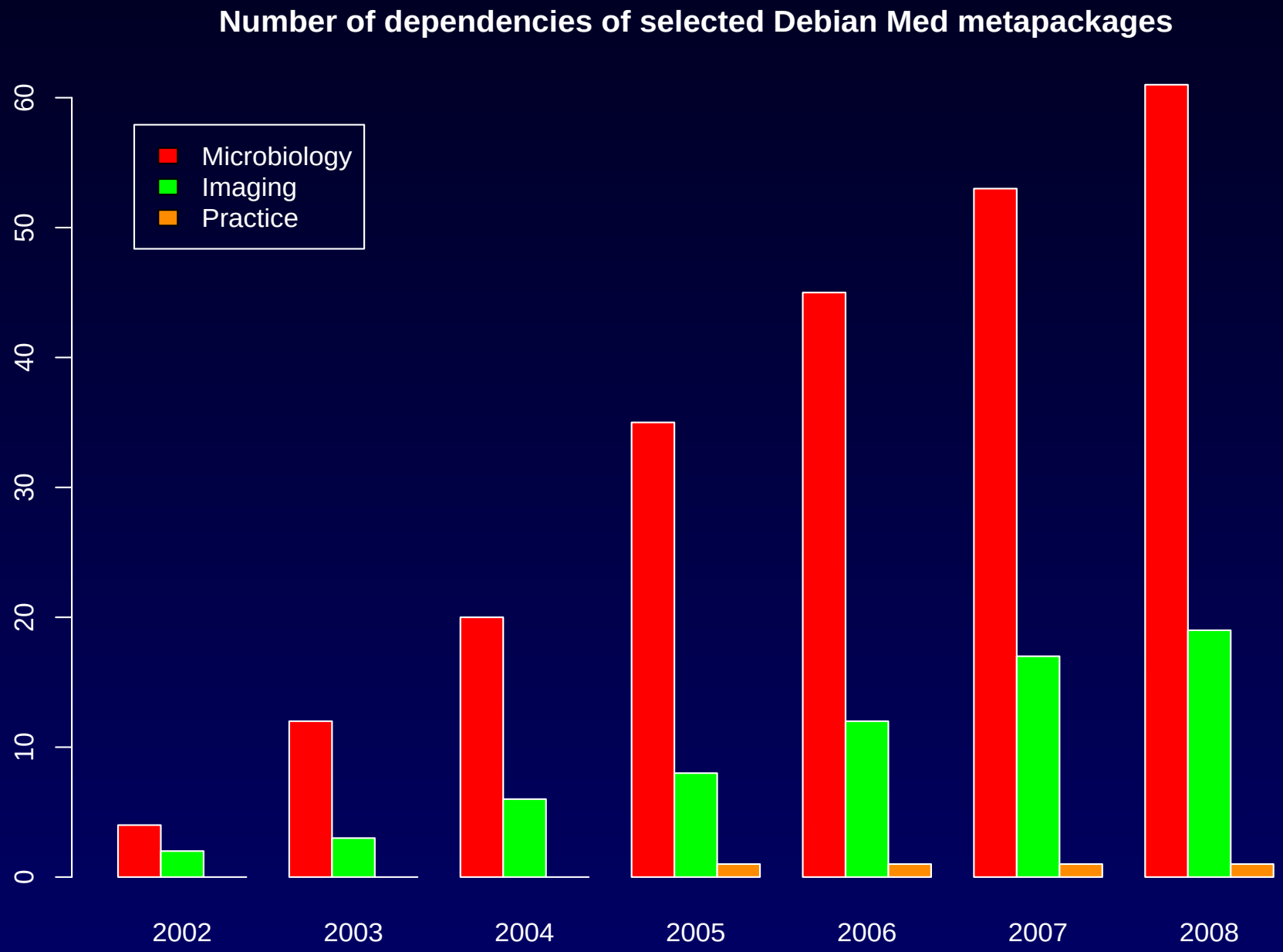
Dependency of metapackages

- Technical output - number of metapackages and their dependencies
- Probably good measure
- Hard to obtain because not everything is in Vcs
- Querying *snapshot.debian.net* possible but really time consuming

Dependency of metapackages

- Technical output - number of metapackages and their dependencies
- Probably good measure
- Hard to obtain because not everything is in Vcs
- Querying *snapshot.debian.net* possible but really time consuming

Selected metapackages of Debian Med

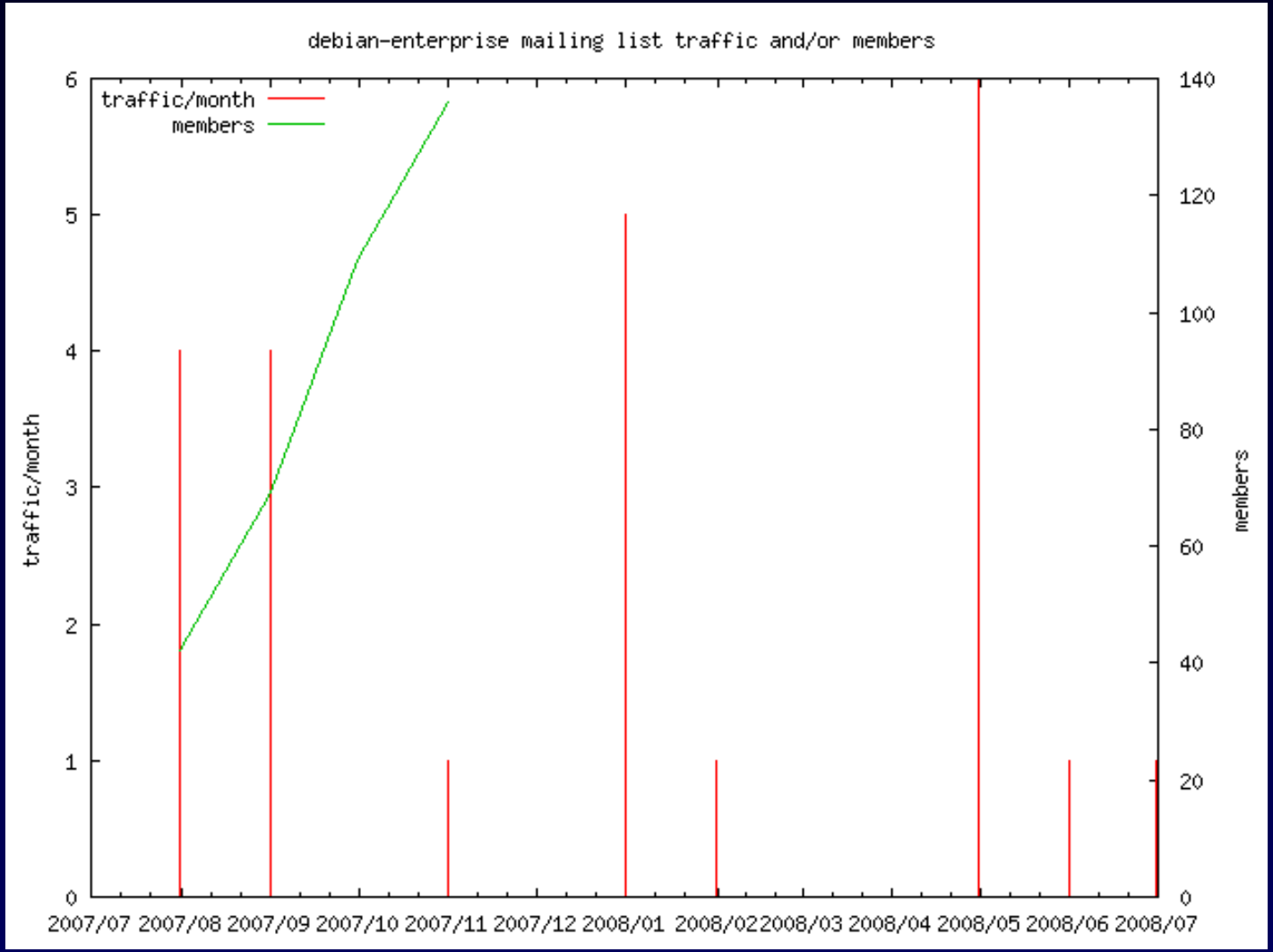


1 History

2 Development
- Measures for development
- Growth of Metapackages
- Activity of members

3 Comparison of CDDs
- Graphing activity
- Conclusion

Number of subscribers?



Activity of communication measure

- Intensity and quality of communication
- Every CDD has a mailing list
- Who are the active posters (except robots)
- Influence of SPAM, noise, flames, etc. is reduced
- Mind the run-over-by-bus factor

Activity of communication measure

- Intensity and quality of communication
- Every CDD has a mailing list
- Who are the active posters (except robots)
- Influence of SPAM, noise, flames, etc. is reduced
- Mind the run-over-by-bus factor

Activity of communication measure

- Intensity and quality of communication
- Every CDD has a mailing list
- Who are the active posters (except robots)
- Influence of SPAM, noise, flames, etc. is reduced
- Mind the run-over-by-bus factor

Activity of communication measure

- Intensity and quality of communication
- Every CDD has a mailing list
- Who are the active posters (except robots)
- Influence of SPAM, noise, flames, etc. is reduced
- Mind the run-over-by-bus factor

Activity of communication measure

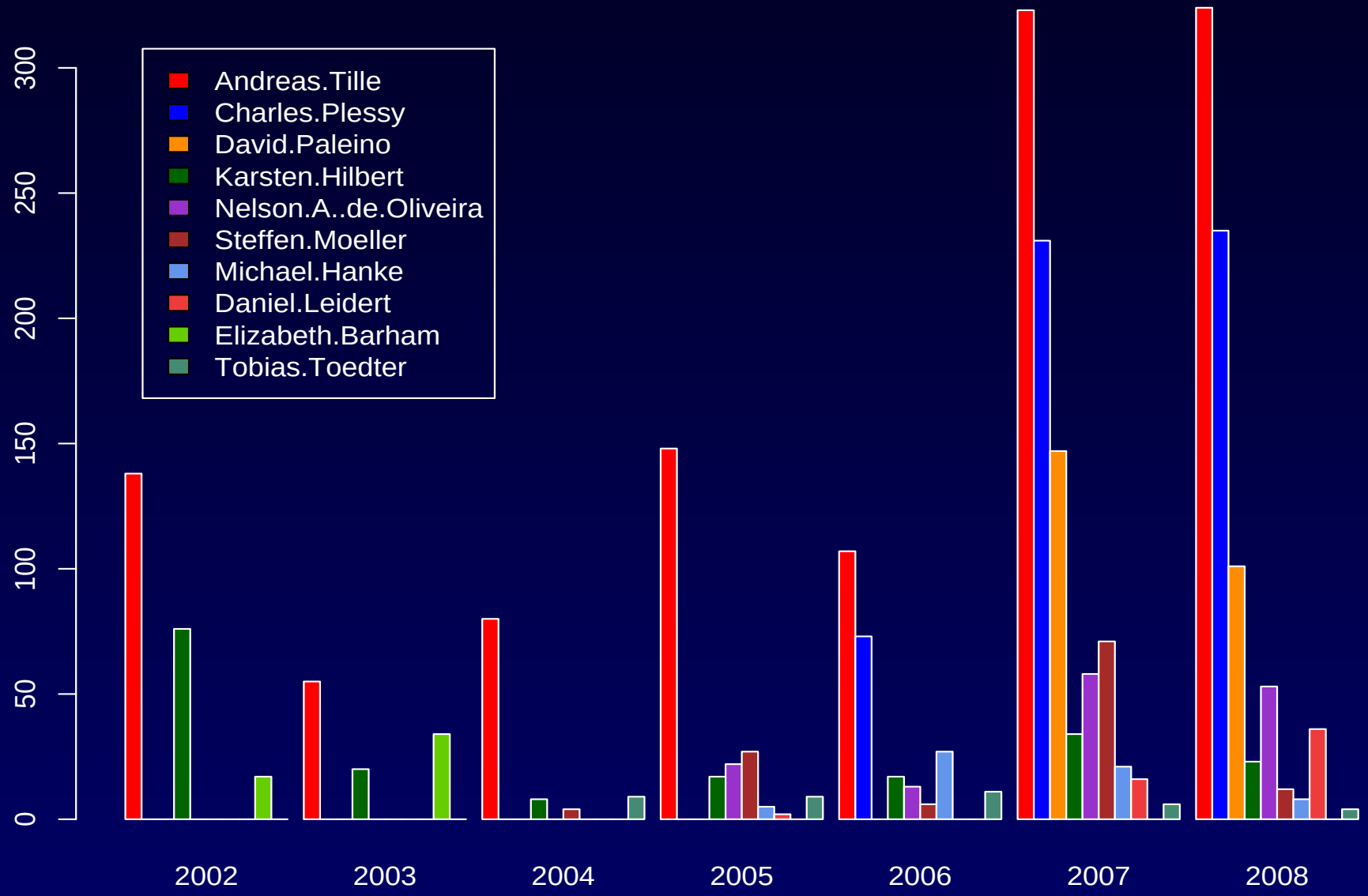
- Intensity and quality of communication
- Every CDD has a mailing list
- Who are the active posters (except robots)
- Influence of SPAM, noise, flames, etc. is reduced
- Mind the run-over-by-bus factor

1 History

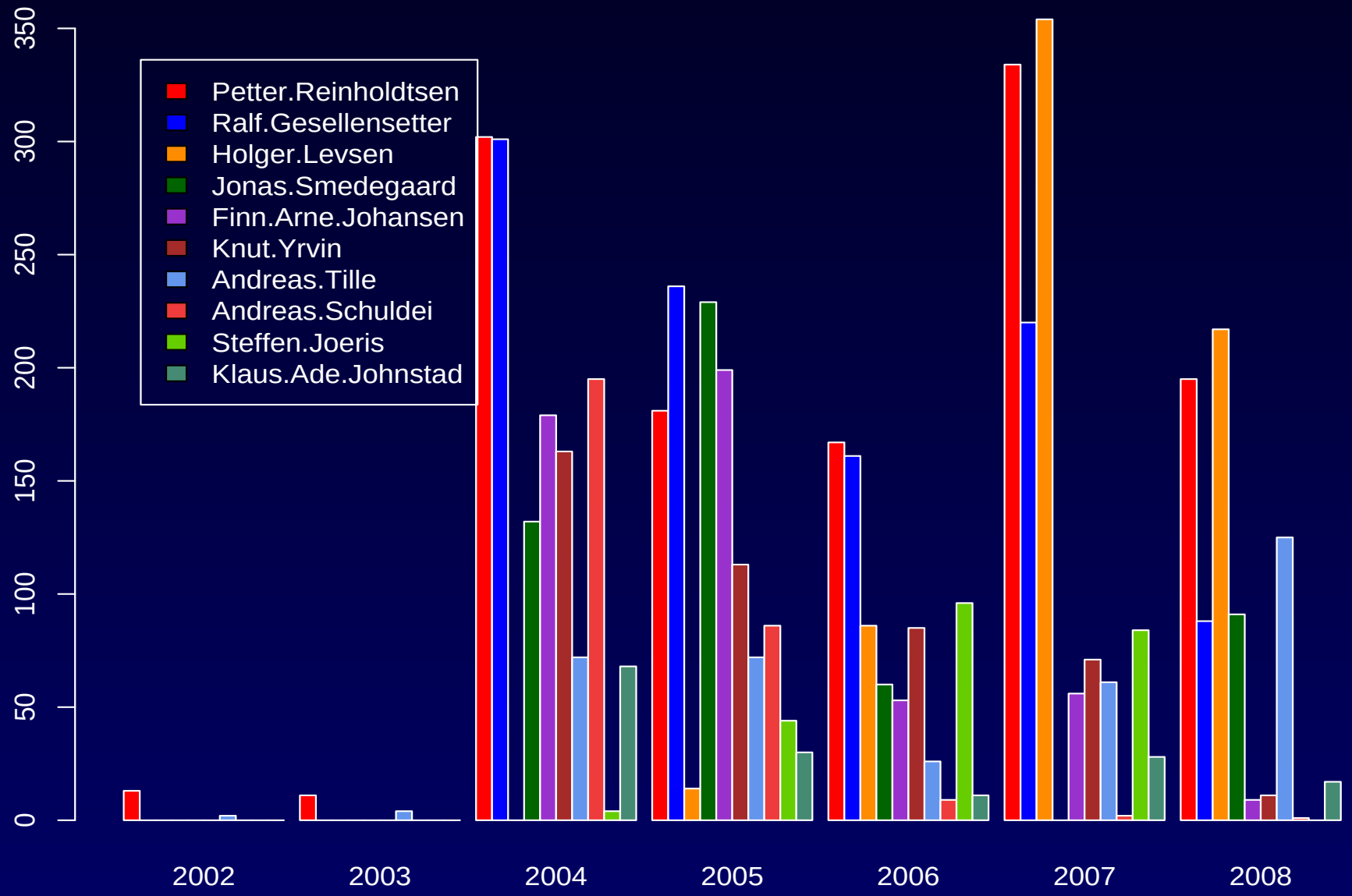
2 Development
- Measures for development
- Growth of Metapackages
- Activity of members

3 Comparison of CDDs
- Graphing activity
- Conclusion

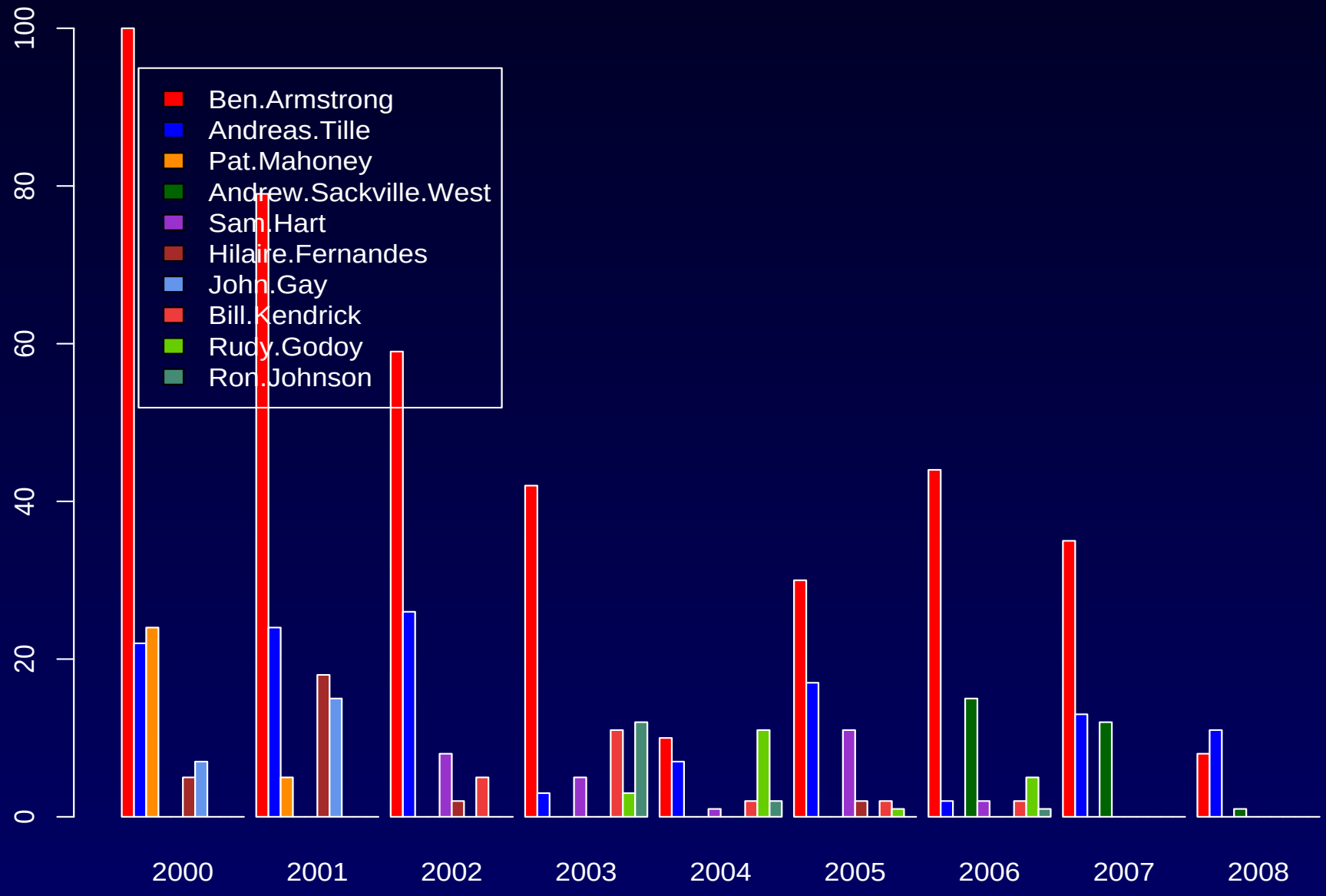
Top 10 posters on *debian-med@lists.debian.org*



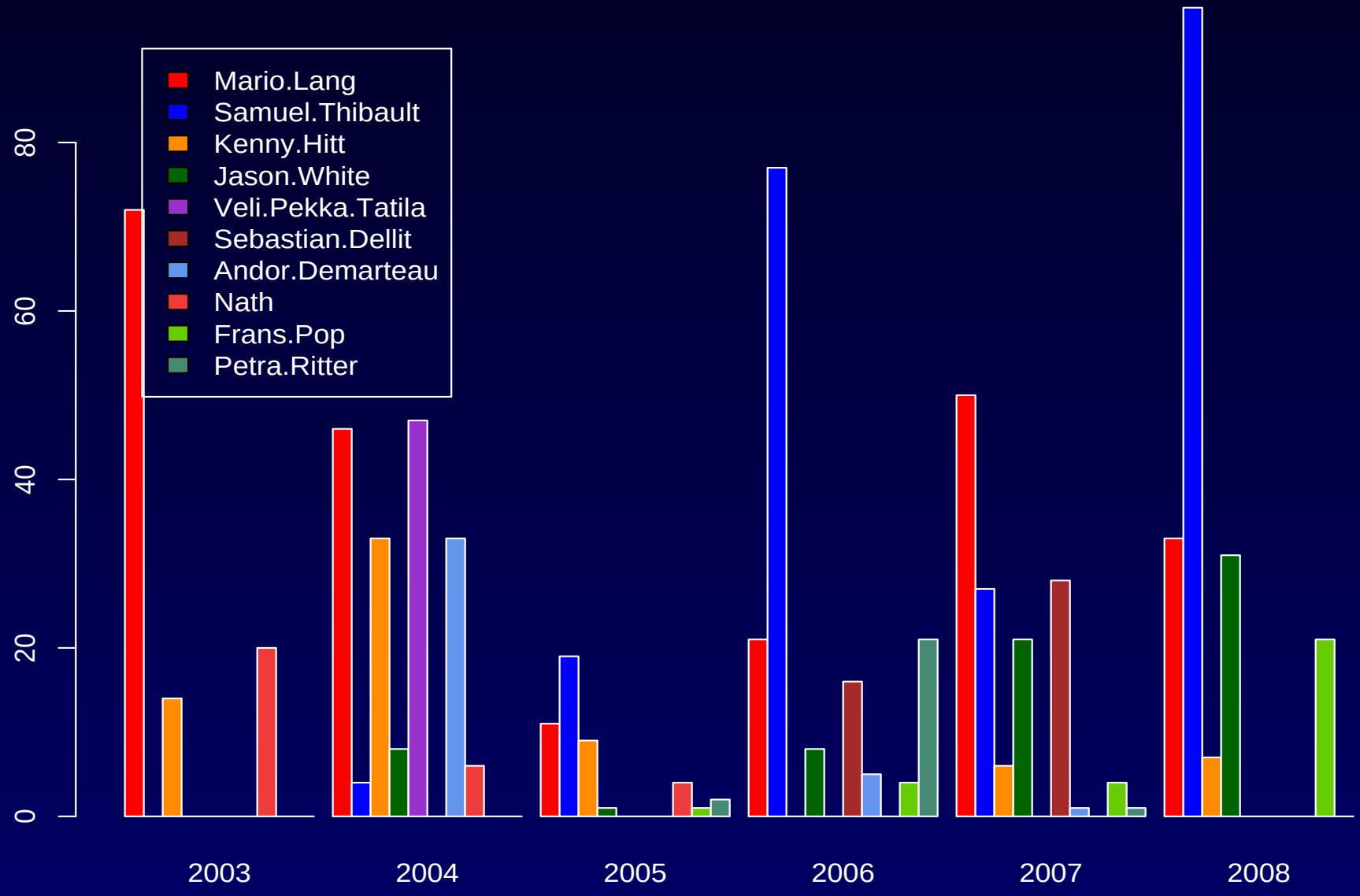
Top 10 posters on *debian-edu@lists.debian.org*



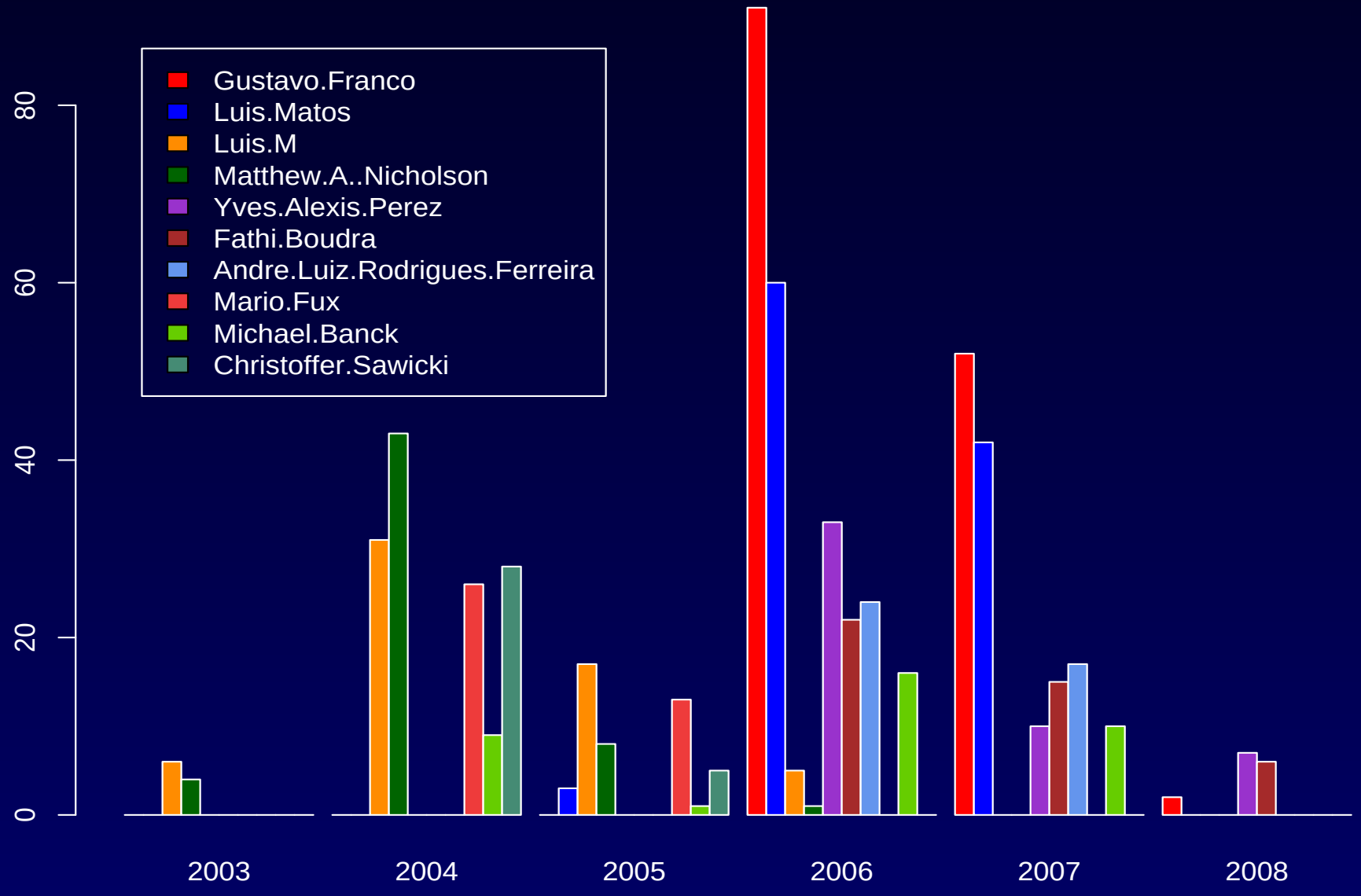
Top 10 posters on *debian-jr@lists.debian.org*



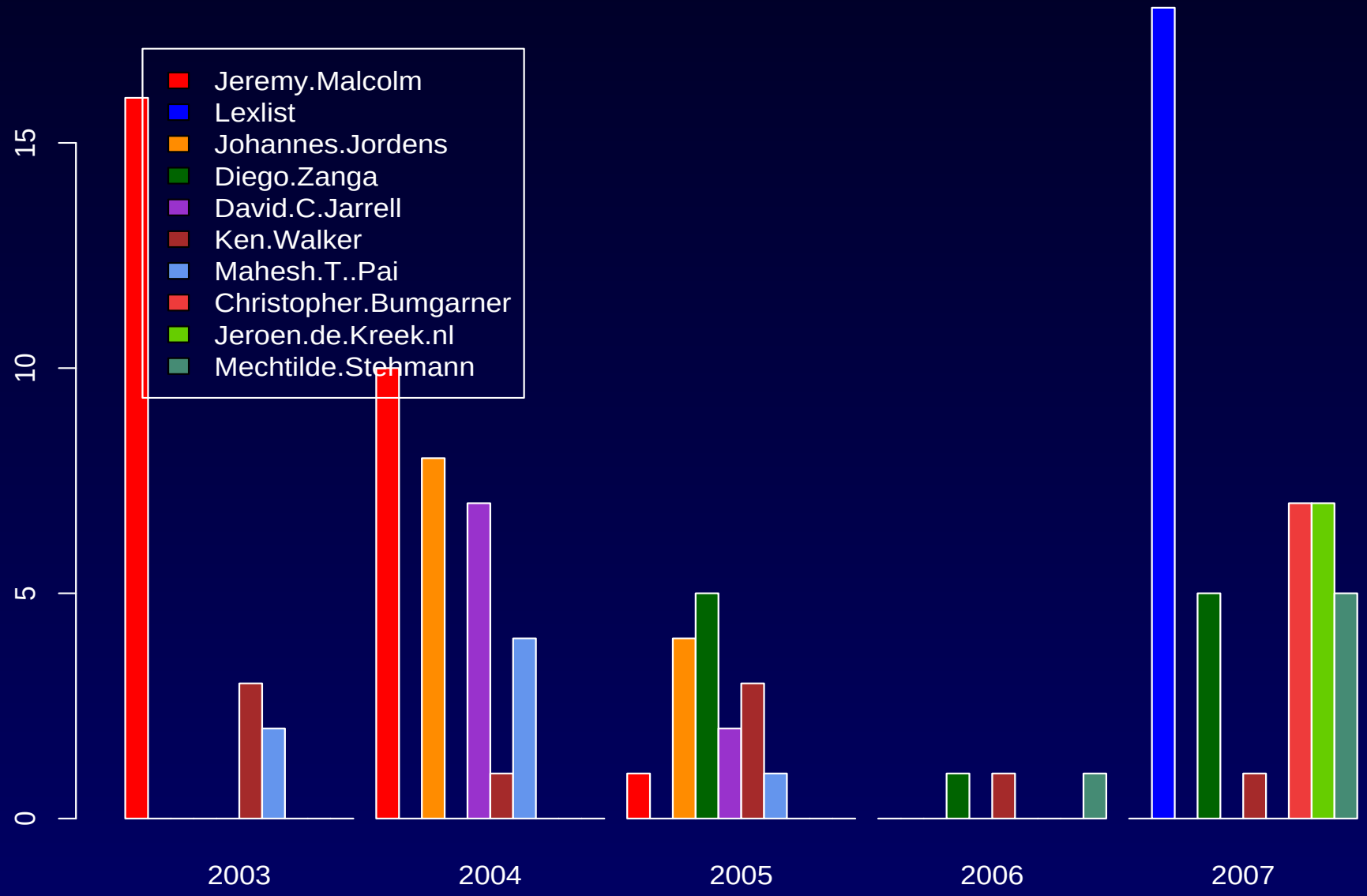
Top 10 posters on *debian-accessibility@lists.d.o*



Top 10 posters on *debian-desktop@lists.debian.org*



Top 10 posters on *debian-lex@lists.debian.org*



- Successfull SPAM protection: in 2008 this list gathered only SPAM
- Peak in November 2007:

Debian-Lex Rebirth?	23
Debian-Lex Rebirth? - python/perl/php?	25
Debian-Lex Rebirth? - wiki	8
Rebirth	1
Rebirth of Debian-lex	2

- Good example for “vanished leader”

- Successfull SPAM protection: in 2008 this list gathered only SPAM
- Peak in November 2007:

Debian-Lex Rebirth?	23
Debian-Lex Rebirth? - python/perl/php?	25
Debian-Lex Rebirth? - wiki	8
Rebirth	1
Rebirth of Debian-lex	2

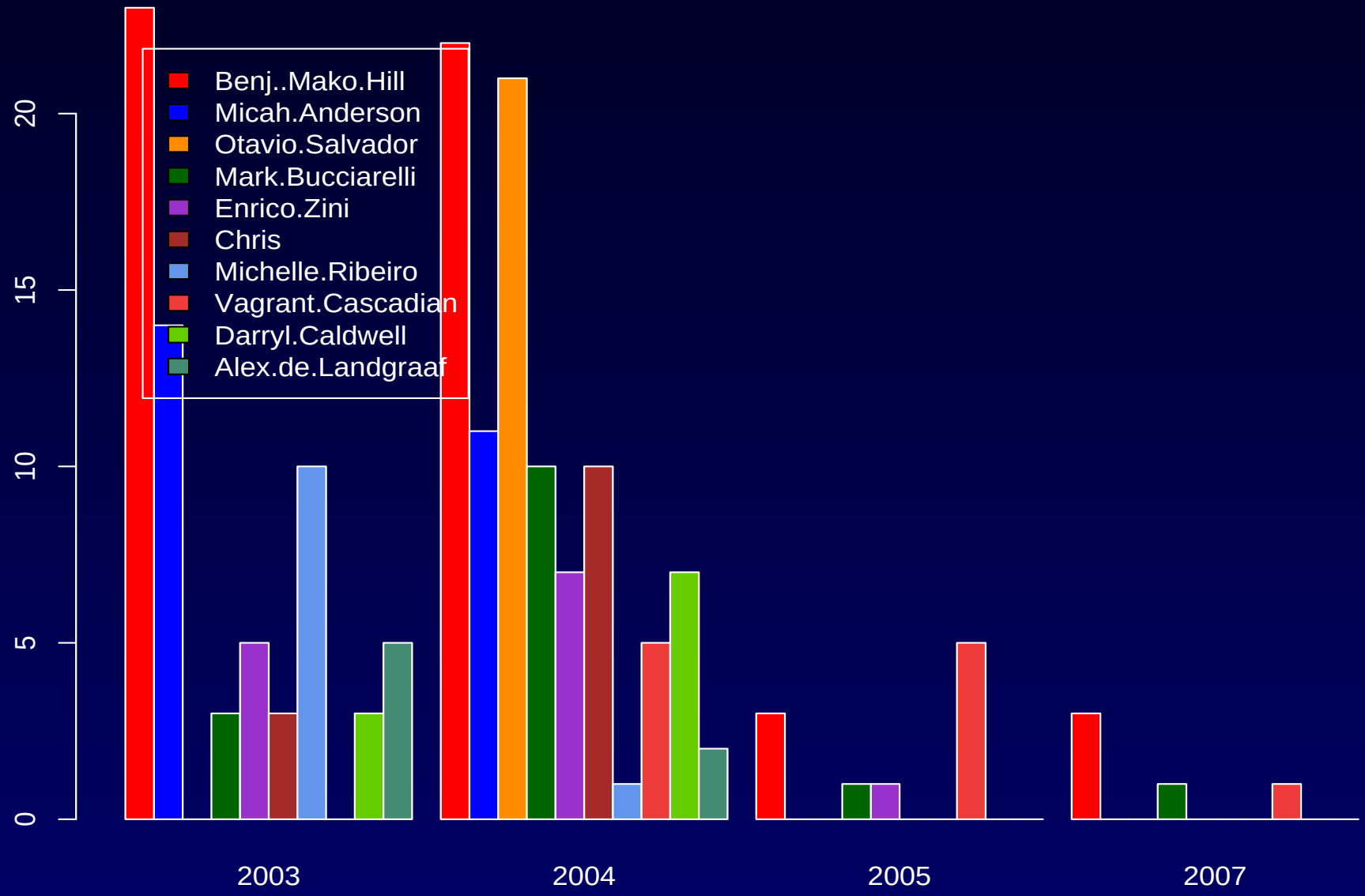
- Good example for “vanished leader”

- Successfull SPAM protection: in 2008 this list gathered only SPAM
- Peak in November 2007:

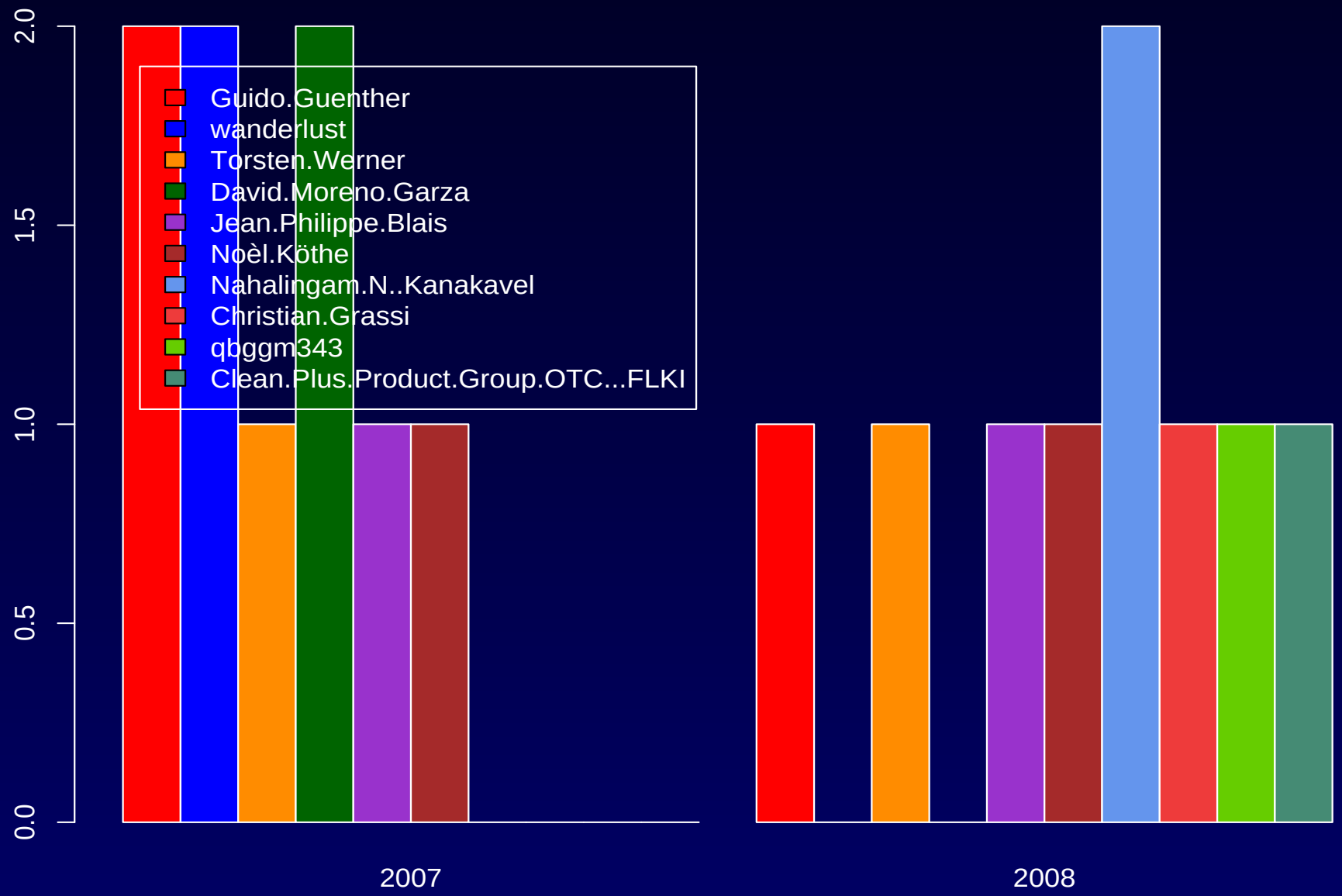
Debian-Lex Rebirth?	23
Debian-Lex Rebirth? - python/perl/php?	25
Debian-Lex Rebirth? - wiki	8
Rebirth	1
Rebirth of Debian-lex	2

- Good example for “vanished leader”

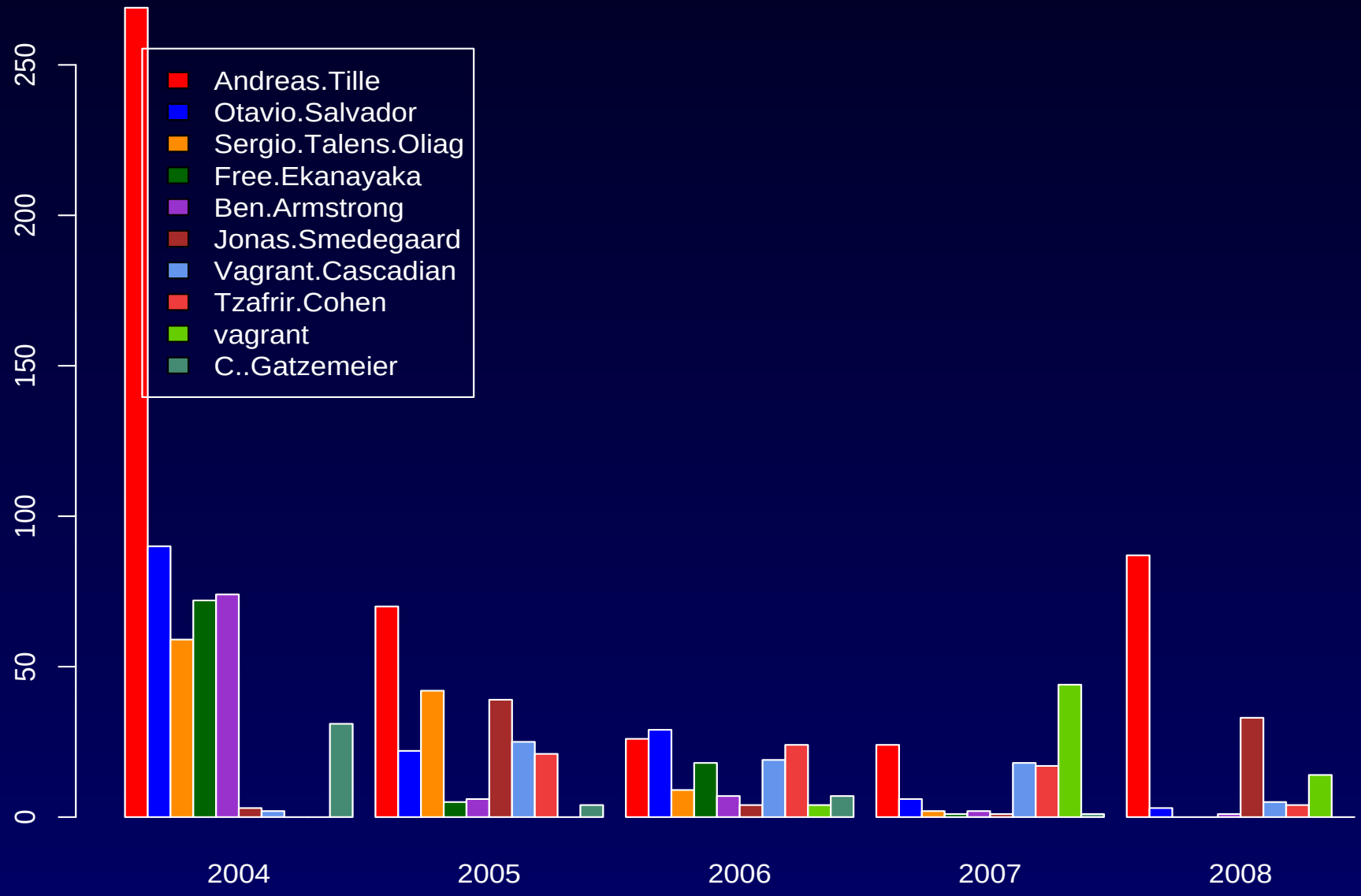
Top 10 posters on *debian-nonprofit@lists.debian.org*



Top 10 posters on *debian-enterprise@lists.debian.org*



Top 10 posters on *debian-custom@lists.debian.org*



1. History
2. Development
   - Measures for development
   - Growth of Metapackages
   - Activity of members
3. Comparison of CDDs
   - Graphing activity
   - Conclusion

Most CDDs depend from single person

- Debian Edu has a really healthy crew
- Debian Med has at least run-over-by-bus factor two
- Others seem to depend from single person
- Need of better management

Most CDDs depend from single person

- Debian Edu has a really healthy crew
- Debian Med has at least run-over-by-bus factor two
- Others seem to depend from single person
- Need of better management

Most CDDs depend from single person

- Debian Edu has a really healthy crew
- Debian Med has at least run-over-by-bus factor two
- Others seem to depend from single person
- Need of better management

Most CDDs depend from single person

- Debian Edu has a really healthy crew
- Debian Med has at least run-over-by-bus factor two
- Others seem to depend from single person
- Need of better management

This talk is available at
http://people.debian.org/~ tille/talks/
Andreas Tille <tille@debian.org>